DESIGN FOR YIELD FOR SUB-22nm FinFET-BASED FPGA

By

Mohamed Mohie El-Din Mohamed Aly Hassan

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Electronics and Electrical Communications Engineering

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

DESIGN FOR YIELD FOR SUB-22nm FinFET-BASED FPGA

By

Mohamed Mohie El-Din Mohamed Aly Hassan

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Electronics and Electrical Communications Engineering

Under the Supervision of

Prof. Dr. Hossam A. H. Fahmy
Professor
Elec. & Comm. Dept.
Faculty of Engineering, Cairo University

Dr. Hassan Mostafa
Assistant Professor
Elec. & Comm. Dept.
Faculty of Engineering, Some University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

DESIGN FOR YIELD FOR SUB-22nm FinFET-BASED FPGA

By
Mohamed Mohie El-Din Mohamed Aly Hassan

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Electronics and Electrical Communications Engineering

Approved by the
Examining Committee

_____
Prof. Dr. First S. Name, External Examiner

_____
Prof. Dr. Second E. Name, Internal Examiner

_____
Prof. Dr. Hossam A. H. Fahmy, Thesis Main Advisor

_____
Dr. Hassan Mostafa, Thesis Advisor

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

Engineer's Name: Mohamed Mohie El-Din Mohamed
Date of Birth: 08/03/1990
Nationality: Egyptian
E-mail: Mohamed.mohie.hassan@gmail.com
Phone: +201068663016
Address: 9 St., El-Maadi, Cairo, Egypt

Registration Date: 17/09/2011
Awarding Date: …./…./2016
Degree: Master of Science
Department: Electronics and Electrical Communication

Supervisors:

Prof. Dr. Hossam A. H. Fahmy
Dr. Hassan Mostafa

Examiners:

Prof. …… (External examiner)
Prof. …… (Internal examiner)
Prof. Dr. Hossam A. H. Fahmy
Dr. Hassan Mostafa

Title of Thesis:

Design For Yield For Sub-22nm FinFET-Based FPGA

Key Words:
Design for Yield; Process Variations; FinFET; FPGA; Technology Scaling; Leakage Power

Summary:

In this thesis, a performance evaluation study for a FinFET-Based FPGA cluster under threshold voltage variation, representing the Die-to-Die variations, is launched with technology scaling starting from 20nm down to 7nm nodes showing the scaling trends of various performance metrics including the average power, delay, and power-delay product. Also some design insights and recommendations are proposed for the designers to achieve yield percentage of 99.87%. The leakage power is also studied for 14nm technology node under threshold voltage and temperature variations. Some solutions are implemented for leakage power control under threshold voltage variations including transistor stacking, minimum leakage vector, and gate sizing.

Acknowledgments

I would like to express my utmost gratitude to Allah for giving me the strength to complete the work.

I would also like to sincerely thank my supervisors, Dr. Hossam Fahmy and Dr. Hassan Mostafa, for their continuous support and guidance throughout my work.

Table of Contents

List of Tables

List of Figures

Nomenclature

ADE	Analog Design Environment
ASIC	Application-Specific Integrated Circuits
BLE	Basic Logic Element
BSIM	Berkley Short-channel IGFET Model
CAD	Computer-Aided Design
CMG	Common Multi-Gate
CMOS	Complementary Metal-Oxide Semiconductor
DIBL	Drain-Induced Barrier Lowering
DLL	Delay-Locked Loop
D2D	Die-to-Die
FinFET	Fin Field Effect Transistor
FPGA	Field-Programmable Gate Array
GIDL	Gate-Induced Drain Leakage
IGFET	Independent-Gate FET
LER	Line-Edge Roughness
LSTP	Low Standby Power Devices
LUT	Look-Up Table
NWE	Narrow Width Effects
OPE	Optical Proximity Effects
PDF	Probability Density Function
PDP	Power-Delay Product
PLL	Phase-Locked Loop
RDF	Random Dopant Fluctuations
SCE	Short Channel Effects
SGFET	Shorted-Gate FET
SoPC	System-on-a-Programmable-Chip
SRAM	Static Random Access Memory
SSTA	Statistical Static Timing Analysis
WID	Within-Die

Abstract

As CMOS technology is scaled towards the deep sub-micron regime, digital circuits' designers are facing increased variability in form of either process variations or environmental variations. Those variations are classified to Die-to-Die variations and Within-Die variations. Our work presented in this thesis aims at evaluating the performance of a FinFET-Based FPGA cluster under threshold voltage variation that represents the Die-to-Die variations with technology scaling starting from 20nm down to 7nm nodes using Berkley Predictive Technology Models, showing the scaling trends of different performance metrics including the average power, delay, and power-delay product. Some design insights and recommendations are proposed for the designers to achieve yield percentage of 99.87%.

Since the leakage power is much more pronounced in advanced technology nodes, we also studied the leakage power and its variation for 14nm technology node under threshold voltage and temperature variations. The results emphasized the log-normal dependency of the leakage power on the threshold voltage and also the exponential dependency with temperature. Some solutions are proposed and implemented for leakage power control under threshold voltage variation including transistor stacking, minimum leakage vector (MLV), and gate sizing. These solutions have shown improvements for both the leakage power and leakage power variation, but also these solutions introduced a minor delay and area overheads which are reported and compared as well.

The FPGA cluster we built for our study is configured to a 2-bit adder benchmark used for both the performance evaluation study with technology scaling and the leakage power evaluation study as well. Cadence Virtuoso and ADE-GXL are used for both FPGA cluster building and simulations respectively.

Chapter 1 : Introduction

This chapter presents a short introduction about the significance of studying the effect of process variations while scaling down CMOS technologies coming to different FinFET technologies. Section 1.1 presents the motive behind this research work. Section 1.2 provides the thesis outline and organization.

1.1. Motivation

More than four decades of scaling CMOS technology has been the biggest driver for electronics industry. CMOS transistor scaling allowed building chips with billions of transistors in modern Integrated Circuits (ICs) and a broad range of electronic products with very high integration levels [1]. Nevertheless, the continued rigorous scaling of CMOS technology in sub-100nm regime has created massive design challenges, especially when transitioning to advanced technology nodes starting from 22nm technology. Assignable to process control limitations, manufacturing tolerances in process technology are not equally scaling as the transistor channel length [2-5]. Furthermore, process variations due to fundamental physical limits such as Line-End Roughness (LER) and Random Dopant Fluctuations (RDF) are significantly increasing with technology scaling [2, 6-11]. Consequently, statistical parameter variations are getting worse with consecutive technology generations, and variability is currently becoming one of the biggest challenges that face the semiconductor industry, resulting in huge yield losses [5]. That variability has been affecting analog design for some considerable time, and currently it is significantly impacting digital design at nanometer technology nodes. In addition, scaling the threshold voltage in nanometer regime posed a massive increase in the sub-threshold leakage current due to the exponential dependence, hence affecting the power efficiency which is becoming the key to sustaining continually enhanced performance for future VLSI circuits,

1.2. Organization of the Thesis

In order to have an overview about the process variations effects on the next generation FPGAs incorporating FinFETs devices in the manufacturing process, it is important to study the variations effects on the basic performance metrics such as the average power, delay, and power-delay product.
An introduction about the process variations in general, their sources and impacts on the digital circuits, is illustrated in Chapter 2 along with an introduction about FPGAs, their structures, and the FPGA cluster we used for our study.

In Chapter 3, we study the impact of threshold voltage variation with, representing D2D variations, with FinFET technology scaling from 20nm down to 7nm over the FPGA cluster we built for our evaluation study. Also some design insights are presented in this chapter that help achieving a yield percentage of 99.87%.

Since the leakage power is becoming much more pronounced with the continued technology scaling, Chapter 4 presents the work done in order to study the leakage power variation with both the threshold voltage and temperature on the FPGA cluster. Also some solutions are proposed and implemented in order to control the leakage power for 14nm node.

And finally, the conclusion and the potential future work are drawn in Chapter 5.

Appendix A illustrates the PTM models we used in our simulation in detail.

Chapter 2 : Literature Review

This chapter is organized as follows: Section 2.1 discusses the variability sources and classifications, the impact of variations on the frequency and power, and some variations mitigation techniques. Section 2.2 discusses the FPGA as the platform of our study, the FPGA architecture in terms of both the logic and routing resources, and also the structure and specifications of the FPGA cluster we used in this research work.

2.1. Variability

Variations in integrated circuits are basically the deviations from the intended or designed values for a structure or circuit parameter in concern. Usually, the variations are caused by two different sources: physical factors and environmental factors. The physical factors cause a permanent variation in device parameters and they are generally caused by the lack of exact controls and statistical variations during the fabrication process [12]. Regarding the environmental factors, they cause variations in the circuit operation while the circuit is functioning, and they include variations in the power supply and temperature.

2.1.1. Classification of variations

The variations sources can generally be categorized into two classes [4, 9, 12, 13]:

2.1.1.1. Die-to-Die (D2D) Variations

The D2D variations are also called inter-die or global variations. They are variations from die to die and, in the same way, they affect all devices on the same chip (for example, they may cause all the transistors' threshold voltages or gate lengths' on the same die to deviate from their nominal values by the same amount). These variations are generally independent and hence, they can be represented by a single value for each die. These variations are generally assumed to have a Gaussian distribution with a given variance [12], and they represent a shift in the parameter's mean from its nominal value. D2D variations in a single process parameter are dealt with using corner-based models which assume that all the devices on a given design sample have a value that is shifted from the mean by a fixed amount [12].

2.1.1.2. Within-Die (WID) Variations

The WID variations are also called intra-die or local variations. These variations cause transistor parameters to vary across different devices within the same die (for example, some devices may have larger channel length than the rest of the devices on the same die). Thus, each device on a die requires a separate random

variable that represents its WID variations. WID variations can be subdivided further into two classes: [13].

1- Random Variations

They are spatially uncorrelated variations which result from statistical quantization effects, such as RDF and LER. The impact of these random variations is expected to be worse with process parameters scaling, and they can be characterized by their statistical distribution. The random variations' impact can be alleviated by means of increasing the logic depth because of the averaging effect. Unfortunately, the trend to boost the clock frequency of a design using aggressive pipelining has resulted in smaller logic depth which, hence, increases the impact of this type of variations.

2- Systematic Variations

These variations are usually caused by physical phenomena such as distortions in the lens and some elements in the lithographic system. This type of variations is quite complicated to be modeled; therefore, they are usually modeled as random variations with certain value of spatial correlation. All the variations that are layout-dependent, such as channel width and length variations, are considered systematic variations as well.

2.1.2. Sources of variability

2.1.2.1. Process Variations (Static Variations)

Process variations impact the device structure and, thus, they alter the circuits' electrical properties. The process variations' sources can be outlined as follows:

1- Random Dopants Fluctuations (RDF)

With CMOS technology scaling, the number of doping impurities in the channel depletion layer decreases, especially with minimum geometry devices. The atomicity of the dopants in the channel does not allow a constant concentration of dopants to appear across the channel as shown in Figure 2.1. Thus, it is very unlikely to have two neighboring transistors with the same number and placement of dopants. This random number and placement of the dopants cause uncertainty in the transistor threshold voltage, V_{th}. The statistical distribution of V_{th} due to RDF is found to follow a normal distribution [14, 15]. The standard deviation of V_{th} distribution due to RDF is modeled as [16, 17]:

$$\sigma_{V_{th}} = \sqrt[4]{4q^3\epsilon_{Si}N_a\varphi_F}\ \frac{T_{ox}}{\epsilon_{ox}}\ \frac{1}{\sqrt{3W*L}} \tag{2.1}$$

where q is the electron charge, ϵ_{ox} and ϵ_{Si} are the dielectric constants of the gate oxide and silicon respectively, Na is the channel dopant concentration, φ_F is the difference between intrinsic level and Fermi level,

T_{ox} is the gate oxide thickness, and L and W are the channel length and width respectively. Equation (2.1) shows that $\sigma_{V_{th}}$ is inversely proportional to the square root of the active device area. Thus, sizing up the transistors can help mitigating these variations, which is one of the most commonly used techniques in analog circuit design to decrease transistors mismatch [18]. Moreover, for SRAM cells which typically have minimum size devices, $\sigma_{V_{th}}$ will be the largest. Figure 2.2 shows that with technology scaling, the RDF variations ($\sigma_{V_{th}}$) experience a large increase which may reach up to 50% of V_{th} in advanced technology nodes, causing a large spread in performance and power.

2- Channel Length Variations

The patterning of design features with smaller dimensions than the light wavelength, used in optical lithography, results in distortions due to light diffraction, which is called Optical Proximity Effects (OPEs) [4, 12]. These effects are expected to worsen with technology scaling as the light wavelength is not scaling with the same pace as the device feature size, as shown in Figure 2.3. The OPEs will make it very challenging to print precise patterns on the Silicon wafer with technology scaling [19], making the lithography at these small feature sizes very challenging. OPEs are layout dependent, resulting in different Critical Dimension (CD) variations depending on neighboring lines and the orientation [13]. Controlling these variations has become very challenging in current technologies, and is expected to increase for future technology nodes [9].

The variation in transistor's channel length has a direct effect on the transistor electrical parameters; however, the most impacted parameters are the transistor's threshold voltage V_{th} [15, 20, 21]. This is due to the exponential dependence of V_{th} on channel length L for short channel devices, specifically due to Drain Induced Barrier Lowering (DIBL) effect. DIBL causes V_{th} to be substantially dependent on L as shown in Figure 2.4. This dependence can be modeled as [15, 20, 21]:

$$V_{th} \approx V_{tho} - (\xi + \eta V_{DS})\exp(\frac{L}{L_{to}}) \quad (2.2)$$

where V_{tho} is the long channel threshold voltage, ξ is the charge sharing coefficient, L_{to} is the characteristic length, and η is the DIBL coefficient.

Accordingly, a slight variation in L will introduce huge variation in V_{th} as shown in Figure 2.4.

3- Line-Edge Roughness (LER)

LER refers to the roughness introduced on the channel edge during the gate patterning, as shown in Figure 2.1, which contributes to the threshold voltage variations. Previously, the amount of this introduced LER was insignificant compared to the dimensions of the transistor channel (on the

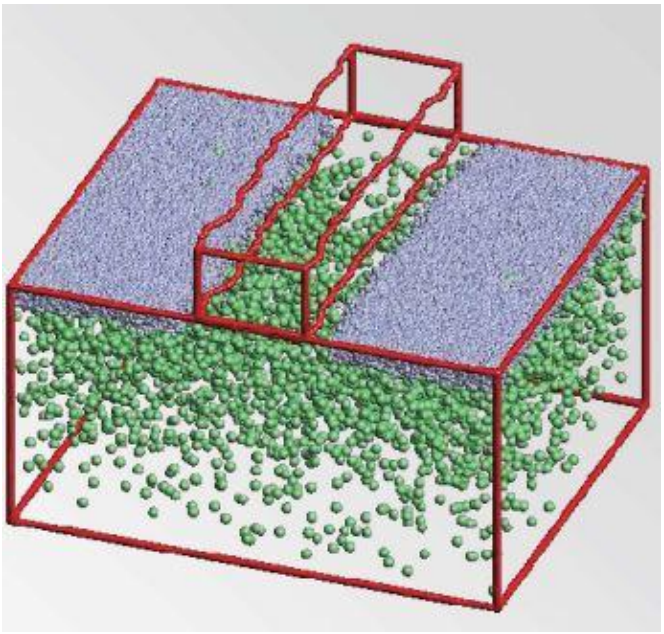


Figure 2.1: Atomistic process simulation incorporating RDF and LER as the sources of intrinsic fluctuations [1]. The green dots indicate the dopant atoms which determine the device's threshold voltage, while the blue dots indicate the drain/source doping

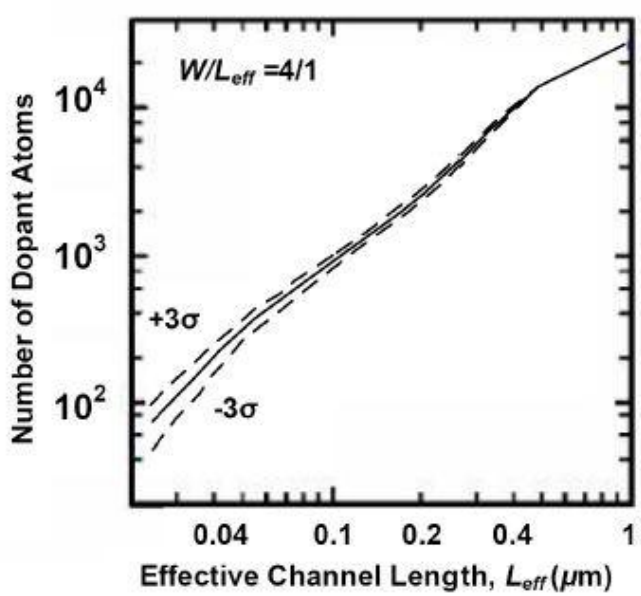


Figure 2.2: Number of dopant atoms in the depletion layer of a MOSFET versus channel length Leff

order of 5 nm) and also much smaller than the CD variations. However, with the continued transistor scaling, that introduced roughness becomes a significant source of variations in transistor's characteristics since it does not scale correspondingly [22]. These random effects cause variations in transistor's threshold voltage. Figure 2.5 shows the predicted variations in threshold voltage due to both LER and RDF versus technology nodes [1, 8] showing that for sub-32 technology nodes, threshold voltage variations due to LER and RDF will be comparable.

4- Gate Oxide Thickness Variations

Variations in the oxide thickness T_{ox} affect many electrical parameters of the device, especially the transistor threshold voltage V_{th}. Therefore, the T_{ox} variations should be considered.

5- Channel Width Variations

Transistor channel width (W) will have variations as well due to the lithography limitations. These channel width variations will contribute to V_{th} variations due to the Narrow-Width-Effects (NWEs), which cause V_{th} to be dependent on W. However, the impact of W variation on V_{th} can be considered very minimal compared to the impact due to L variations since W is typically 3-4 times larger than L [23].

2.1.2.2. Environmental Variations (Dynamic Variations)

Environmental variations impact the circuit operation while the circuit is functioning. They include variations in both the supply voltage and the temperature of the chip or across the chip [6, 9]. Variations in power supply are mainly due to the switching activity variations within the die that are dependent on the input vectors. A reduced power supply lowers the drive strength of the transistors and, consequently, causes performance degradation [13]. This reduction in the power supply will be problematic with technology scaling since the headroom between the supply voltage and the device's threshold voltage is being reduced consistently [24]. WID temperature variations are considered one of the major performance and packaging challenges as both device and interconnect exhibit temperature dependence that results in performance degradation at higher temperature. Moreover, temperature variations across different blocks communicating on the same die may result in performance mismatch, which may lead to functional failures [4]. Figure 2.6 shows WID temperature fluctuations for a microprocessor unit, with the core exhibiting a hot spot of 120℃ [25].

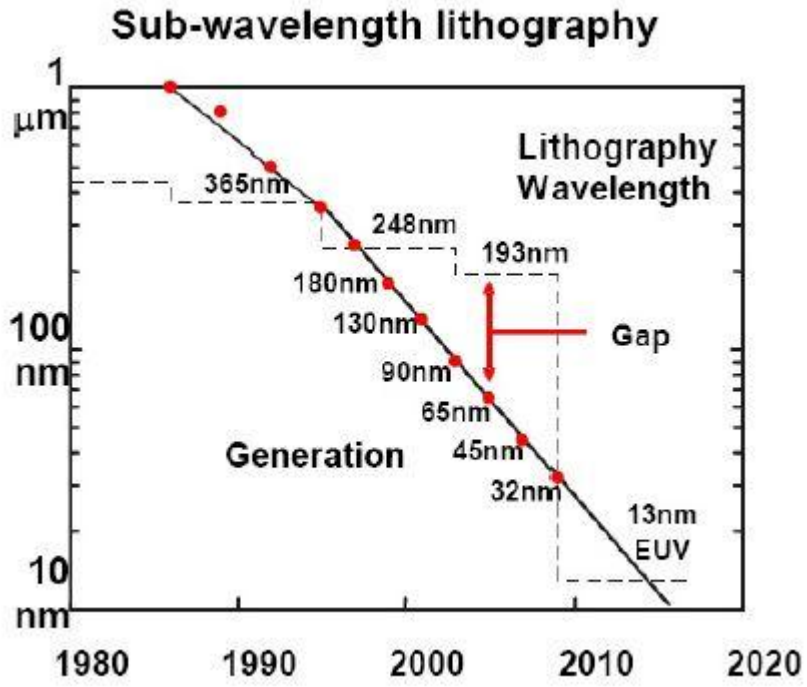


Figure 2.3: Lithography wavelength scaling for different technology nodes [6]

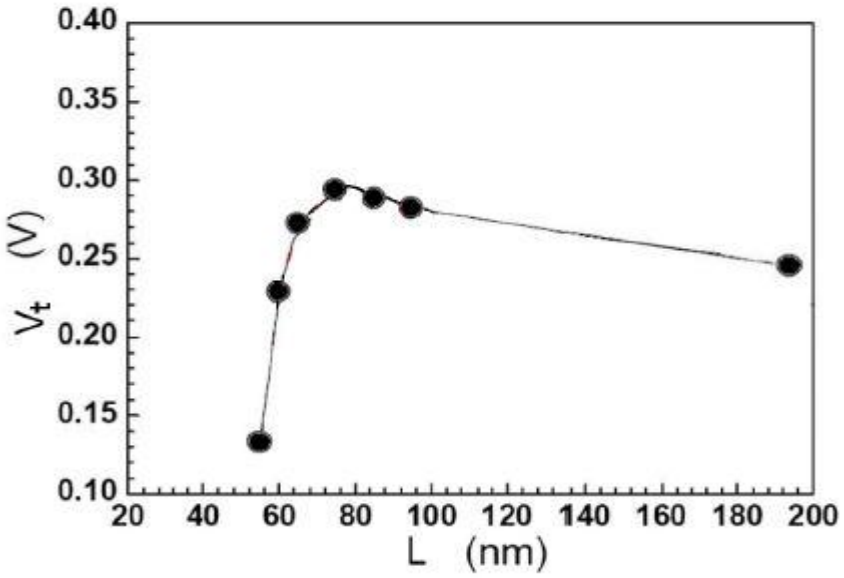


Figure 2.4: Measured Vth versus channel length L for a 90nm CMOS technology with shows strong short channel effects causing sharp roll-off for Vth for shorter L [15]

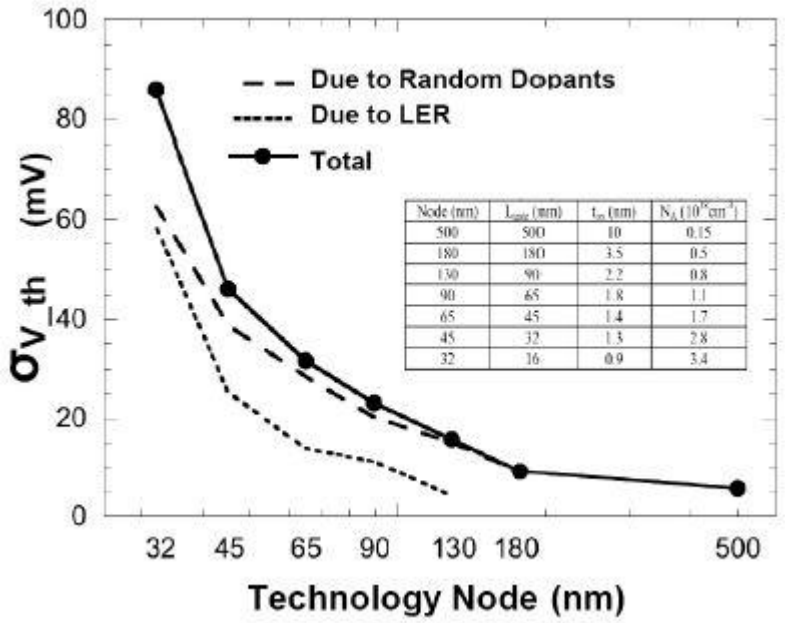


Figure 2.5: Predicted $\sigma_{V_{th}}$ including RDF and LER versus technology nodes for the smallest transistor. The inset shows the technological parameters used [8]

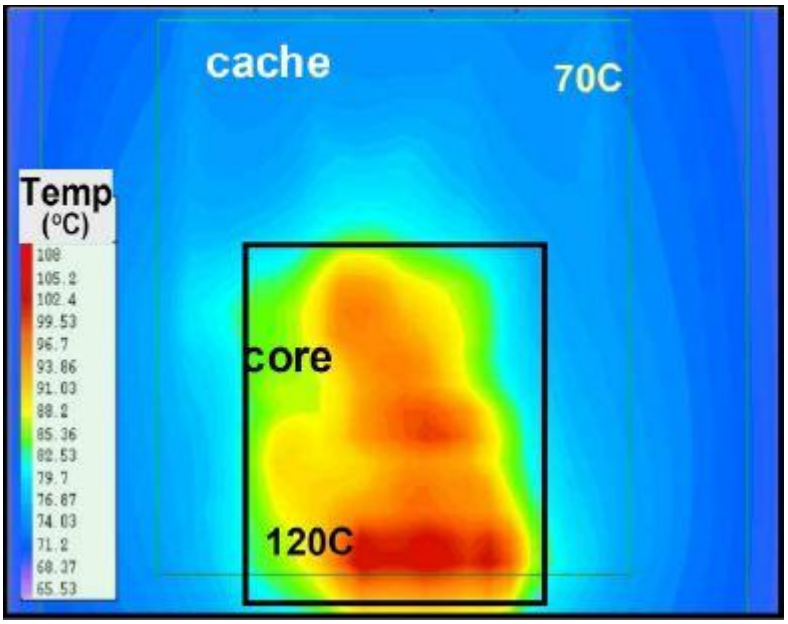


Figure 2.6: Thermal profile showing WID temperature variation for a microprocessor. Hot spots with temperatures as high as 120℃ are shown [25]

2.1.3. Impact of Variability on the Frequency and Power

In the nanometer regime, the reduction of the threshold voltage causes a substantial increase in the device sub-threshold leakage current which flows between the drain and source of a transistor when V_{GS} is less than the transistor's threshold voltage V_{th} [7, 26]. Sub-threshold leakage current has an exponential dependence on the threshold voltage. Furthermore, sub-threshold leakage is also very sensitive to temperature, doubling for every $8°K$ to $10°K$ temperature increase [27]. Leakage power is considered a very significant portion of the total power consumed in sub-90nm technology nodes. It is expected that the leakage power can reach up to more than 50% at 45nm technology, as shown in Figure 2.7.

The large variability in advanced CMOS technology nodes plays an important role in determining the total chip leakage [28]. This has underlined the need to take statistical leakage variations into consideration during the design cycle [28, 29]. Figure 2.8 shows the measured variations for both leakage power and frequency for 65nm

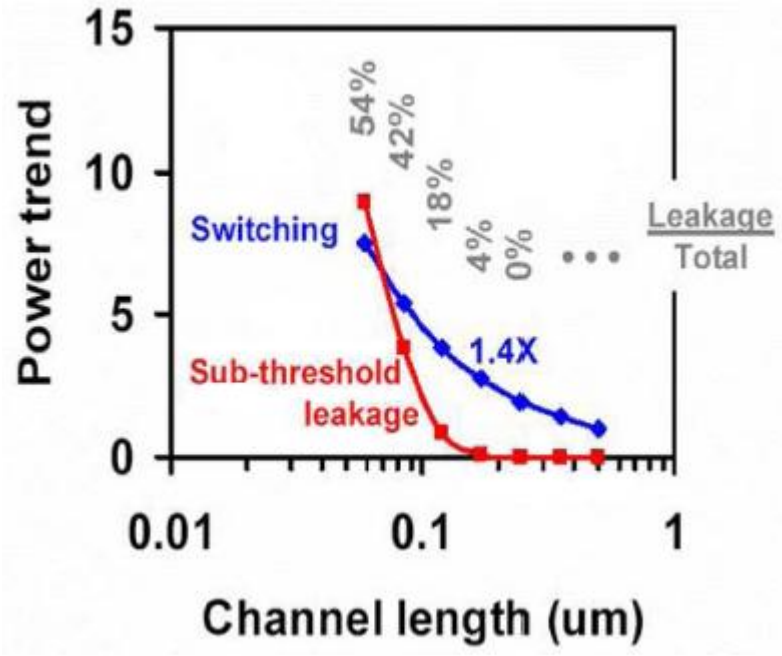


Figure 2.7: Dynamic (switching) and static (leakage) power versus technology scaling, showing the exponential increase in leakage power [26]

technology node which illustrates that there is a leakage variation about 10X for a 50% variation in chip frequency [30].

According to [25], a large percentage of the chips that meet the required operating frequency constraint dissipate a huge amount of leakage power. This makes them inconvenient for usage and, accordingly, causes yield degradation. This is due to the trade-off between leakage current and circuit performance. For devices with smaller V_{th} than nominal due to channel length variations, the sub-threshold leakage current increases exponentially. Meanwhile, the circuit delay decreases with increasing the driving current, I_D, since the overdrive voltage ($V_{DD} - V_{th}$) is increased. Thus, those

chips have higher operating frequency, but they suffer from huge leakage which makes them unacceptable [6, 25, 31].

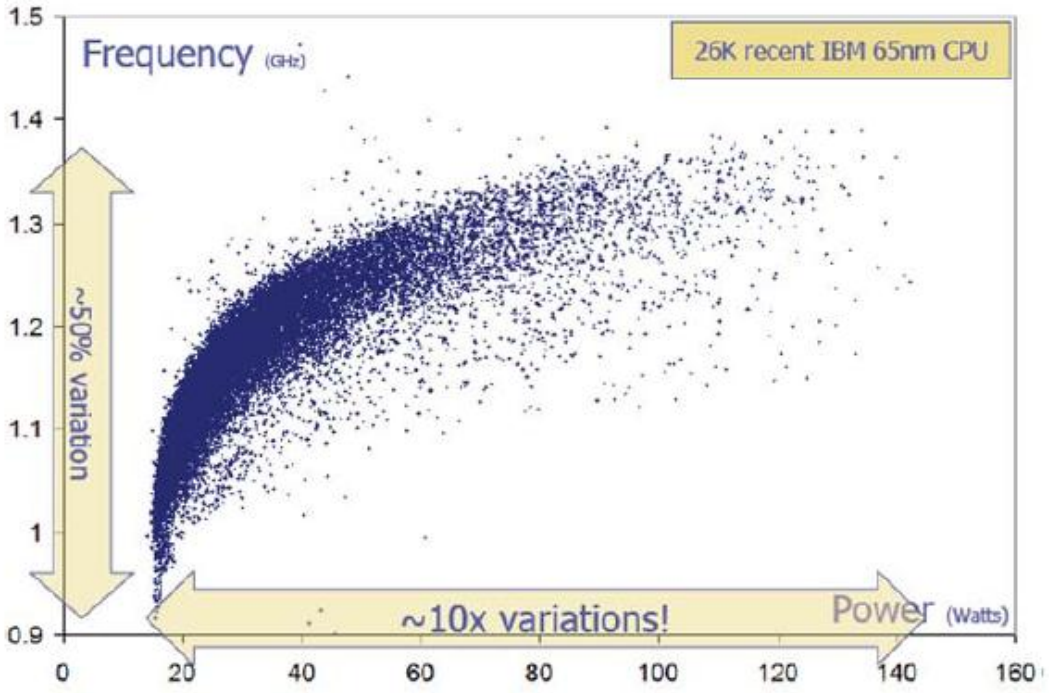


Figure 2.8: Leakage and frequency variations for IBM processor in 65nm technology [30]

2.1.4. State-of-Art Variations Mitigation Techniques

In this subsection, we discuss couple of state-of-art related research dealing with the increase in variability in nanometer regime in order to improve the yield. The first method is using Computer Aided Design (CAD) tools and statistical design which attempt to model the variations in the design flow cycle. And the second method attempts to deal with variations at the architecture level.

2.1.4.1. CAD Tool and Statistical Design

Recently, a large number of research work has been made in the area of CAD tools that attempt to account and model the random variations at the design flow level. One of the most researched topics in this area is Statistical Static Timing Analysis (SSTA) [9, 13, 21, 32, and 33]. In SSTA, the circuit delay is

Basically considered a random variable and SSTA then calculates the probability density function (PDF) of the delay at a certain defined path [9]. Similar to the SSTA which is used to model the delay variations, few research work targets modeling the process variations effects on other metrics including leakage power, noise margins, and soft errors [34-38].

Statistical design aims at statistically altering the circuit parameters at the design phase in order to reduce the process variations' impact and increase both the circuit robustness and the yield. One of the most common statistical design techniques is statistical gate sizing at which either the length or width of the transistor is tweaked to modulate the current drive capability. For example, process variations may increase the circuit delay, the statistical gate sizing algorithms are proposed to reduce the mean and standard deviation of the delay variations and, thus, improve the timing yield [32, 33].

2.1.4.2. Variations Mitigation at the Architecture Level

One of the first pieces of work that related variability to architecture was the work introduced by Bowman *et a.l* [39-41], and presented a statistical predictive model for the maximum operating frequency (FMAX) distribution

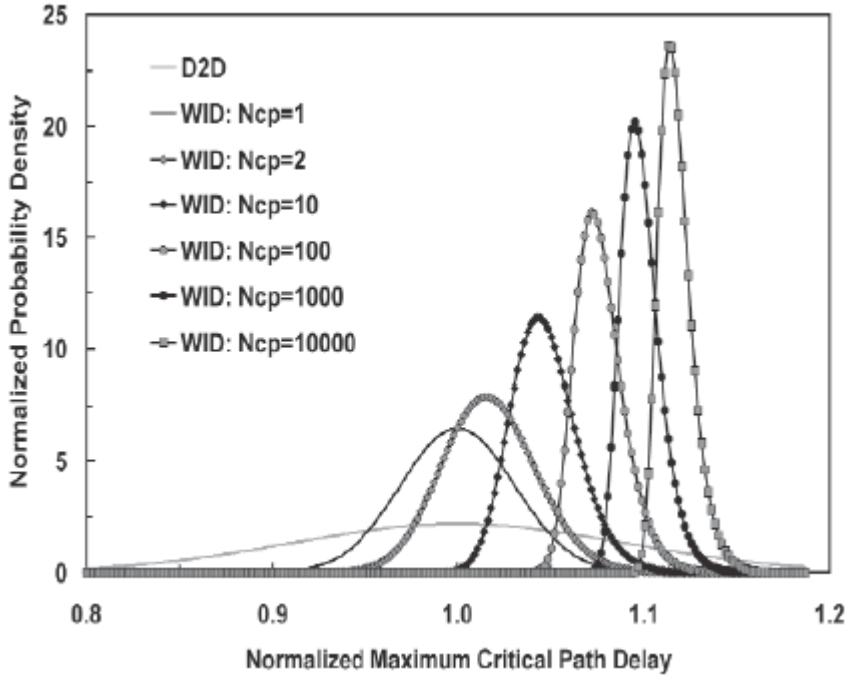


Figure 2.9: The WID maximum critical path delay distribution for different values of independent critical paths Ncp. As Ncp increases, the mean of maximum critical path delay increases [52]

in the presence of process variations in a chip. This technique provides insights on the impact of different components of variations on the FMAX distribution. The WID delay distribution heavily depends on the total number of independent critical paths for the entire chip N_{cp}. For a larger number of critical paths, the mean value of the maximum critical path delay increases, as shown in Figure 2.9. As the number of critical paths increases, the probability that anyone of them will be strongly impacted by process

variations becomes higher, and as a result, increases the mean of critical path delay. On the other hand, the delay's standard deviation (or delay spread) decreases with larger N_{cp}, thus making the spread of the overall critical path determined mainly by D2D variations. The results showed that WID variations directly affect the mean of the maximum frequency, while D2D fluctuations affect the variance.

Another factor that impacts the delay distribution is the logic depth per critical path. The impact of logic depth on delay distribution is different when dealing with random or systematic WID variations. Random WID variations have an averaging effect on the overall critical path distribution, while systematic WID variations affect all the gates on the path, hence, increase delay spread.

Other variation-tolerant research work at the architectural level was presented in [42], where a statistical methodology for pipeline delay analysis was presented. The importance of logic depth in variability research studies was accentuated, and it was shown that the change in logic depth and imbalance between stage delays can greatly improve the yield of a pipeline. Techniques such as deep pipelining and the push for high clock speeds decreases logic depth and have an undesirable impact on design variability [42].

2.2. **FPGAs**

In order to sustain reprogrammable and easily reconfigurable designs with flexible prototyping capabilities and enhanced performance, taking advantage of hardware parallelism, Field-Programmable Gate Arrays (FPGAs) are being used in production as reliable candidates in electronic design in terms of performance, time to market, and long-term maintenance. In addition, FPGA-based hardware solutions are considered the most cost-aware solutions when it comes to decreasing the nonrecurring engineering (NRE) expenses. This is mainly due to the continuous change in system requirements and design specifications over time, compared to custom Application-Specific Integrated Circuits (ASIC) designs which endure far more NRE expenses.

There are two main platforms dominating the programmable logic devices market; FPGAs and Complex PLDs (CPLDs). FPGAs mostly incorporate Look-up Tables (LUTs) to implement the logic functions, while CPLDs use the sum-of-products approach for implementing the logic functions. Lately, FPGA vendors offered a comprehensive, alternative platform to FPGA for large volume production demands called structured ASICs [43, 44].

Traditionally, FPGAs basic structure, as shown in Figure 2.10, consist of input/output pads, array programmable logic resources embedded in a sea of programmable interconnects that are configurable to implement any logic function with the possibility of augmenting memory and multiplier blocks. However, state-of-the-art FPGAs usually include Digital Signal Processing (DSP) blocks, embedded memory, Phase-Locked Loops (PLLs), and other special feature blocks as shown in Figure 2.11. These features made FPGAs to be an appealing alternative for some System-on-a-Programmable-Chip (SoPC) designs.

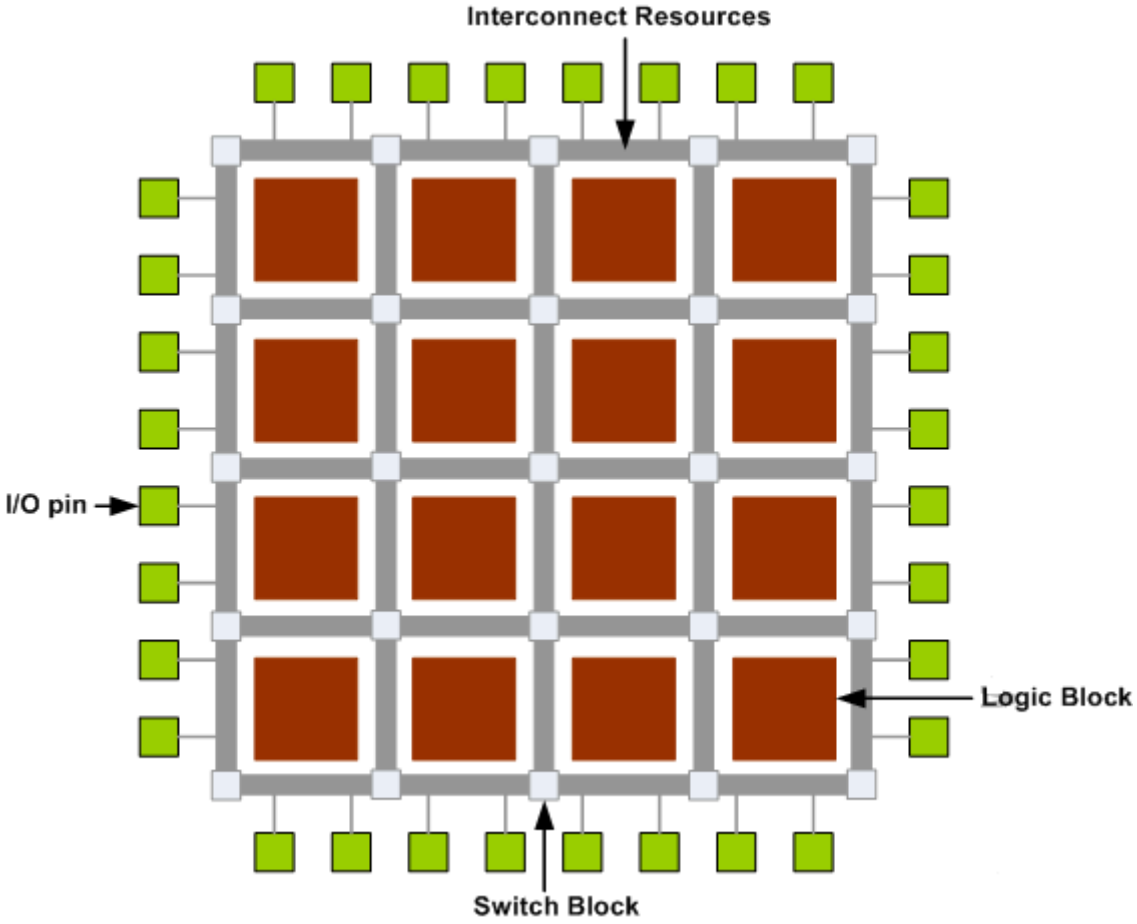


Figure 2.10: Basic FPGA structure

The technology used in programming both logic and the interconnect resources within the FPGAs can be flash memory [45], antifuse [46,47], or Static Random Access Memory (SRAM). For the SRAM-based FPGAs, they offer in-circuit re-configurability at the expense of volatility, while antifuse-based FPGAs are write-once devices. Flash-based FPGAs provide an intermediate solution by providing re-configurability as well as non-volatility. The most widely used programming technology in FPGAs now is SRAM as shown in Figure 2.12.

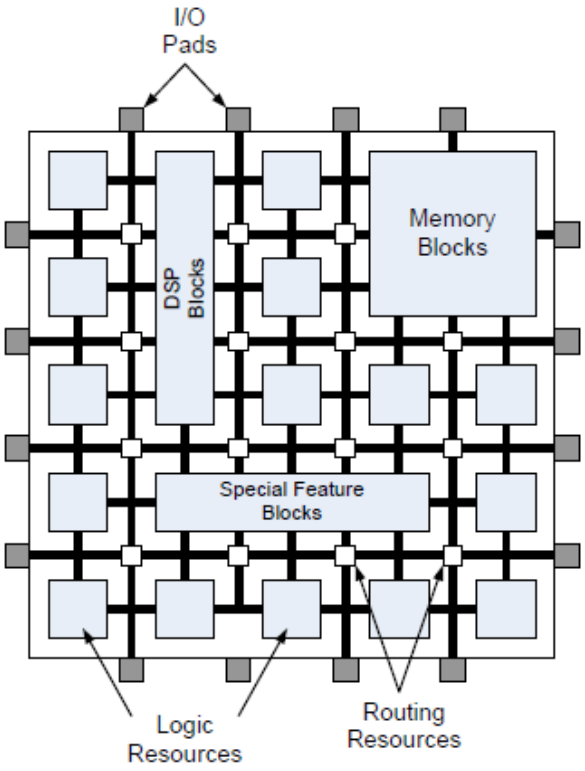


Figure 2.11: Modern FPGA fabric

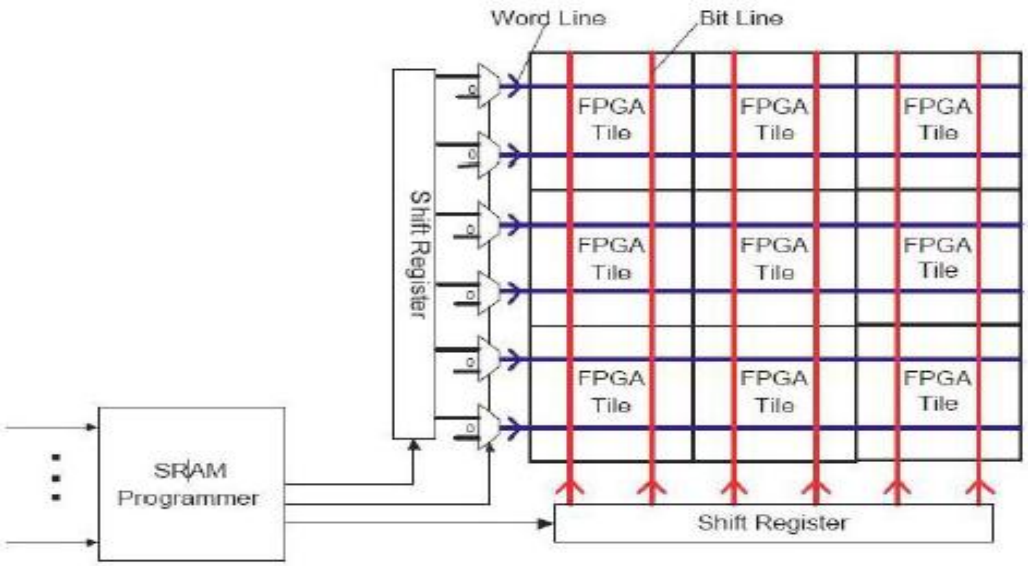


Figure 2.12: SRAM Programmer for logic and routing resources

2.2.1. FPGA Logic Resources Architecture

The logic blocks within FPGAs are mainly responsible for implementing the functionality needed by each application. Increasing the logic blocks' size, i.e., increasing the number of inputs to each logic block, increases the number of logic functions that could be performed by each logic block and also improves the delay/area performance of the logic block [48, 49]. However, this comes on the expense of wasted resources since logic blocks will not have all of their inputs fully utilized.

Currently, most commercial FPGAs incorporate LUTs to implement the logic blocks. A *k*-input LUT basically consists of 2^k configuration bits by which the required truth table is programmed during the configuration stage. The almost standard number of inputs for LUTs is four, which was proven optimum for both area and delay objectives [49]. However, this number can vary depending on the targeted application.

2.2.2. FPGA Under Study

The targeted FPGA built used in our study is an Island-style FPGA which consists of 2-dimensional array of repeated tiles, each consists of a logic cluster block, routing channels to connect inputs and outputs to the clusters, and Inter-cluster routing in order to connect clusters with each other. One level down of hierarchy of the tile is shown in Figure 2.13. Some previous work was done to study different configurations of FinFET-based FPGA LUTs [50] implemented using 16nm technology and simulated using HSPICE. The metrics used to evaluate the different candidate LUTs were the delay, energy, and the layout area.

In this research work, we have built FinFET-based FPGA logic cluster on schematic level using Cadence Virtuoso as shown in Figure 2.14. It consists of three basic logic elements (BLE), each BLE encapsulates a LUT, as shown in Figure 2.15, with size of 4 (four inputs), D-Flip-Flop, and 2-to-1 multiplexer to select either the registered or unregistered LUT output. Both the cluster size, N and the LUT size, K have been obtained by experimentally deriving the relationship between the number of cluster logic inputs required to achieve utilization percentage of 98% as a function of K and N [49]. This is $= \frac{K}{2} \times (N + 1)$, where I is the number of distinct cluster inputs (8 in our case, as reported in Table 2.1). Generally, we have in our design 11 inputs to the logic cluster, eight of them are distinct inputs while the other three LUTs which makes the output of each LUT available for direct connection to one of the inputs of the nearby LUTs in the same cluster which implies the "fully connected" approach; This means that all I cluster inputs and N outputs can be connected to each of the K inputs on every LUT. This, as a result, increases the FPGA speed by saving the number of inputs and bypassing the long capacitive routing channels as shown in Figure 2.16.

The SRAM cell with its sizing used within the FPGA cluster and the Transmission Gate Flip Flop (TG-DFF) built are shown at Figures 2.17 and Figure 2.18 respectively. Figure 2.19 shows the overall schematic of the cluster we built for our research work including three BLEs and 12 16-to-1 multiplexer units.

The FPGA cluster built for this work has been configured to build 2-bit adder benchmark circuit by manually programming the SRAM cells in LUTs accordingly and configuring the selection lines of the multiplexers to allow fully-connecting the BLEs.

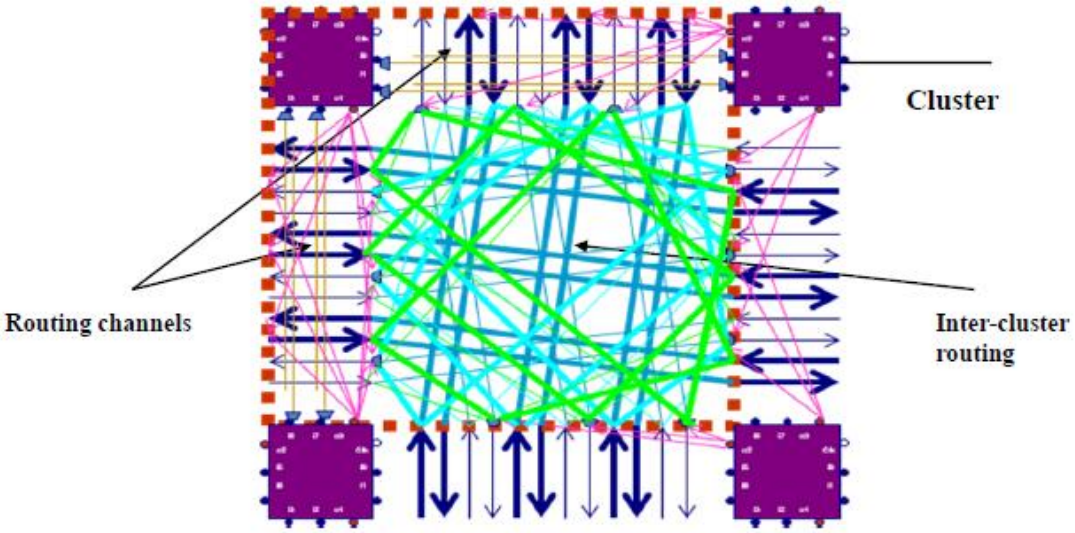


Figure 2.13: A closer look at the tile of Island-Style FPGA

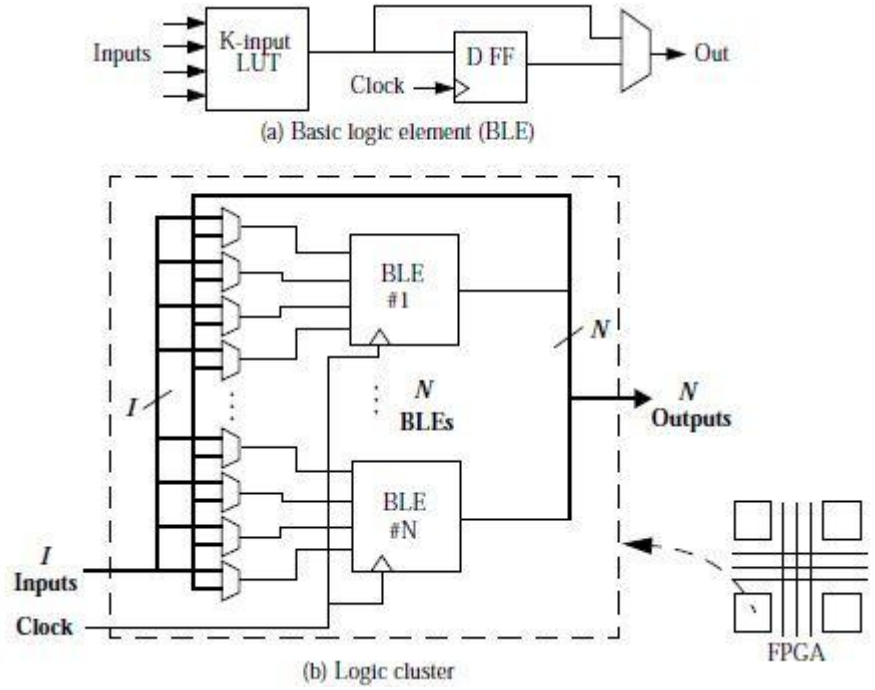


Figure 2.14: Structure of (a) Basic Logic Element (BLE) and (b) Logic cluster

Table 2.1: Architecture decisions for the FPGA

Parameter	Value
LUT size (K)	4
Cluster size (N)	3
Number of cluster inputs (I)	8

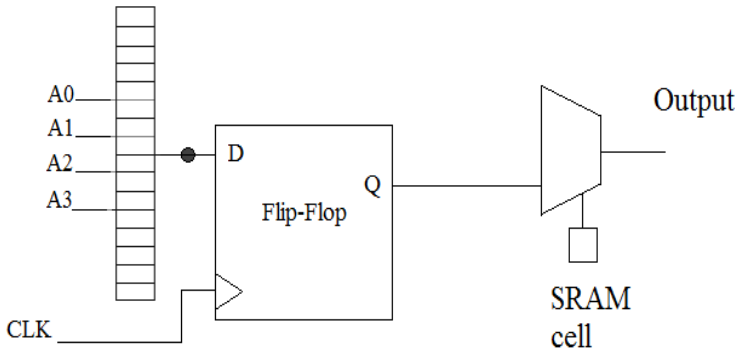


Figure 2.15: Lookup table with 4 inputs and 16 SRAM cells

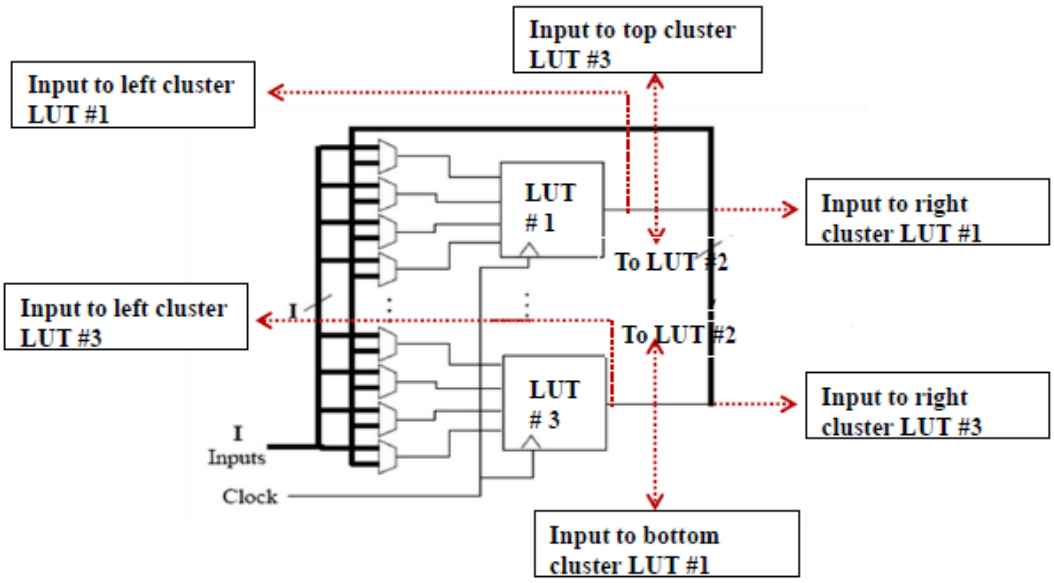


Figure 2.16: Sneak-path design in FPGA cluster

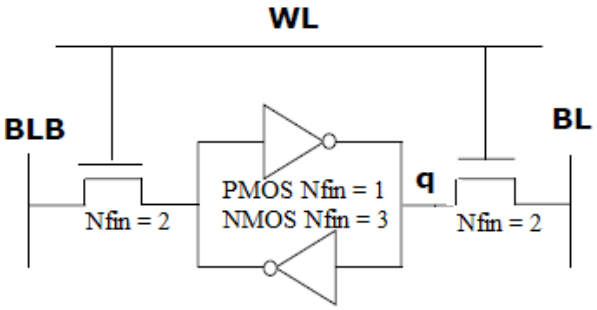


Figure 2.17: SRAM structure and sizing

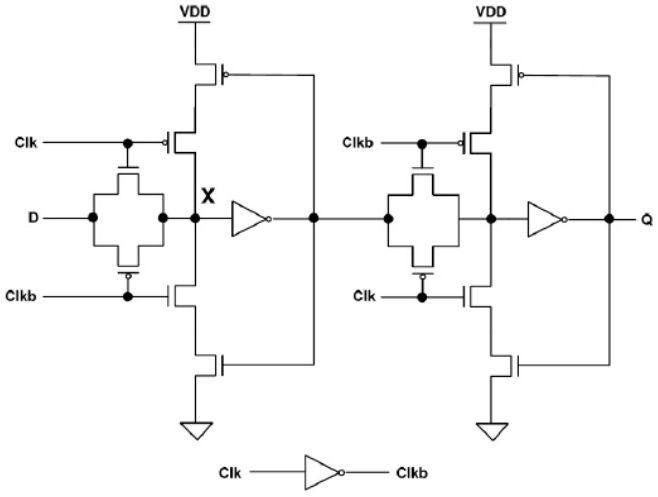


Figure 2.18: Transmission Gate Flip-Flop

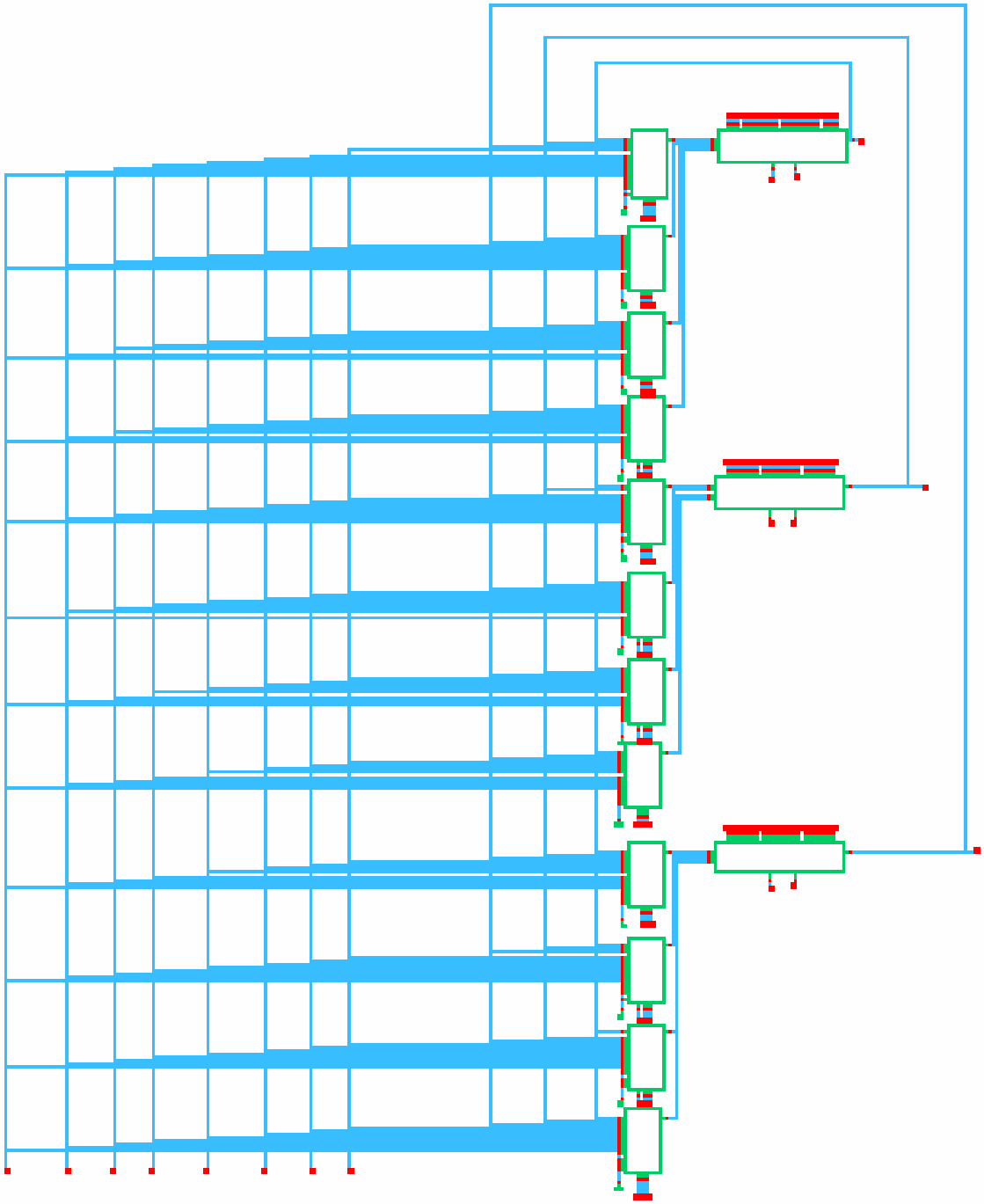


Figure 2.19: FinFET-based FPGA cluster with 3 BLEs and 12 16-to-1 multiplexers

Chapter 3 : Performance Evaluation of FinFET-Based FPGA Cluster Under V_{th} Variation

The performance of FinFET-based FPGA cluster is evaluated with technology scaling for channel length from 20nm down to 7nm showing the scaling trends of basic performance metrics. The impacts of threshold voltage variation representing D2D variations, along with the temperature variation, on the delay, power, and power-delay product are reported after simulating 2-bit adder benchmark. Simulation results show an increasing trend of the average power and power-delay product variations with threshold voltage as we go down with technology node. On the other hand, the delay shows the least percentage of variations with threshold voltage at the most advanced node of 7nm.

This chapter is organized as follows: section 3.1 gives detailed introduction about FinFETs and their dominance in the production for future technology nodes. Section 3.2 presents the simulation methodology used to measure the performance metrics. Section 3.3 presents the results and discussions of the performance evaluation study. Some design insights are given at Section 3.4. Finally, the conclusion is presented at Section 3.5.

3.1. Introduction

Rigorous scaling of planar MOSFETs towards deep sub-micron regime has delivered ever-increasing transistor density and performance to ICs. However, the continuation of MOSFETs scaling in nanometer technologies is becoming extremely challenging because of the dramatic increase in the sub-threshold leakage current [7, 26, 51]. With deeply scaled MOSFETs the channel lengths are becoming very narrow and, as a result, the drain voltage starts to dominate the electrostatics of the channel and, accordingly, the gate starts to lose sufficient control over the channel. Consequently, the gate is unable to completely shut off the channel while operating in the off-mode, which increases I_{OFF} between the source and the drain. Using high-k dielectric materials and thinner gate oxides helps alleviating this problem by increasing the gate-channel capacitance. However, thinning the gate oxides is limited by the deterioration in Gate-Induced Drain Leakage (GIDL) and gate leakage [52-54]. Multiple-Gate Field-Effect Transistors (MGFETs), which are an alternative to planar MOSFETs, show better screening of the drain potential from the channel because of the proximity of the additional gate(s) to the channel (which means higher gate-channel capacitance) [5, 55-59]. This makes MGFETs superior to planar MOSFETs in short-channel performance metrics, such as threshold voltage (V_{th}) roll-off, DIBL, and sub-threshold slope (S). Improvement in these performance metrics implies less degradation in the transistor's V_{th} with continued technology scaling, which then implies less degradation in $I_{OFF.}$

Among all MGFETs, FinFETs (as a type of DGFET) and Trigate FETs (another popular MGFETs with three gates) have emerged as the most desirable and attractive alternatives to planar MOSFETs due to their simple structure and ease of fabrication [60-68]. Figure 3.1 shows a conventional planar MOSFET and a FinFET. Two or three gates wrapped around a vertical channel (Fin) enable manageable alignment of gates and compatibility with the standard CMOS fabrication process. In Trigate FETs, an

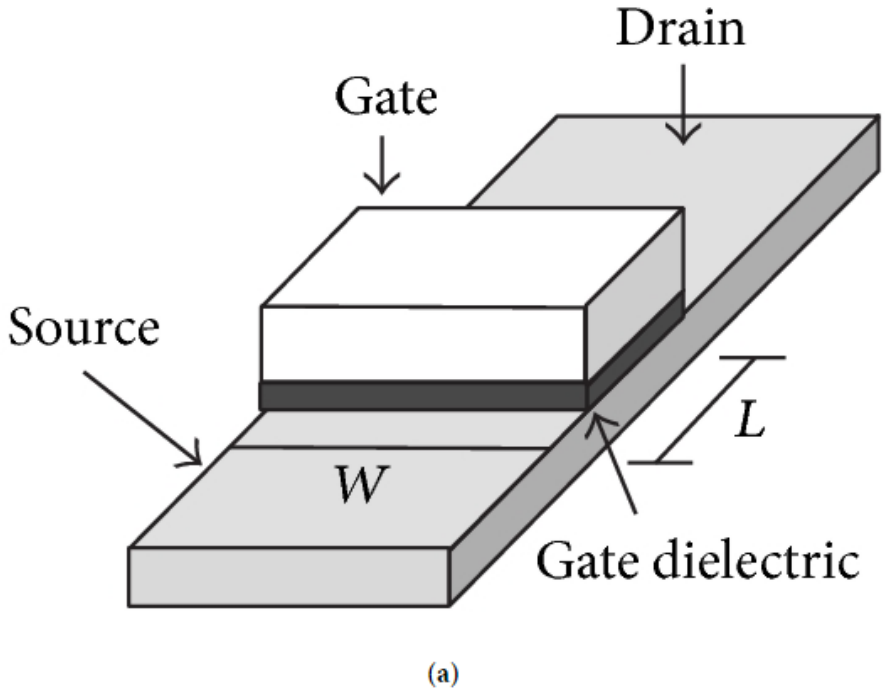
additional optional etching step of the hard mask is involved in order to create the third gate on top of the channel. This third gate leads to some advantages like additional transistor width and reduced fringe capacitances despite adding to process complexity [69-71]

Over the past decade, FinFETs have attracted increasing attention because of the degrading short-channel behavior of planar MOSFETs [60-65]. Figure 3.2 demonstrates the superior short-channel performance of FinFETs over planar MOSFETs with the same channel length. While the planar MOSFET channel is horizontal, the FinFET channel (Fin) is vertical. Hence, the channel height (H_{fin}) determines the FinFET width (W). This leads to a special property of FinFETs known as width quantization. This property says that the FinFET width must be a multiple of H_{fin}, that is, widths can be increased by using multiple fins. Thus, arbitrary FinFET widths are not possible. Although smaller fin heights offer more flexibility, they lead to multiple fins which, as a result, leads to more silicon area. On the contrary, taller fins lead to less silicon footprint, but they may also result in structural instability. Typically, the height of the fin is determined by the process engineers and is generally kept below four times the fin thickness [72, 73].
Although FinFETs implemented on SOI wafers are very popular, FinFETs have been also extensively implemented on conventional bulk wafers [74-76]. Figure 3.3 shows FinFETs implemented on bulk and SOI wafers. Unlike bulk FinFETs where all the fins share a common Si substrate/bulk, fins in SOI FinFETs are physically isolated. Some foundries prefer the bulk technology because it is much easier to migrate to bulk FinFETs from the conventional bulk MOSFETs. However, FinFETs on both types of wafers are quite comparable in terms of yield, performance, and cost. The rest of the discussion will be limited to SOI FinFETs.

Intel firstly introduced Trigate FETs, or interchangeably referred to as FinFETs, at the 22 nm node in the Ivy-Bridge processor in 2012 [69, 77]. Figure 3.4 shows a Trigate FET along with a FinFET. The thickness of the dielectric on top of the fin is reduced in Trigate FETs in order to create the third gate. Due to the existence of the third gate, fin thickness adds to the channel width as well. Thus, Trigate FETs have a slight width advantage over FinFETs. Trigate FETs also have less gate-to-source capacitance compared to FinFETs due to additional current conduction at the top surface, but this advantage is diminished by increased parasitic resistance [70].

3.1.1. FinFET Classification

There are two main types of FinFETs: Shorted-Gate (SG) and Independent-Gate (IG). SG FinFETs are also known as three-terminal FinFETs and IG FinFETs as four-terminal FinFETs. In SG FinFETs, both the front and back gates are physically shorted, while in IG FinFETs, the gates are physically isolated (Figure 3.5). Thus, in SG FinFETs, both gates are jointly used to control the channel electrostatics. Hence, SG FinFETs show higher on-current (I_{ON}) and also higher off-current (I_{OFF} or the sub-threshold current) compared to IG FinFETs. On the other hand, IG FinFETs offer the flexibility of applying different voltages to their two gates. This enables using the back-gate bias to linearly modulate the V_{th} of the front gate. However, IG FinFETs have larger area overhead because of the need for placing two separate gate contacts.



(a)

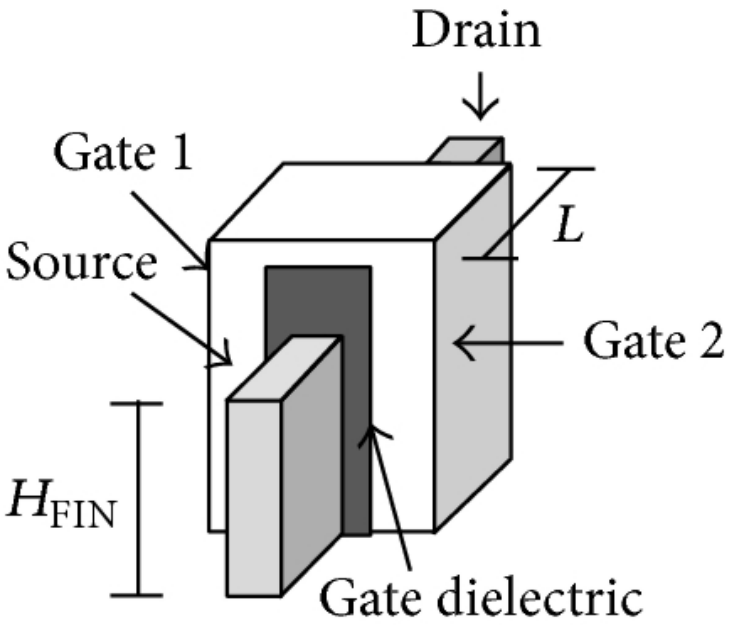


Figure 3.1: Structural comparison between (a) planar MOSFET and (b) FinFET

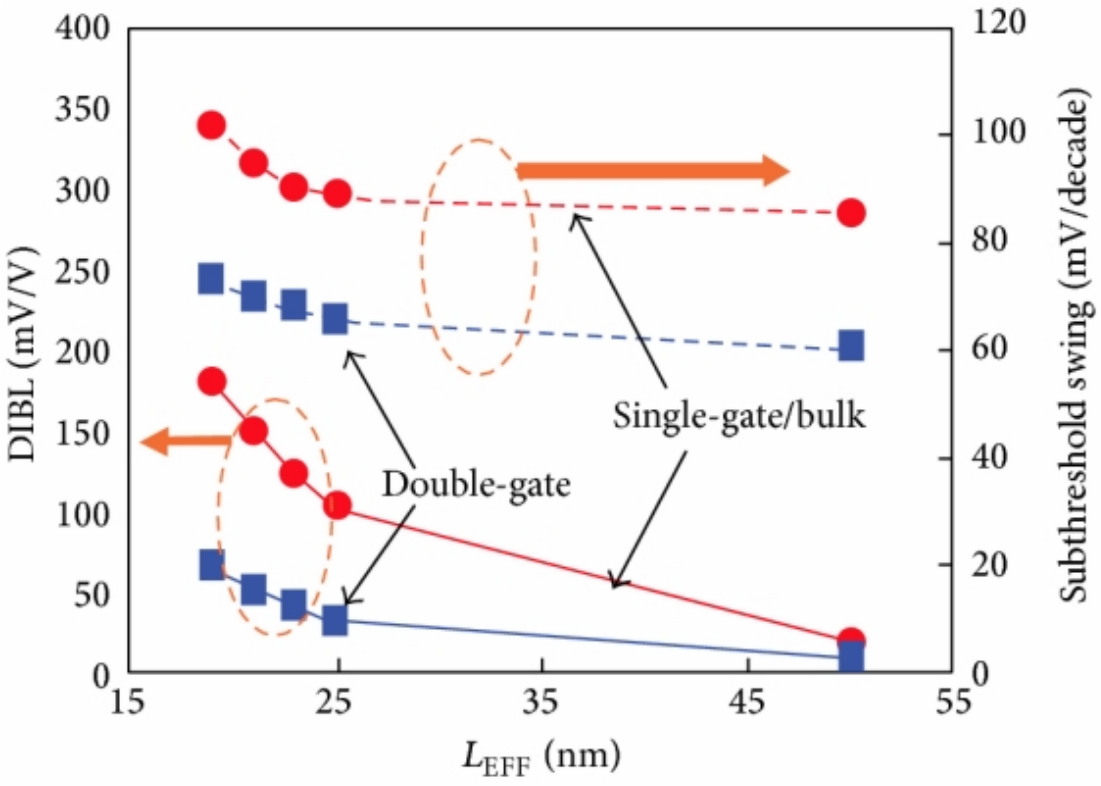


Figure 3.2: DIBL and sub-threshold swing (S) versus effective channel length for double-gate (DG) and bulk-silicon nFETs. The DG device is designed with an undoped body and a near-mid-gap gate material [59]

SG FinFETs can be further subdivided based on asymmetries in their device parameters. Normally, the work functions (Φ_m) of both the front and back gates of a FinFET are the same. However, the work functions can also be made different. This leads to an asymmetric gate-work function ASG FinFET [78, 79]. ASG FinFETs have very promising short-channel characteristics with I_{OFF} that is two orders of magnitude lower than that of an SG FinFET, with only somewhat lower I_{ON} than that of an SG FinFET [80]. They can be fabricated with selective doping of the two gate-stacks. Figures 3.6 and 3.7 show comparisons of the drain current versus front-gate voltage curves for SG, IG, and ASG nFinFETs and pFinFETs, respectively, demonstrating the advantages of ASG FinFETs.

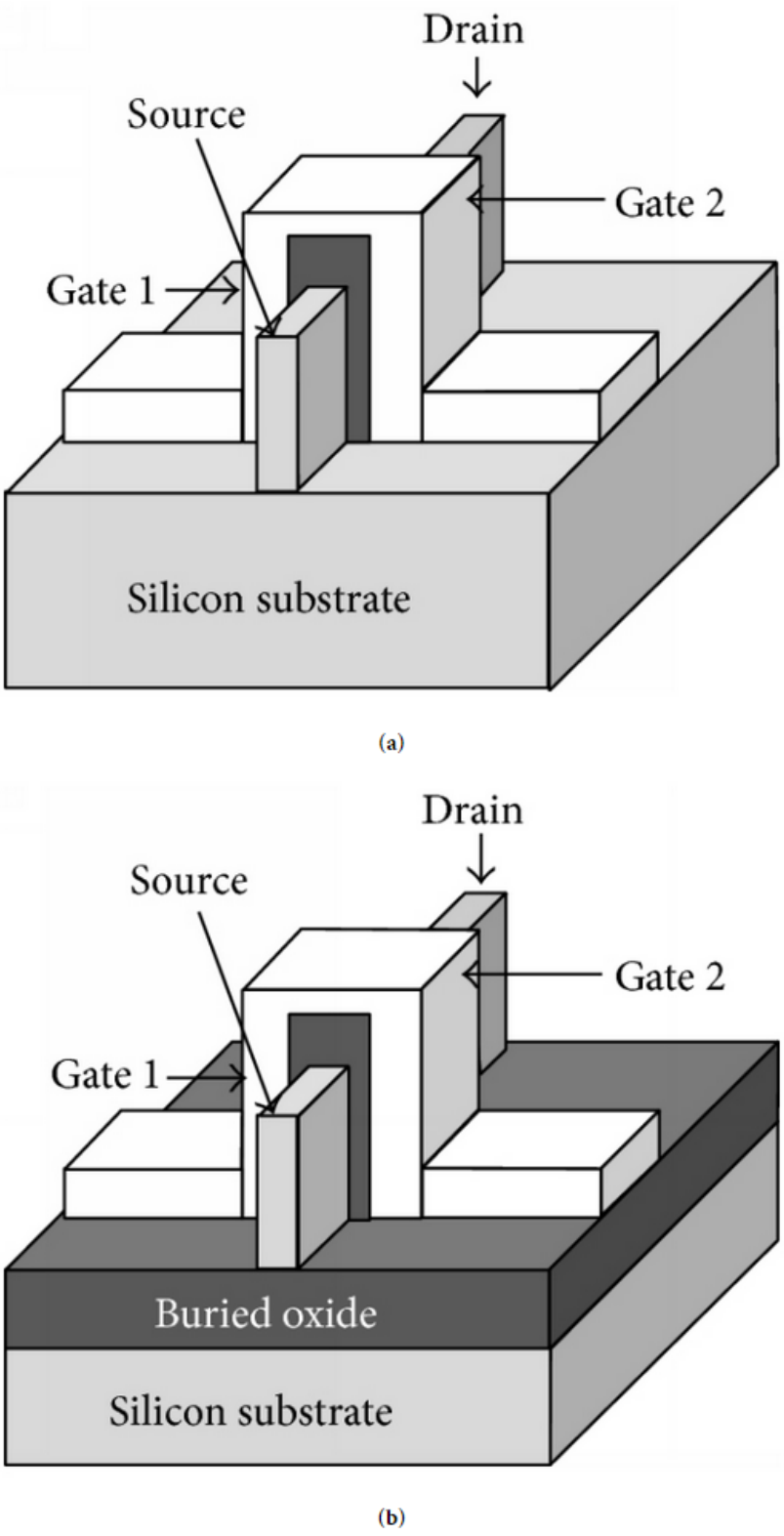


Figure 3.3: Structural comparison between (a) bulk and (b) SOI FinFETs

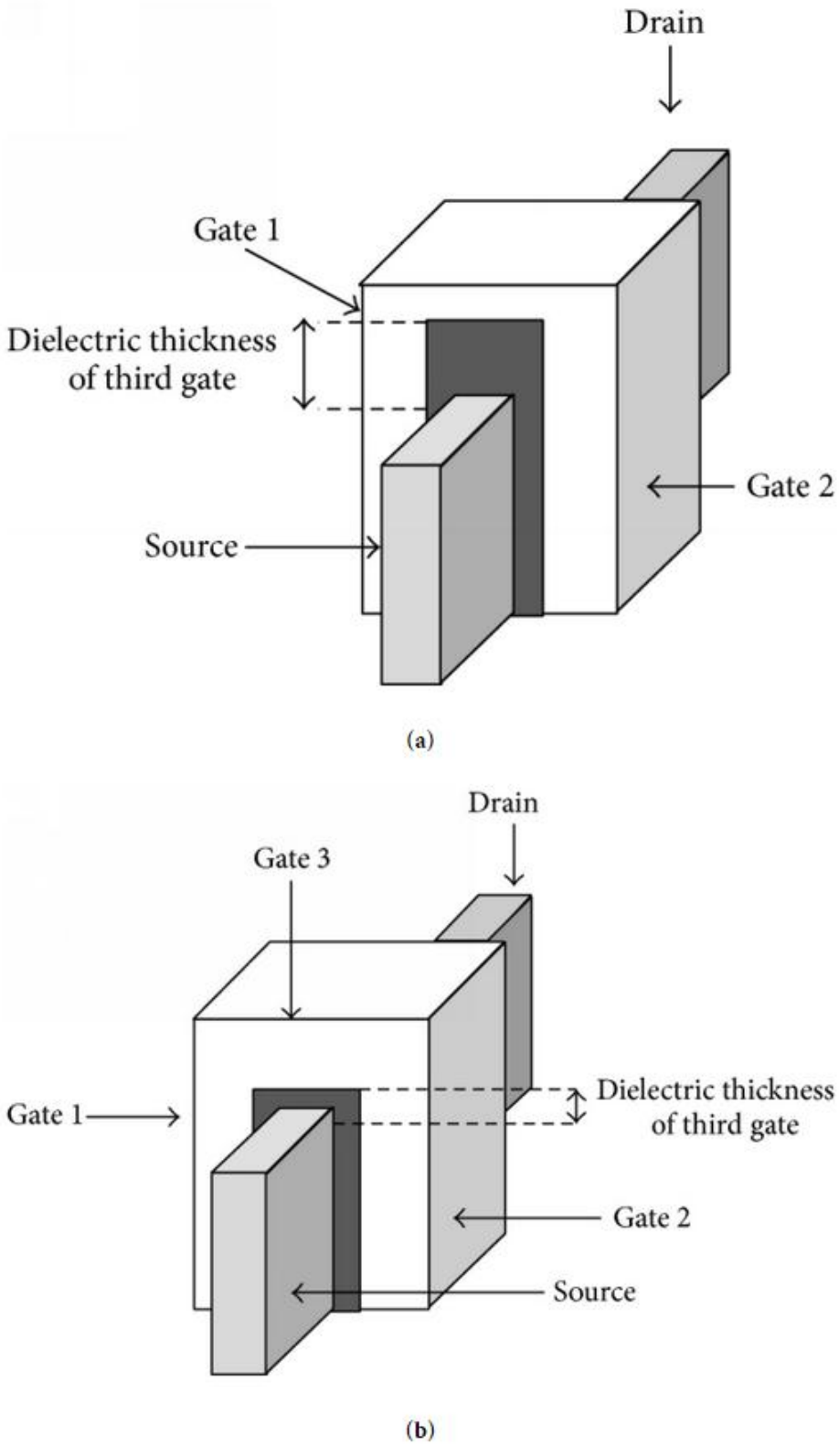


Figure 3.4: Structural comparison between (a) FinFET and (b) Trigate FET

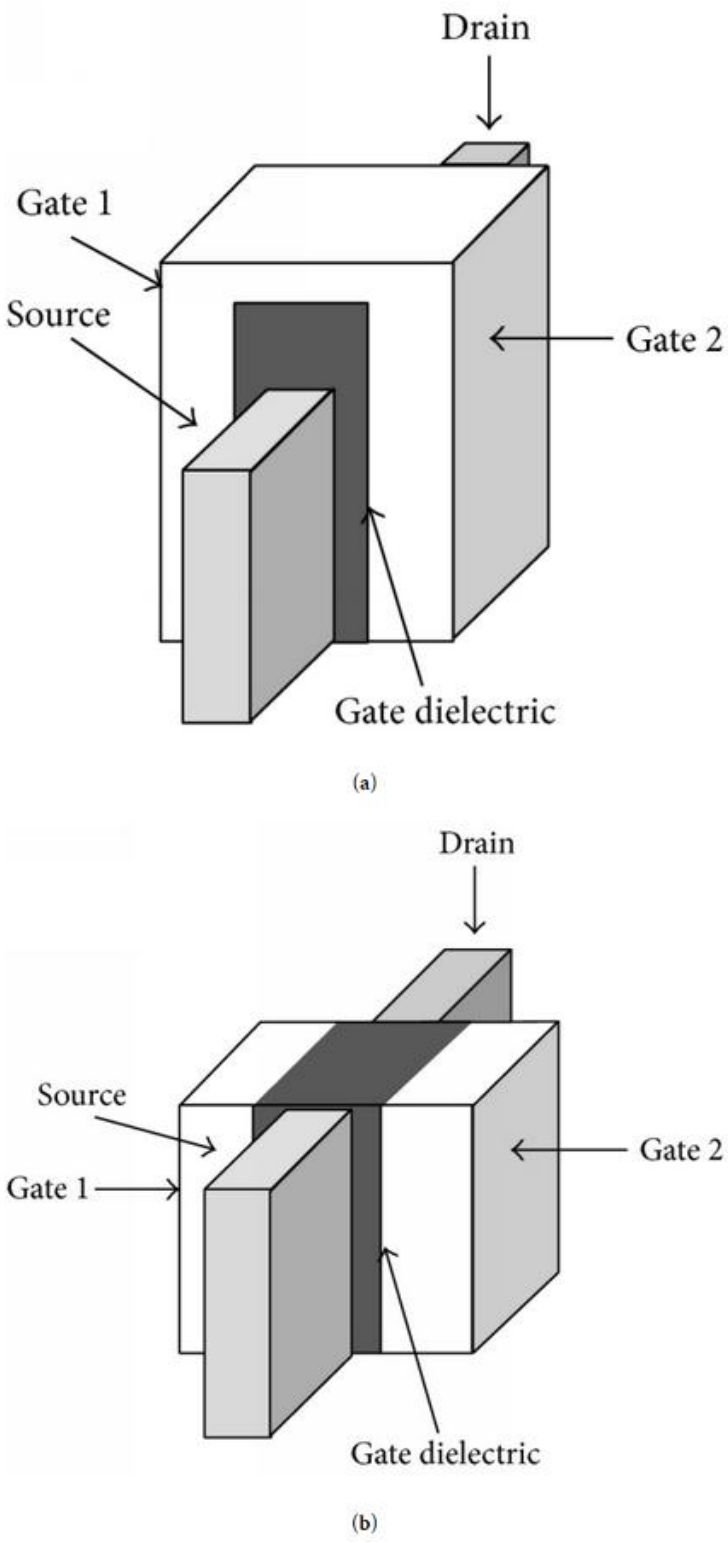


Figure 3.5: Structural comparison between (a) SG and (b) IG FinFET

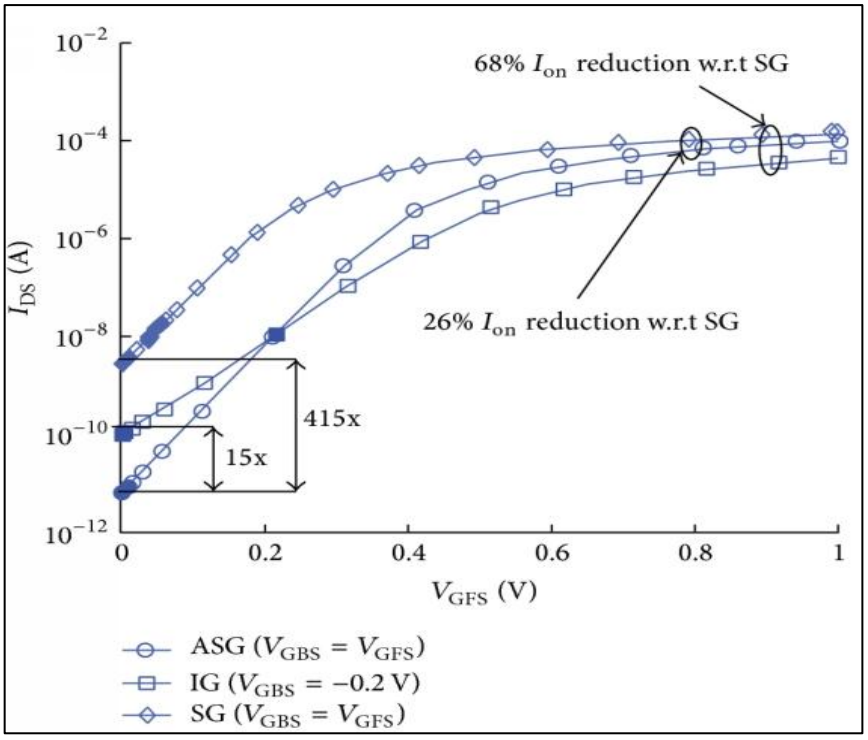


Figure 3.6: Drain current (IDS) versus front-gate voltage (VGFS) for three nFinFETs [80]

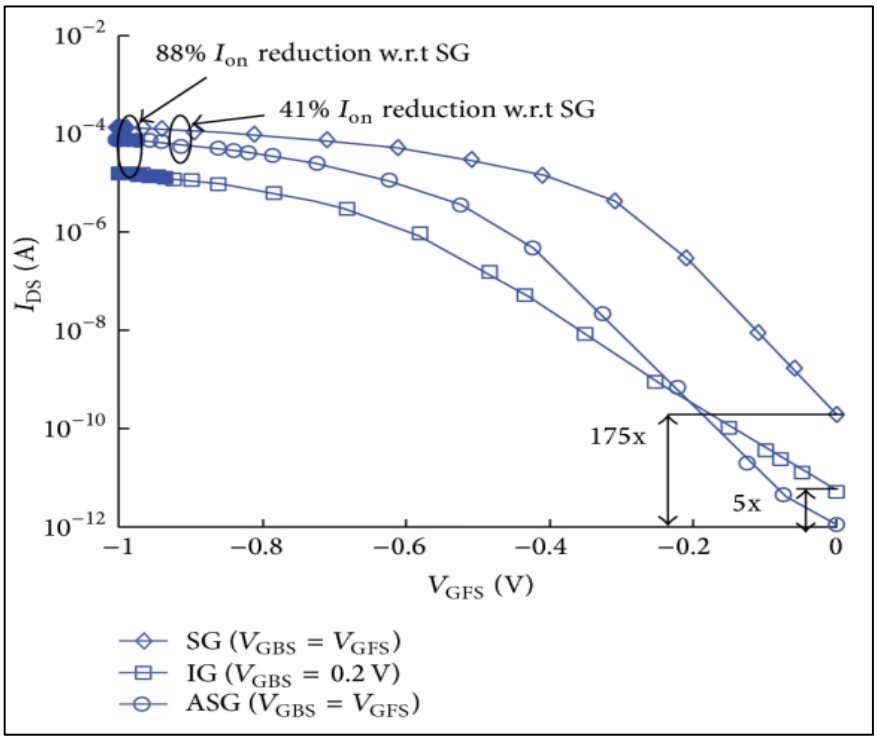


Figure 3.7: Drain current (IDS) versus front-gate voltage (VGFS) for three pFinFETs [80]

3.1.2. Process Variations for FinFET

In order to tackle the Sort-Channel Effects (SCEs) in planar MOSFETs, a sufficient number of dopants must be injected into the channel. However, this means that RDF may lead to a significant variation in V_{th}. For instance, at deeply scaled technology nodes, the $3(\sigma/\mu)$ variation V_{th} in caused by discrete impurity fluctuations can be greater than 100% [81]. Since FinFETs enable better SCEs performance due to the existence of a second gate, they do not need a high channel doping to ensure a high V_{th}. Hence, designers can keep the thin channel (fin) at nearly intrinsic levels (10^{15} cm^{-13}). This, in turn, reduces the statistical impact of RDF on V_{th}. The desired V_{th} is obtained by engineering the gate work function instead. Low channel doping also ensures better mobility of the carriers inside the channel. Thus, FinFETs emerge superior to planar MOSFETs by alleviating a major source of process variation.

Due to their small dimensions, along with the lithographic limitations, FinFETs do suffer from other process variations and they are subjected to many important physical fluctuations, such as variations in gate length (L_{GF}, L_{GB}), gate-oxide thickness (T_{OXF}, T_{OXB}), fin-thickness (T_{SI}), and gate underlap (L_{UN}) [81–87]. For example, gate oxide is

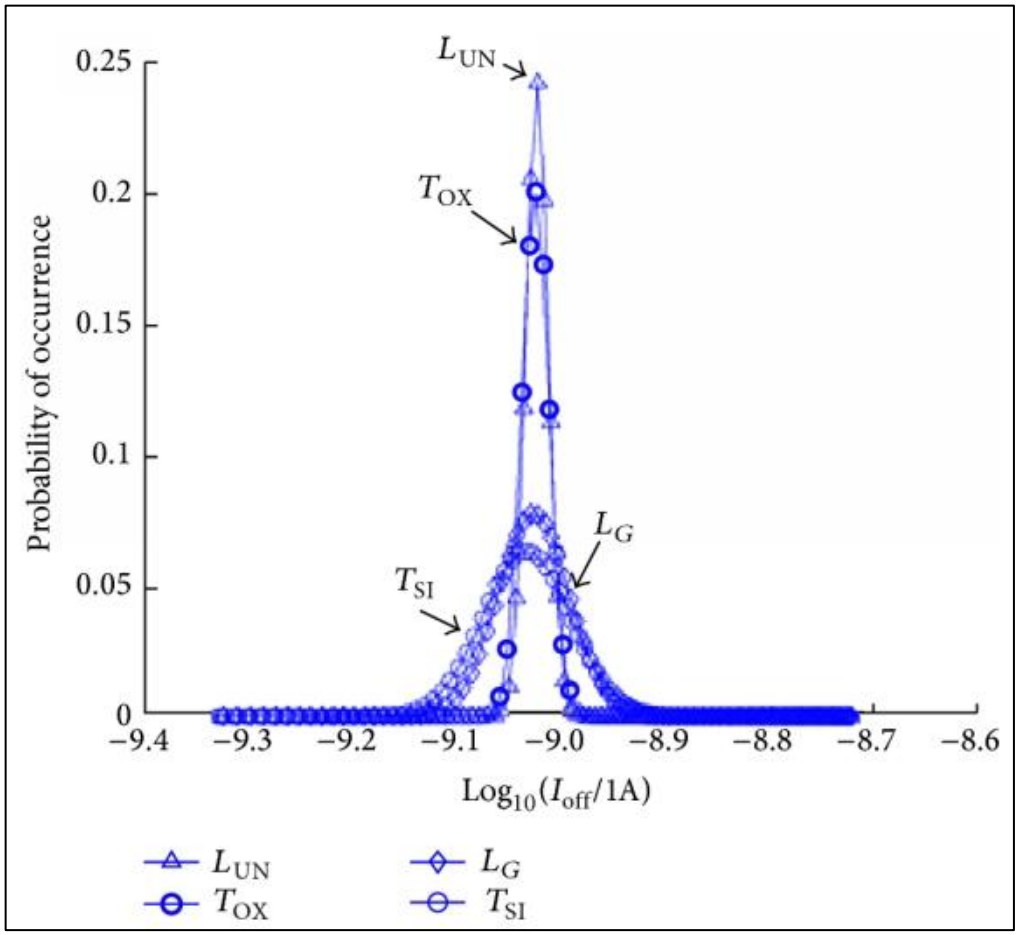


Figure 3.8: Distribution of leakage current (I_{OFF}) for different process parameters, each varying independently [84]

on the etched sidewall of the fin, and may suffer from non-uniformity. The degree of non-uniformity depends on the fin's Line-Edge Roughness (LER) fin. LER also causes variations in fin thickness. Figure 3.8 illustrates the impact of parametric variations on the sub-threshold current (I_{OFF}) of an nFinFET.

Choi et al. have also studied temperature variations in FinFET circuits in the presence of the aforementioned physical parameters variations [88]. They showed that even under moderate process variations ($3(\sigma/\mu)$ = 10%) in L_{GF}, L_{GB}, and $T_{SI,}$ thermal runaway is possible in more than 15% of ICs when having primary input switching activity of 0.4. The effect of temperature variation is more severe in SOI FinFETs because the oxide layer under the fin suffers from poor thermal conductivity. Thus, heat generated in the fin cannot dissipate easily in SOI FinFETs. Bhoj and Jha have evaluated SG, IG, and ASG FinFETs under temperature variation and found that although I_{OFF} degrades for all three FinFETs at a higher temperature, ASG FinFETs still remain the best and retain a 100X advantage over SG FinFETs, as shown in Figure 3.9 [80]. They also showed the distribution of I_{OFF} under process variations for the three FinFET types, as shown in Figure 3.10.

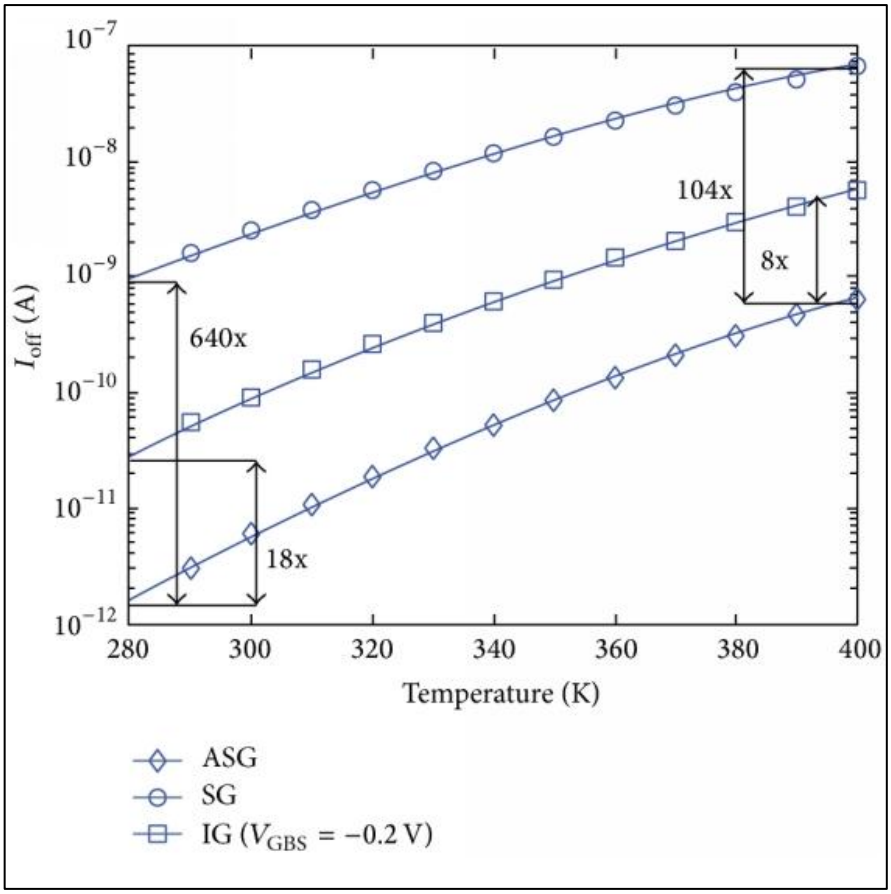


Figure 3.9: I_{OFF} versus temperature for three nFinFETs [80]

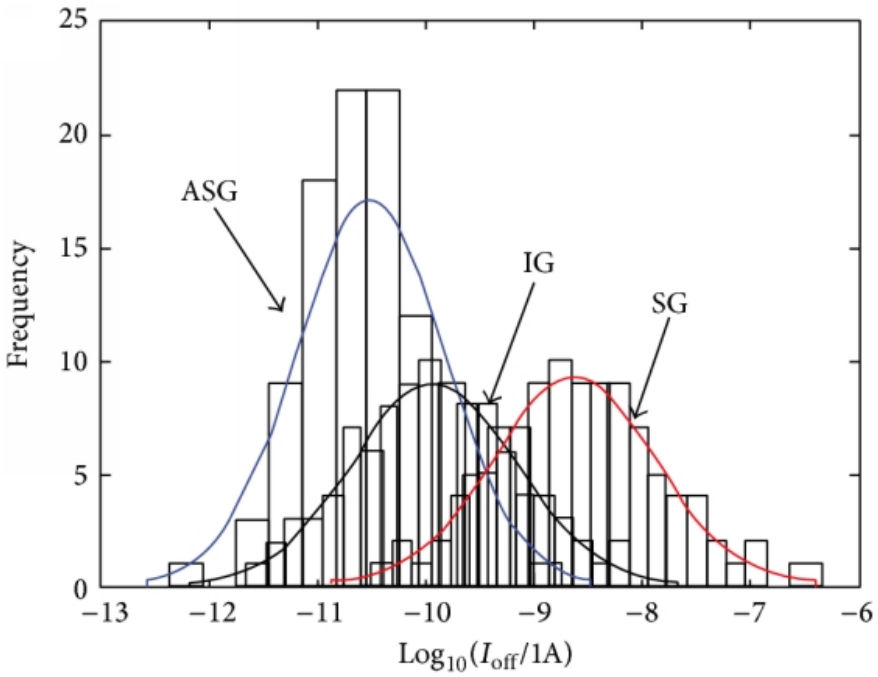


Figure 3.10: Distribution of I_{OFF} under process variations for three nFinFETs [80]

Table 3.1: Simulated Device Parameters

Device	TG-FinFET				
Lg (nm)	20	16	14	10	7
T_{fin} (nm)	15	12	10	8	6.5
H_{fin} (nm)	28	26	23	21	18
V_{DD}	0.9	0.85	0.8	0.75	0.7

3.2. **Simulation Methodology**

In this work, we used the predictive technology model for multi-gate transistors (PTM-MG) [89], described in Appendix A, starting from 20nm down to 7nm technology nodes for low-standby power devices (LSTP), and these models are based on BSIM-CMG compact models. PTM-MG models follow the scaling approach of not only scaling the channel length (L) but also the fin height (H_{fin}), fin thickness (T_{fin}), and supply voltage (V_{DD}), as reported in Table 3.1. Thus, the FinFET is used such that the effective channel width ($W_{eff} = 2\ H_{fin} + T_{fin}$).

Table 3.2: Threshold Voltage Variations

Node (nm)	Threshold Voltage (mv)			
	Nominal	% change values (of nominal)		
		±6%	±12%	±18%
7	268	16	32	48
10	292	17.5	35	52.2
14	311	18.6	37.3	55.9
16	320	19.2	38.4	57.6
20	330	20	40	60

The performance metrics used for cluster's performance evaluation under V_{th} and temperature variations were the average power, delay, and power-delay product (PDP). Average power is basically calculated by multiplying the average value of the current drawn from the source by the value of the supply voltage. Delay is calculated by taking into consideration the most critical path from the input to the output from the 3rd BLE, which represents the carry output resulting from the 2-Bit addition operation. And the PDP is the multiplication of both the average power and delay. This is calculated for each technology node in our study starting from 20nm down to 7nm. And for each technology node, we studied the performance metrics with V_{th} variation within range ±18% with 6% step of the nominal V_{th} of each technology node accordingly. The simulated values of V_{th} variations are reported in Table 3.2. We also studied the performance metrics' variations with temperature variations with range (-100% to 300%) with step 100% [90].

3.3. **Results and Discussions**

3.3.1. Average Power

The simulation results show that the average power variation percentages with V_{th} variation increase with FinFET technology scaling. Figure 3.11 shows the average power variation percentages with three different V_{th} variation percentages for all the technology nodes considered in this study. For each technology node, the variation percentages of the average power decrease with increasing the V_{th} variation percentages from -18% to -6%. Figure 3.12 shows the average power variation percentages as well with temperature variations from 100% to 300%, showing the same trend with maximum variation percentages for the most advanced technology node of 7nm.

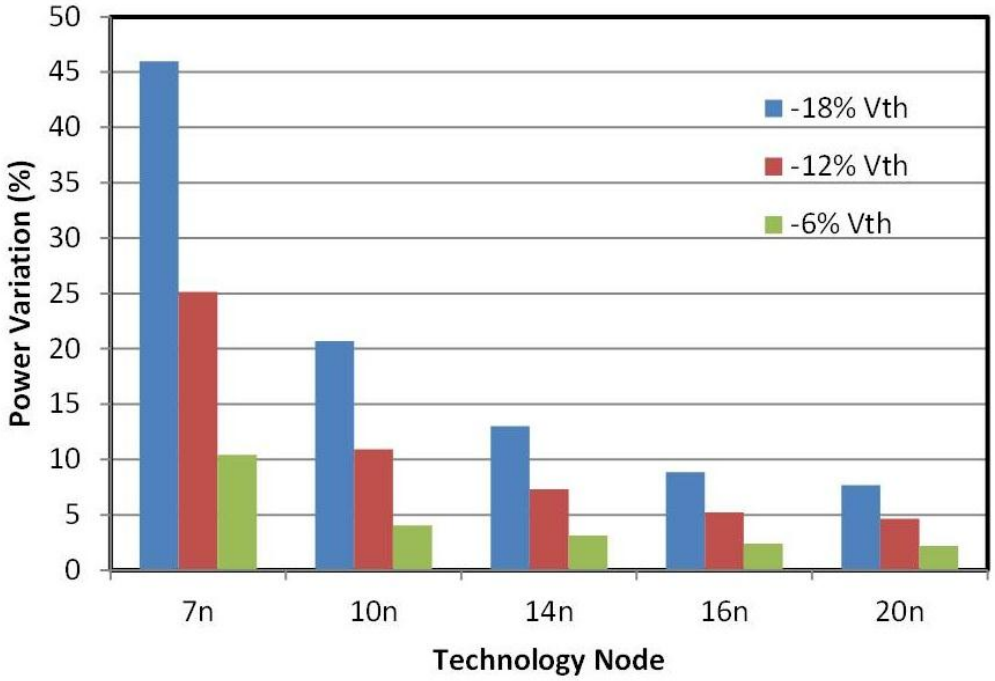


Figure 3.11: Average power variation percentages with Vth variation for various technology nodes [90]

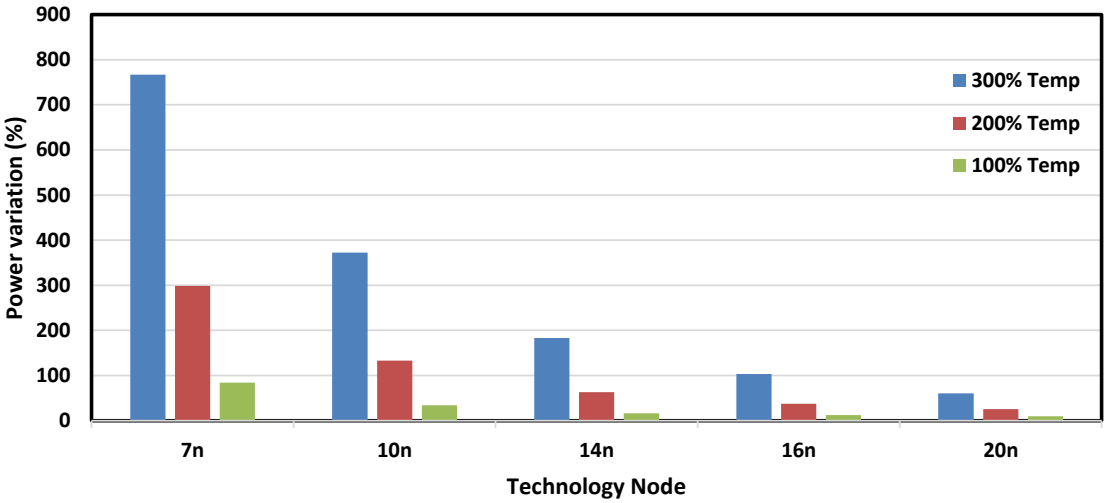


Figure 3.12: Average power variation percentages with temperature variation for various technology nodes

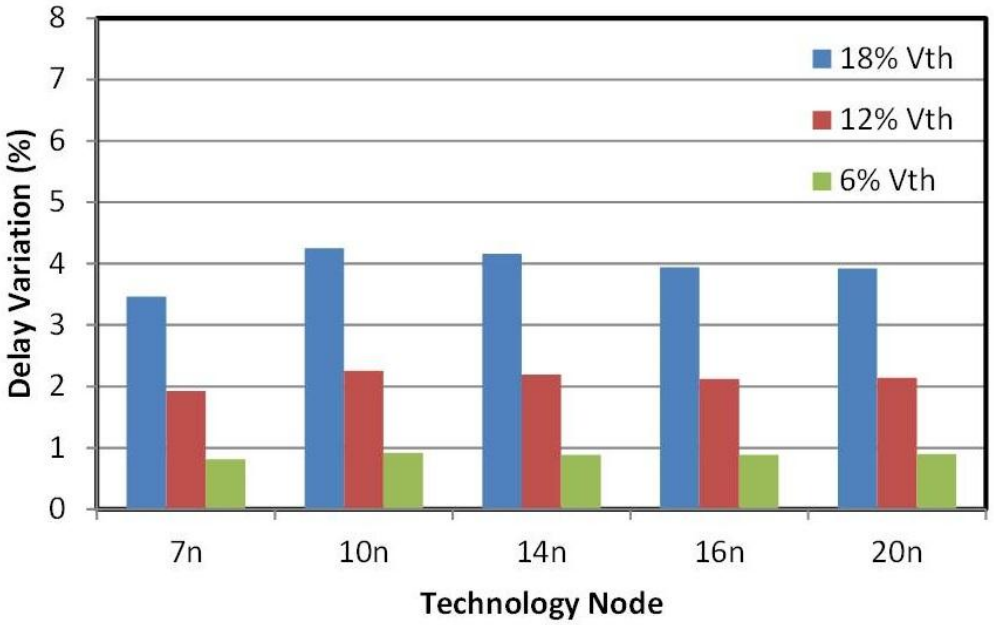


Figure 3.13: Delay variation percentages with Vth variation for various technology nodes [90]

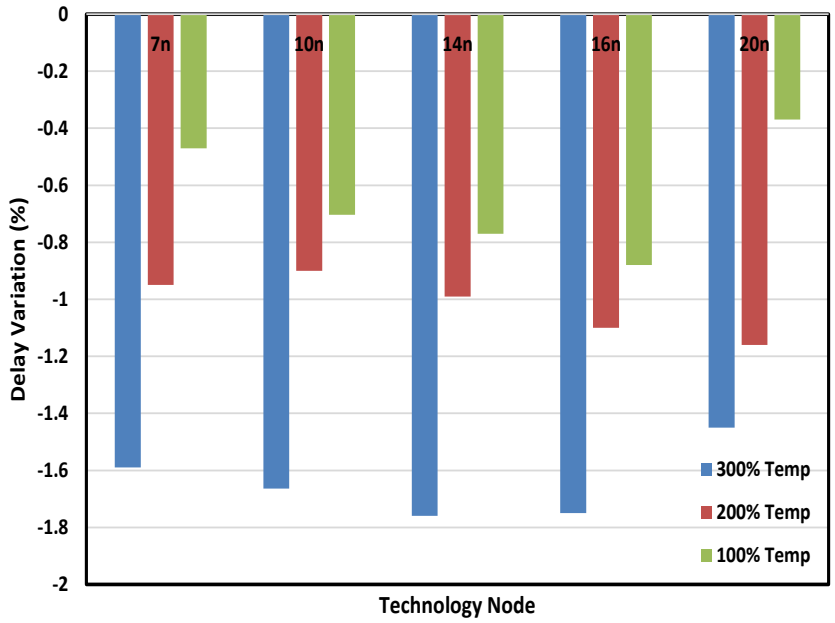


Figure 3.14: Delay variation percentages with temperature variation for various technology nodes

3.3.2. Delay

Regarding the simulation results for delay variation with V_{th} and temperature, they show that 7nm node has the least delay variation percentage with V_{th} variation compared to other technology nodes, as shown in Figure 3.13. About the other technology nodes, the delay variation percentages are nearly comparable. Also, the delay variation percentage increases with increasing the V_{th} variation percentage from 6% to 18% for each technology node. Figure 3.14 shows the delay variation percentages with temperature variation as well.

3.3.3. Power-Delay Product

Variation percentages of PDP with V_{th} variation and temperature variations are shown in Figures 3.15 and 3.16 respectively. PDP variation with V_{th} and temperature variation increase with FinFET technology scaling. The PDP behavior is following the same trend as the average power variation with technology nodes. Thus, the power variation is considered the dominant contributor in the PDP calculation compared to delay.

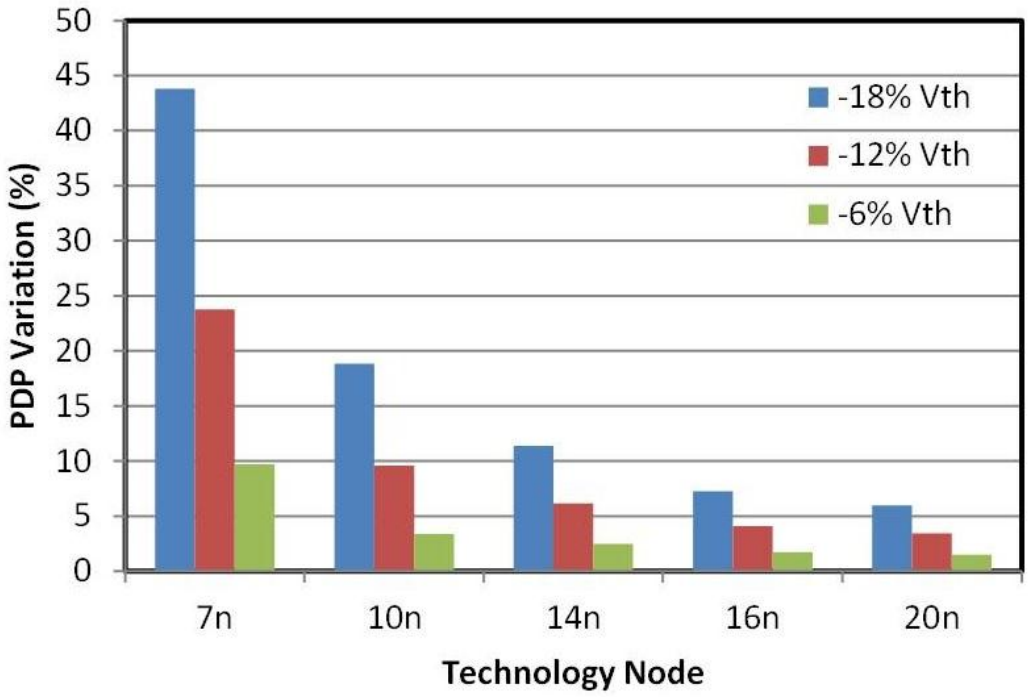


Figure 3.15: PDP variation percentages with Vth variation for various technology nodes [90]

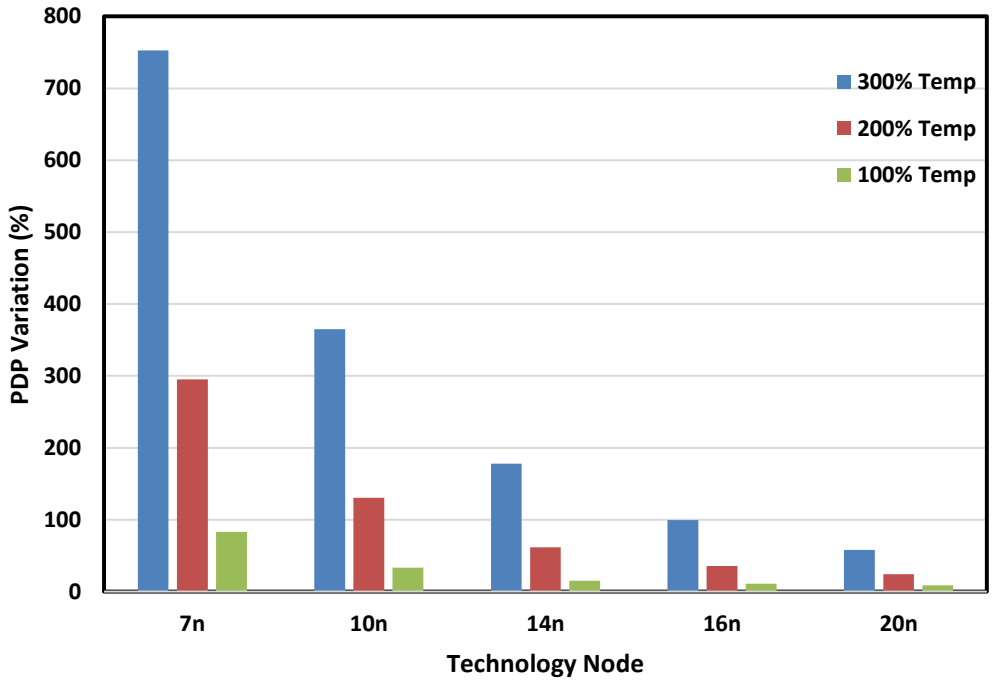


Figure 3.16: PDP variation percentages with temperature variation for various technology nodes

3.4. Design Insights

In our study, we defined a targeted yield percentage of 99.87% for which we determined the design constraints of different performance metrics. This targeted yield percentage represents the 3σ value, or three standard deviations of the mean, for a particular technology node; The mean value μ here is the nominal value (the metric value at zero percentage change in the V_{th} for this node), and σ here is calculated by calculating the standard deviation between each metric's values for different V_{th} variation percentages from -18% to 18% with 1% step (total of 37 corners including the nominal condition). Figures 3.17 to 3.19 show the design constraints values for average power, delay, and PDP for all the technology nodes calculated as $\mu \pm 3\sigma$. The large gap between the design constraints within the power and PDP curves starting at 14nm node and increasing till 7nm node emphasizes the further increase in the variations with technology scaling as previously mentioned.

3.5. Conclusion

The performance of FinFET-based FPGA cluster is evaluated with technology scaling by configuring the cluster to be 2-Bit adder benchmark. The impacts of both V_{th} and temperature variations on the basic performance metrics are evaluated and reported. The results show that both the average power variations and the PDP variations with V_{th} and temperature variations increase with technology scaling, while the delay variation with V_{th} and temperature variation is not following a certain trend with the

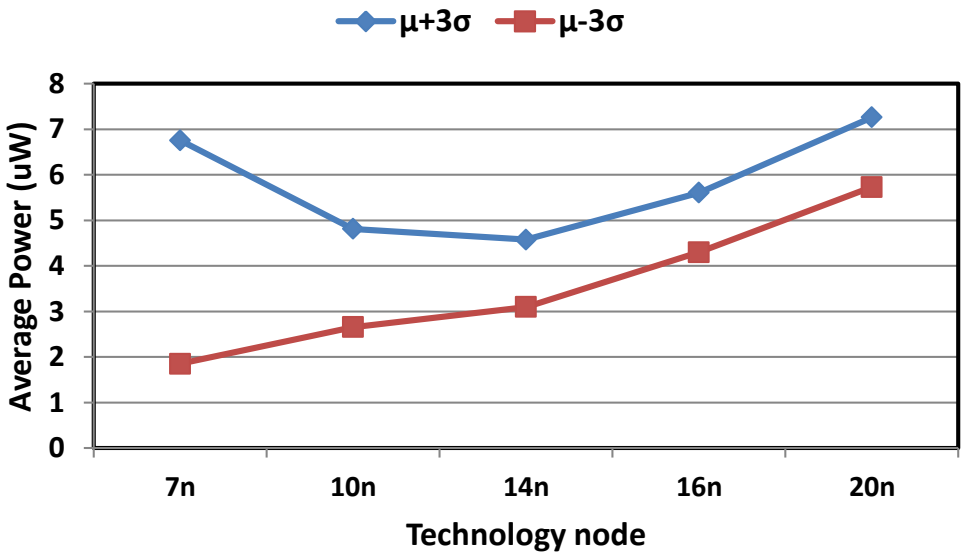


Figure 3.17: Power constraints with Vth for various technology nodes [90]

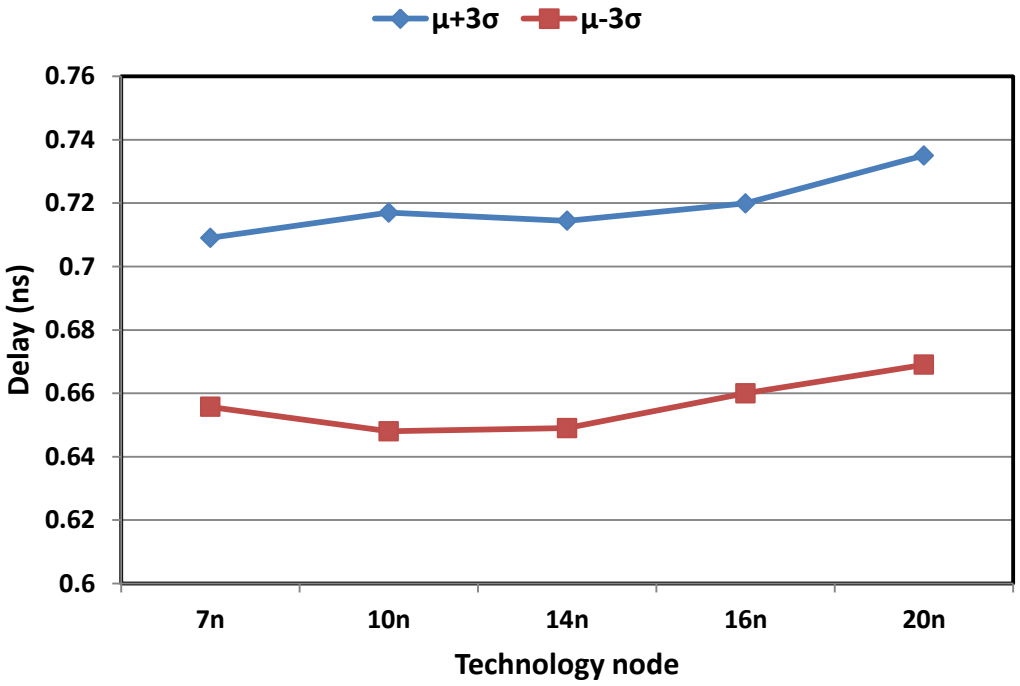


Figure 3.18: Delay constraints with Vth for various technology nodes [90]

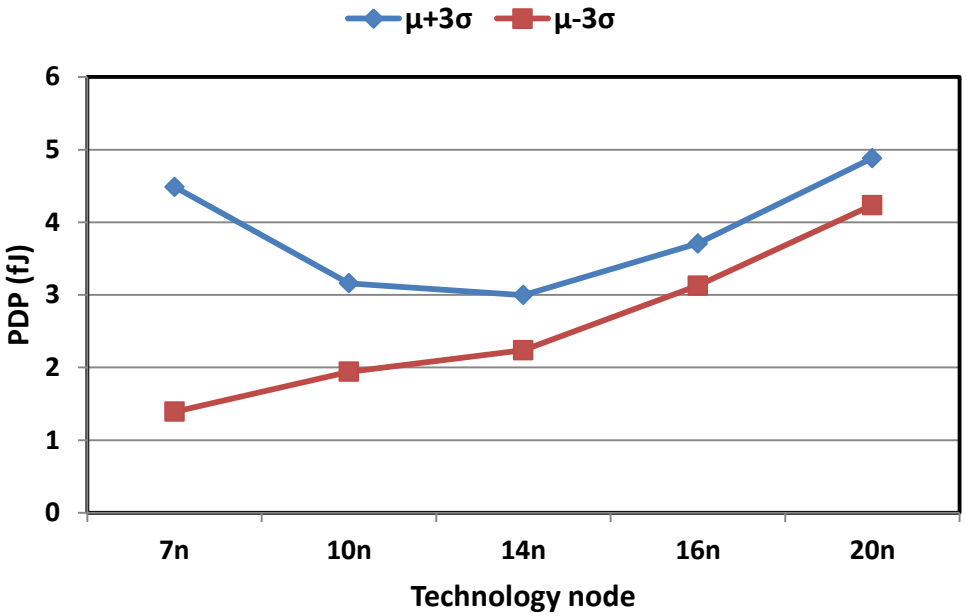


Figure 3.19: PDP constraints with Vth for various technology nodes [90]

technology scaling.
Design constraints values are reported for all the performance metrics included in this study to give the designers some useful insights and recommendations.

Finally, our work suggests a future work of building an FPGA tile that utilizes the cluster we have designed, with the associated routing channels and Inter-cluster routing to be considered in simulations.

Chapter 4 : Leakage Power Evaluation of FinFET-Based FPGA Cluster Under V_{th} Variation

The leakage power of FinFET-based FPGA cluster is studied and evaluated with technology node 14nm. The impact of V_{th} variation, representing D2D variations, on the leakage power is reported after simulating a 2-Bit adder benchmark and comparing the results with the dynamic power consumption. Simulation results show, with the leakage power segmentation, that the multiplexers are the most consuming components for leakage power in the FPGA cluster architecture. Some leakage power control techniques are investigated including transistor stacking, minimum leakage vector, and gate sizing. The effect of each technique on the leakage power, leakage power variation, and the delay is reported and compared with the original design.

This chapter is organized as follows; Section 4.1 gives an introduction about the leakage power and its different sources. Section 4.2 explains the simulation setup and methodology. Results and discussions for the leakage power evaluation study under threshold voltage and temperature variations are discussed in Section 4.3. Some proposed solutions to control the leakage power and investigating their effects on the leakage power variation are explained in Section 4.4. And finally the conclusion is drawn at Section 4.5.

4.1. Introduction

The continuous scaling for MOS devices in order to increase the performance and density and to lower the power consumption is resulting in efficient chip functionality at higher speeds and also in transistor delay reduction by 30% for each technology generation. However, the scaling poses some challenges by causing severe Short-Channel Effects (SCE) such as DIBL, GIDL, Band-to-Band Tunneling (BTBT), and V_{th} roll-off [91]. Scaling is applied on the supply voltage as well in order to keep control over the power consumption. Thus, the transistor's V_{th} has to be proportionately scaled in order to maintain a high drive current capability and achieve high performance improvement. However, scaling V_{th} results in dramatic increase in the leakage current, especially the sub-threshold leakage current. Thus, adversely affecting the power efficiency which is becoming the key to sustaining continually enhanced performance for future VLSI circuits.

4.1.1. Leakage Current Sources

There are many leakage current sources, as shown in Figure 4.1. I_1 is the BTBT current, which mainly flows between the source/drain and the substrate through the reverse biased P-N junction at OFF state. I_2 is the sub-threshold leakage current which is the major contributor and caused by weak inversion, DIBL, and body effect. I_3 & I_4 represent gate oxide tunnelling current and gate current due to hot carrier injection respectively, and both currents flow between gate to or from the substrate and both diffusion terminals. Both I_3 & I_4 are caused due to gate oxide thickness (t_{ox}) reduction causing more carriers to tunnel through the gate oxide. I_5 is the GIDL which is mainly current flowing between drain and well or substrate because of the high field effect in

the drain junction. I_6 is the punch-through current, which is considered as the sub-surface version of the DIBL current, and it flows within the substrate between the source and the drain because the depletion regions of the drain and source become close enough deep in the channel to conduct.
Out of these different leakage current sources experienced by current MOS devices, sub-threshold and gate leakage currents are the most dominant leakage sources in advanced technology nodes. Moreover, the contribution of sub-threshold leakage to the total leakage is much higher than that of gate leakage [47], especially at above room temperature operating conditions.

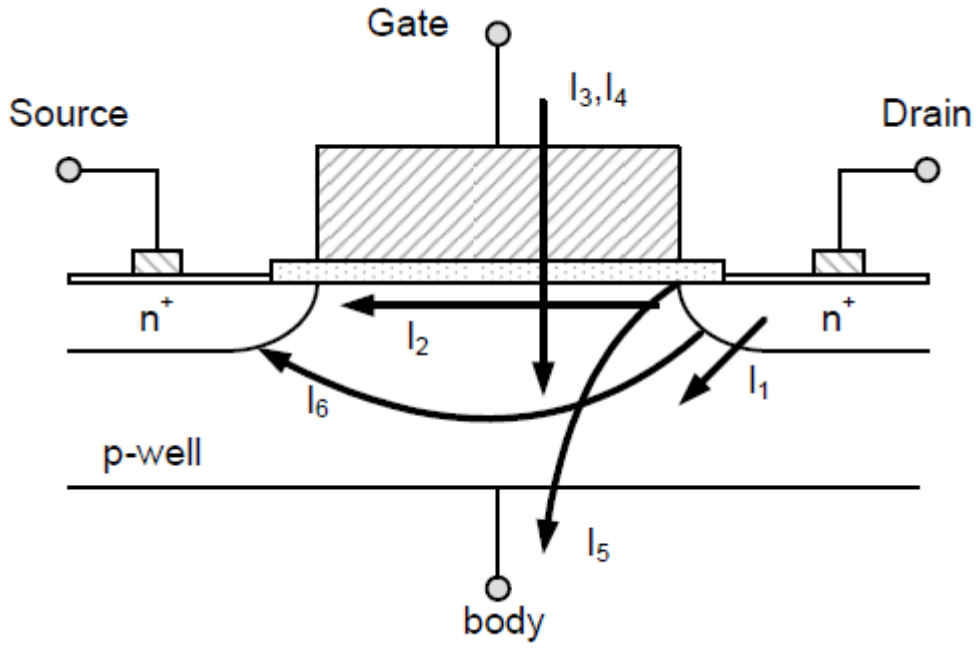


Figure 4.1: Leakage current sources in deep submicron devices [47]

4.1.1.1. Sub-threshold Leakage Current

The sub-threshold leakage current is defined in equation 4.1, and it is mainly current flowing from the drain to the source while in stand-by mode (when the Gate-to-Source voltage V_{GS} is less than V_{th})

$$I_{SUB} = \mu \frac{W_{eff}}{L_{eff}} \sqrt{\frac{q\varepsilon_{si}N_a}{2\varphi_s}} V_T^2 \left(1 - \exp\left(\frac{-V_{DS}}{V_T}\right)\right) \exp\left(\frac{V_{GS}-V_{th}}{nV_T}\right), \quad (4.1)$$

Where I_{SUB} is the sub-threshold current, μ is the electron surface mobility, q is the electron charge, ε_{si} is the silicon permittivity, N_a is the doping concentration in the substrate, $V_T = kT/q$ is the thermal voltage; k is the Boltzman constant, T is the absolute temperature, and φ_s is the surface potential, V_{DS} is the drain-source voltage,

and n is the sub-threshold swing parameter. Sub-threshold current is essentially influenced by V_{th}, the physical dimensions of the channel, gate oxide thickness, drain/source junction depth, channel/surface doping profile, and the supply voltage. For the sub-threshold leakage, the input state dependency is seen from equation 4.1 in the dependence of I_{SUB} on V_{GS} and $V_{DS.}$
There are two dominant factors affecting the input dependency of sub-threshold leakage current: DIBL and Body Effect [92]. Sub-threshold leakage current also has a significant dependence on the temperature.

4.1.1.2. Gate Leakage

The gate leakage current, which is considered as the second most critical leakage source after the sub-threshold leakage in nanometre technologies, exists in both the ON and OFF states of the CMOS transistors [26]. The gate leakage value strongly depends on both V_{GS} and V_{DS} that is large values of $|V_{GS}|$ and small values of V_{DS} cause a huge gate leakage accordingly. Figure 4.2 shows the two configurations of dominant gate leakage and their dependence on the input state. It should be noted that the gate leakage does not depend on the temperature and, as a result, it stays constant with the change chip temperature.

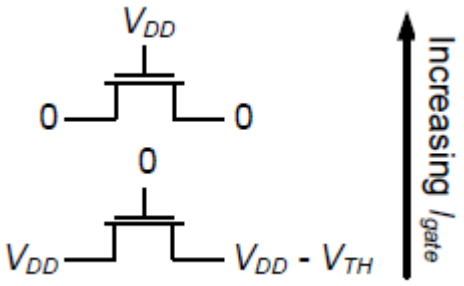


Figure 4.2: Gate leakage dominant states in FPGA pass-transistor device

4.1.2. Standby Leakage Reduction Techniques

Controlling the standby leakage current essentially could be done by several methods; it could be done topologically by introducing some special circuitry such as transistor stacking and minimum leakage vector (MLV) methods, reducing the total effective width is another way to control the leakage current by using efficient transistor structures, optimizing power-delay points, less concurrency in designs, or introducing asynchronous designs. Also adjusting transistor threshold voltages is another method for leakage control by introducing Dual-V_{th} designs [93] and different body biasing techniques. Here we review some techniques used to reduce the standby leakage current:

4.1.2.1. Multi-Threshold CMOS (MTCMOS)

This techniques aims at increasing the threshold voltage to reduce the leakage, along with other effective methods to increase the threshold voltage such as increasing the doping concentration, increasing the oxide gate thickness, and applying reverse body biasing.

Multi-threshold voltage [94-98] uses high-threshold (HVT) devices as sleep transistors while low-threshold (LVT) devices are used to implement the logic as shown in Figure 4.3. When the sleep transistor is in OFF state, the circuit's sub-threshold leakage current becomes only limited to that of the sleep transistor which is considerably low. Hence, the circuit benefits from the high performance of the LVT pull-down networks when the sleep transistor is turned ON, while limiting the circuit sub-threshold leakage current when the sleep transistor is turned OFF.

Practically speaking, one sleep transistor per gate is used, but larger granularities are also used which require fewer but larger sleep transistors. Normally, the NMOS sleep transistor is preferable because the on-resistance of NMOS is smaller than that of PMOS with the same width; hence, NMOS has size advantage over PMOS. This technique, however, comes with performance and area penalties.

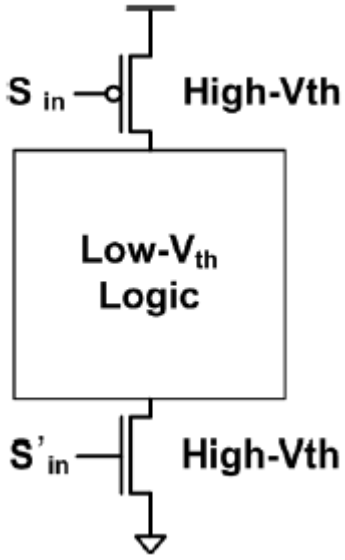


Figure 4.3: Multi-Threshold CMOS (MTCMOS)

4.1.2.2. Dual-Threshold Voltage

Dual-threshold voltage technique [99-101] assigns different threshold voltages to gates depending on whether a gate is on critical or non-critical path as shown in Figure 4.4. Low threshold voltage assigned along the critical path is used to maintain the performance, while high threshold voltage assigned along non-critical path reduces the leakage current.

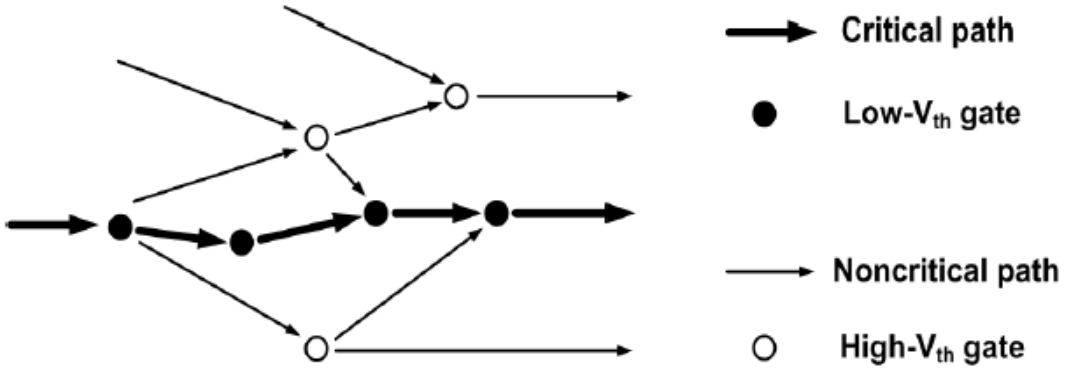


Figure 4.4: Dual-threshold voltage technique

4.1.2.3. Reverse Body Biasing

Reverse Body Biasing (RBB) [102-104] is an effective method of reducing the leakage in standby mode. This technique works by increasing the threshold voltages of MOS transistors (by making the substrate, or body, voltage higher than supply voltage for PMOS transistors and lower than ground for NMOS transistors, based on equation 4.2); reverse biasing the body-to-source junction of a MOS transistor widens the bulk depletion region and, in turn, increases the threshold voltage. RBB is applied to suppress the leakage current when circuits are in standby mode, and is deactivated to restore the transistors' nominal performance when circuits are in active mode as shown in Figure 4.5.

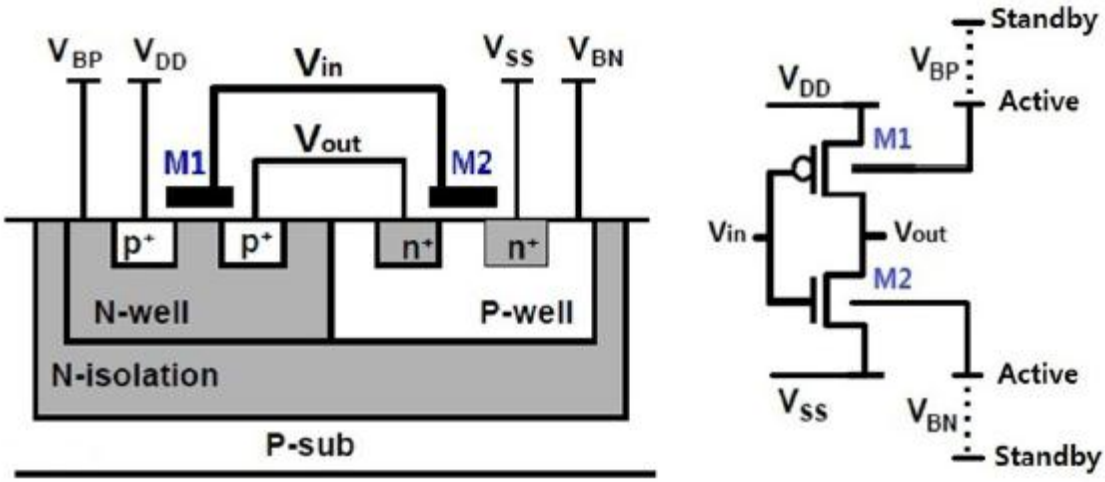


Figure 4.5: Reverse Body Biasing (RBB)

Adaptive body biasing (ABB) techniques have been introduced as more advanced extensions of the body biasing techniques [104-108] in order to alleviate the impact of

D2D and WID parameter variations on microprocessor leakage and frequency. These techniques aim at meeting the power and delay constraints in each die through post silicon tuning; forward body bias is applied to the slow and less leaky devices in order to boost the performance while reverse body bias is applied to the fast and highly leaky devices in order to reduce the leakage. Therefore, the impact of parameter variations is alleviated by post-silicon tuning, resulting in reducing the process variations effects as well as improving the total yield.

Table 4.1: Simulated Device Parameters

L (nm)	T_{fin} (nm)	H_{fin} (nm)	N_{fin}	V_{DD}
14	10	23	1	0.8

4.2. Simulation Methodology

All of these aforementioned leakage challenges have been vastly tackled by incorporating advanced multi-gate MOS devices such as fin field-effect transistors (FinFETs) into production. FinFETs have shown superior device performance at aggressively scaled device parameters as explained in chapter 3, and this made FinFETs compelling candidates to be used as alternatives to the conventional planer MOS transistors for sub 22nm nodes. Tri-gate FinFETs also help alleviating the SCE, hence renovating the chip production industry. And this is due to their high current drive capability, and superior sub-threshold leakage control, causing substantial power savings and leakage current reduction.

The possible sources for V_{th} variations in FinFETs represented by V_{th} in our study are the gate length, fin thickness, oxide thickness, RDF, and gate work function (Φ_m) which is the largest contributor in V_{th} variations as shown in Figure 4.6.

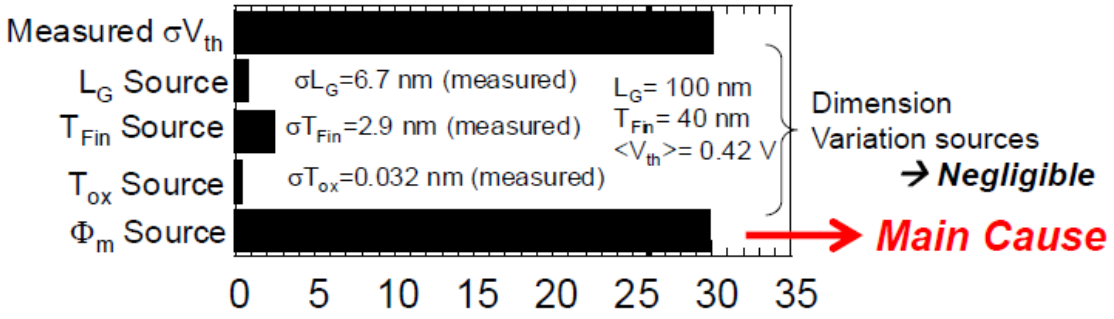


Figure 4.6: Vth variations sources in FinFET devices, σVth [mV]

In this work, the PTM-MG models are used for 14nm simulations. Table 4.1 reports the nominal device parameters supported by PTM-MG for 14nm technology node where (L) is the channel length, (V_{DD}) is the supply voltage, (T_{fin}) is fin thickness, (H_{fin}) is fin height, and (N_{fin}) is the fin count.

Table 4.2: Threshold Voltage Variations

Node(nm)		Threshold Voltage (mv)			
	Nominal	% change values (of nominal)			
		±6%	±12%	±18%	±24%
14	311	18.6	37.3	55.9	74.5

The leakage power is calculated by multiplying the average current drawn from the source, while enforcing standby mode, by the supply voltage value. This is calculated with V_{th} variations within range ±24% with 6% step of the nominal V_{th} for the 14nm technology node. The simulated threshold voltage values are reported in Table 4.2.

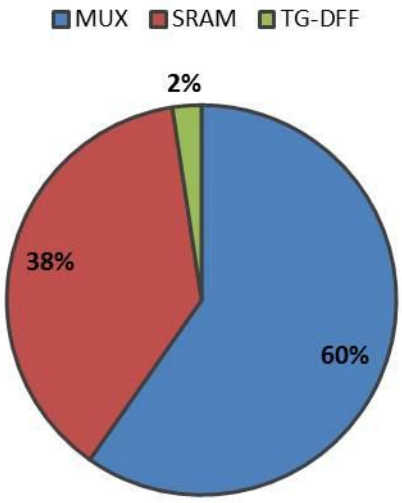


Figure 4.7: Leakage power segmentation

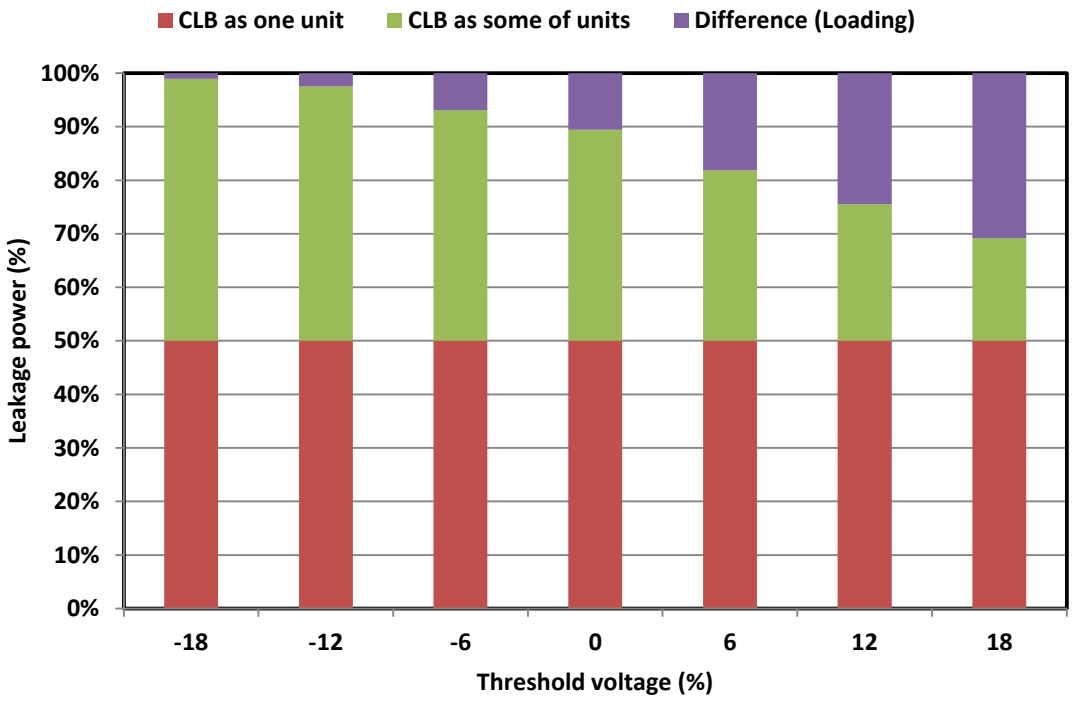


Figure 4.8: Leakage power consumed by the entire cluster vs the sum of leakage power of the units comprising the cluster with the difference representing the loading effect

4.3. Results and Discussions

4.3.1. Leakage Power Segmentation and Loading Effect

After simulating the FPGA cluster to evaluate the leakage power, the results show that the largest portion of leakage power (around 60%) is mainly consumed by the multiplexers embedded inside the cluster including the input multiplexers and those inside Look-Up Tables (LUTs) within Basic Logic Elements (BLEs) as shown in Figure 4.7. The simulation results showed also that the second largest consuming components for the leakage power are the SRAM slices inside the LUTs (around 38%). Then come the D Flip-Flops (DFFs) to consume the least amount of leakage power (around 2%). These segmentation percentages are mostly due to the count of the different components comprising the cluster.

Figure 4.8 also shows the difference between the cluster leakage power consumed by the whole cluster and the leakage power calculated separately by summing the amount of leakage power consumed by every unit comprising the cluster acting alone. The difference represents the loading as a result of adding these units to build the entire cluster.

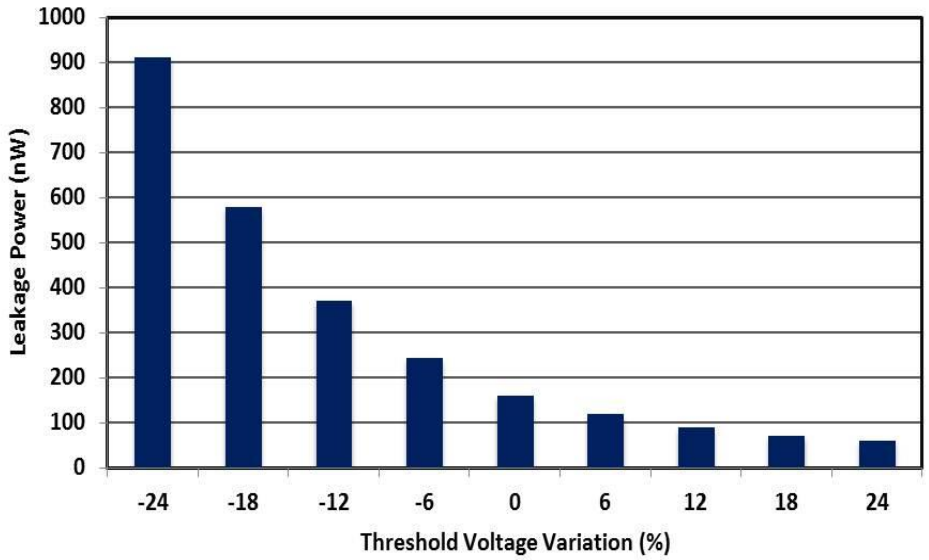


Figure 4.9: Leakage power variation with Vth variation

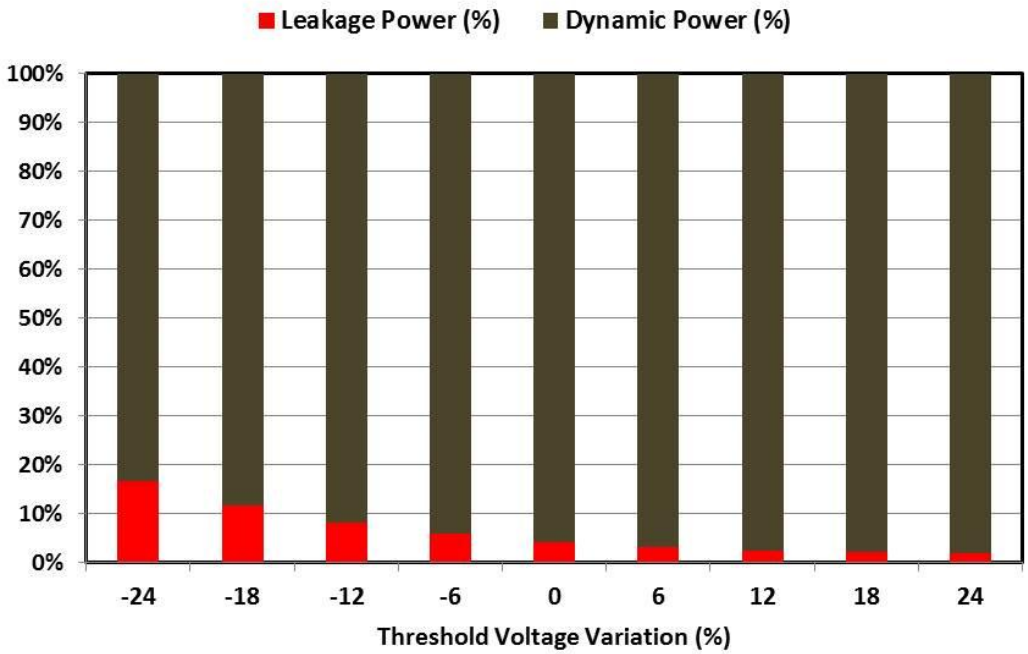


Figure 4.10: Dynamic and leakage power consumption percentages with Vth variation

4.3.2. Leakage Power Variation with V_{th} and Temperature

Regarding the simulation results for leakage power variation with V_{th} , Figure 4.9 shows that the leakage power decreases from 912nW to 60nW by increasing the threshold voltage percentages from -24% to 24% implying the log-normal distribution that reflects the relation between the sub-threshold leakage current and the threshold voltage. The stacked bar chart shown in Figure 4.10 reports the contribution percentages of both dynamic and leakage power consumptions to the total power consumption with the V_{th} variation, showing a larger contribution of the leakage power while decreasing the threshold voltage.

Since the digital circuits usually operate at high temperatures because of the power dissipation, it is important to study the temperature dependence of the sub-threshold leakage current as well. Figure 4.11 plots the temperature dependency of the leakage power showing an exponential trend with increasing the temperature with step 30^{o} C starting from -30^{o} C up to 120^{o} C.

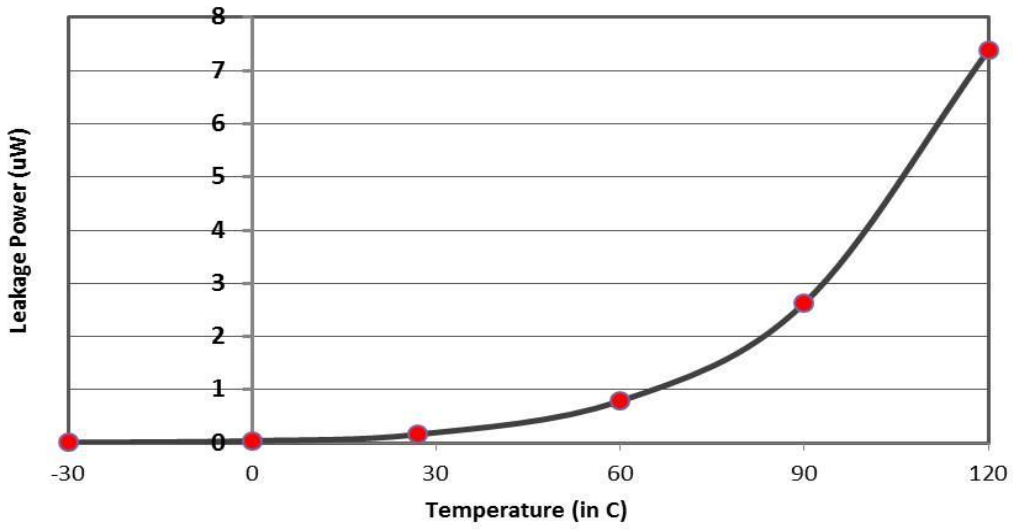


Figure 4.11: Temperature dependency of leakage power

4.4. Proposed Leakage Power Control Techniques

Due to the limitations of the predictive models capabilities used in our simulations, only limited number of solutions for leakage control are allowable for implementation. Three solutions are investigated and implemented in our work including transistor stacking, minimum leakage vector, and gate sizing. The leakage power and leakage power variation improvement versus V_{th} variation, along with the delay overhead, are shown in Figures 4.12, 4.13, and 4.14 for the three solutions and the original design respectively. Each solution will be subjected to a detailed explanation as follows:

4.4.1. Transistor Stacking

Transistor stacking is a well-known technique used to control the standby leakage power which mainly relies on the fact that the sub-threshold leakage current decreases when it flows through a switched-off stack of transistors connected in series fashion [110]. Similar to CMOS devices, stacking the transistors in FinFETs is either done by stacking NFET transistors in Pull-Down Networks (PDNs) or PFET transistors in Pull-Up Networks (PUNs). The only unit that contains PDN and PUN in our cluster design is the inverter, so this is the only circuit at which the stacking solution is applied. It is worth to mention that the transistor stacking solution might not be the optimal solution when it comes to the critical paths since introducing extra transistors would increase the delay of the logic path.

Upon applying the stacking solution on the inverter, the leakage now depends on the input vector applied. Table 4.3 reports the inverter leakage power values with both "1" and "0" inputs and with stacking single NFET and single PFET transistors. The results show that stacking a single PFET transistor yields less leakage compared to stacking an NFET transistor. That difference in the average leakage power between stacking NFET or PFET stems from the difference between the average mobility factors between NFET and PFET transistors with equal sizes. Stacking more than one PFET or NFET would definitely help in reducing the leakage power even more but at the expense of the circuit delay and area.

In order to roughly estimate the area overhead after adding an extra transistor for the stacking purpose, we estimated the transistor area from Intel's 14nm core M processor [111] which stated that the die area of the processor with 1.3 Billion transistors was $82mm^2$. Hence, the transistor area is roughly about $0.006um^2$. Assuming that the PFET area is equivalent to NFET area, our designed cluster contains 780 transistors + 204 Inverter units.

Before stacking, each inverter consists of two transistors. So, the total transistor count before stacking is 780 + (2*204) = 1188 transistors. Giving that transistor area is $0.006um^2$, so now the cluster area is around $7.128um^2$.

After applying the stacking solution, each Inverter now has three transistors. So, the total transistor count in our new design becomes 780 + (3*204) = 1392 transistors. Hence, giving a total area of $8.352um^2$. Calculating the area overhead due to stacking solution, that would give an overhead of ***17.2***%

Table 4.3: Leakage Power Values upon Stacking NFET and PFET

	NFET		PFET	
Input	0	1	0	1
Leakage Power	14pW	317.5pW	240.6pW	12.35pW
Average Leakage Power	165.75pW		126.5pW	

4.4.2. Minimum Leakage Vector (MLV)

This method aims at finding the input vector that minimizes the leakage power during the standby periods by maximizing the off-transistors count in the stacks across the whole circuit. Upon doing exhaustive simulation with all the possible input vectors, the results indicated that increasing the number of ones in the input vector increases the leakage current [112]. We could not report all the simulation results because we are having 11 inputs which means up to 2048 input vectors were tried to get the input vector with the least leakage. Consequently, the All-zeros input vector yielded the least leakage current drawn from the FPGA cluster. This vector not only reduced the leakage power but also its variation with threshold voltage around the nominal condition. That MLV is then being driven by means of inserting input 2-to-1 multiplexers prior to the FPGA cluster to enforce these input values while in standby mode. And these multiplexers insignificantly contributed in increasing the overall leakage values and the delay as well.

Since we have eight distinct inputs, we had to insert eight 2-to-1 multiplexers at the inputs to the cluster to apply the MLV solution. Since each 2-to-1 multiplexer contains eight transistors, we have now 64 extra transistor with extra area of $64*0.006um^2$ = $0.384um^2$, resulting in total design area of $7.512um^2$. Giving an area overhead due to the MLV solution of ***5.387***%

4.4.3. Gate Sizing

Another solution for leakage power control which is more towards the process ad design parameters optimization is gate sizing. This solution works by optimizing W_{eff} to find the optimum gate parameters that result in minimum leakage power and leakage power variation across with V_{th} variations, besides maintaining acceptable delay ranges. Since our design is FinFET-Based, this implies that we have three fin parameters to optimize per transistor; N_{fin}, H_{fin}, and T_{fin}. In our evaluation work, we ruled out N_{fin} optimization due to the introduced area overhead "(2 H_{fin} + T_{fin}) per extra fin" as increasing N_{fin} would boost the driving current, hence increasing the leakage power. Accordingly, we assumed a single fin for all transistors in our design. For the H_{fin}, changing its value in PTM models does not affect the results, they would still be the same as of H_{fin}'s default value of 23nm. Besides, H_{fin} is found to be less dominant in controlling the SCE compared to T_{fin} [113] since the leakage current occurs in the middle of the fin, and it increases as we increase the T_{fin} due to reduced control of the side gates over the channel, while the leakage decreases with decreasing the T_{fin} because the middle part of the fin will get more control from the side gates. In addition, and from fabrication point of view, H_{fin} is fixed for a process since it is limited by the etching technology, so varying H_{fin} is not really an option at a given node. Eventually, we come up with optimizing T_{fin} only for leakage power control.

For 14nm technology, T_{fin} value ranges from 8nm [111] up to 14nm, representing the corner values for fin thickness. So, we used Cadence ADE GXL global optimizer by sweeping over T_{fin} from 8nm to 14nm with step 2nm. And the objective functions which we applied the optimization for were leakage power, leakage power variation, and delay to be minimized. The optimization is done on the basic elements comprising the cluster like Inverter, 2-to-1 multiplexer, and SRAM cell. Figures 4.15 to 4.17, along with Tables 4.4 to 4.6, show the schematic of the basic units used in gate sizing solution, and the optimized values for Inverter, 2-to-1 multiplexer, and SRAM cell, respectively.

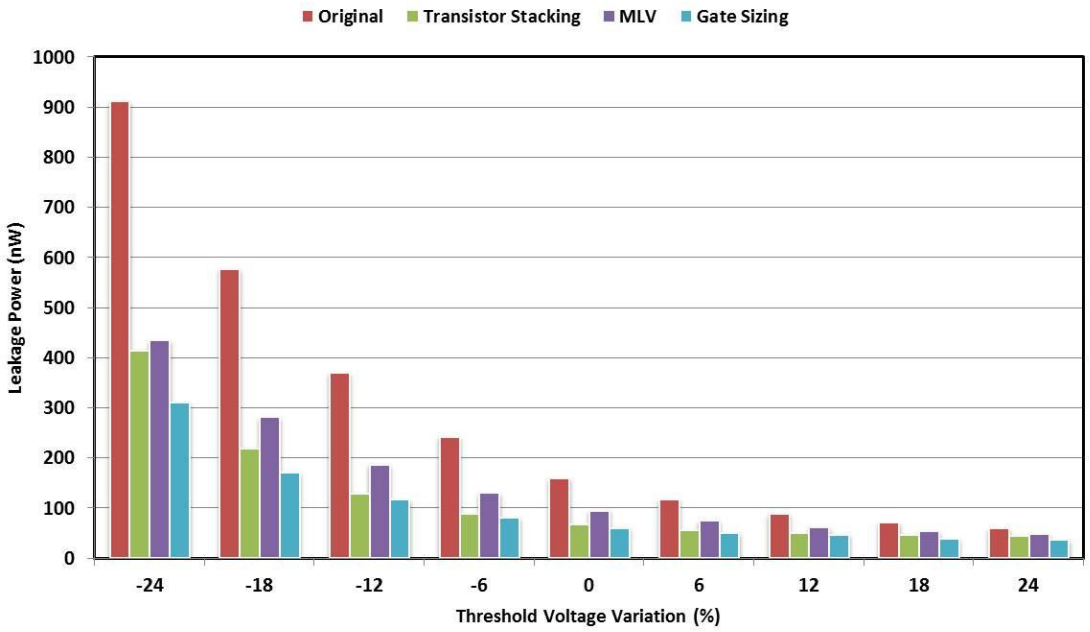


Figure 4.12: Leakage power with Vth variation for the three solutions

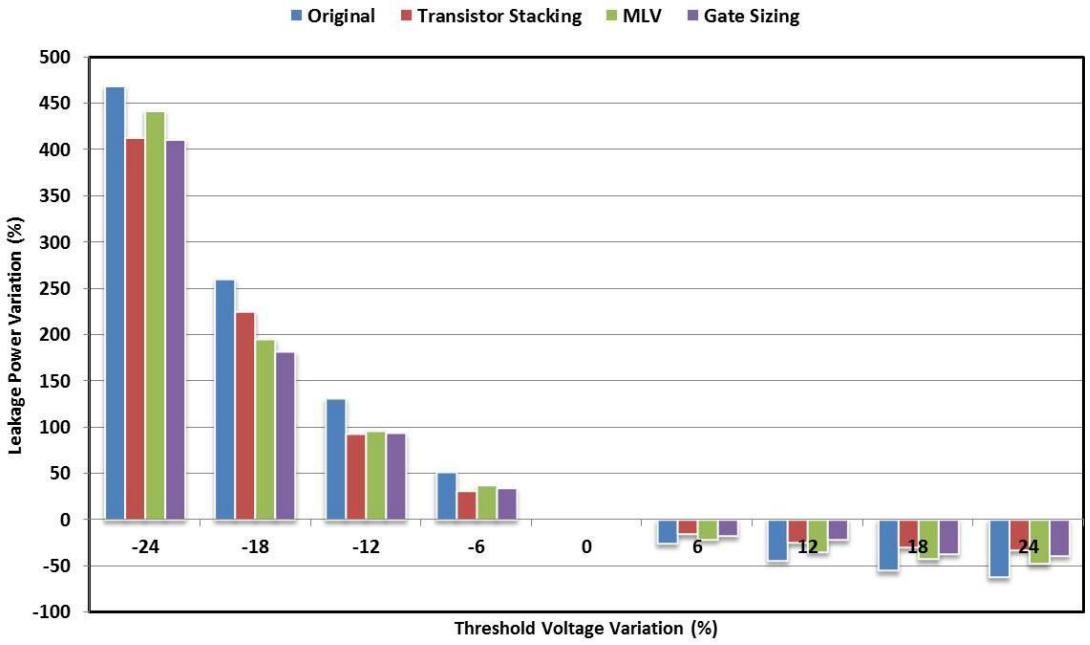


Figure 4.13: Leakage power variation with Vth variation for the three solutions

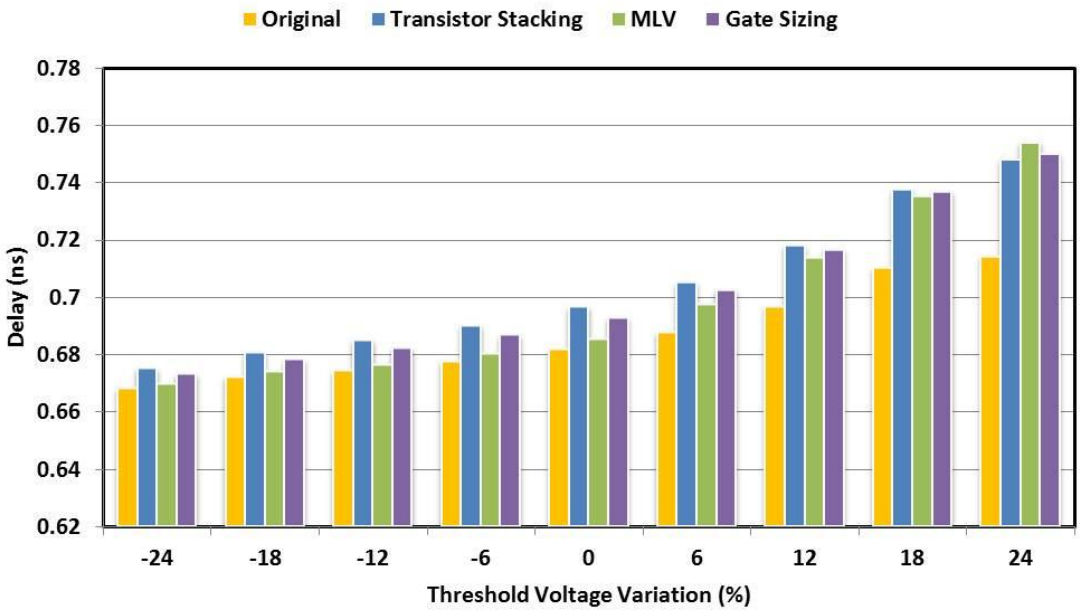


Figure 4.14: Delay overhead with Vth variation for the three solutions

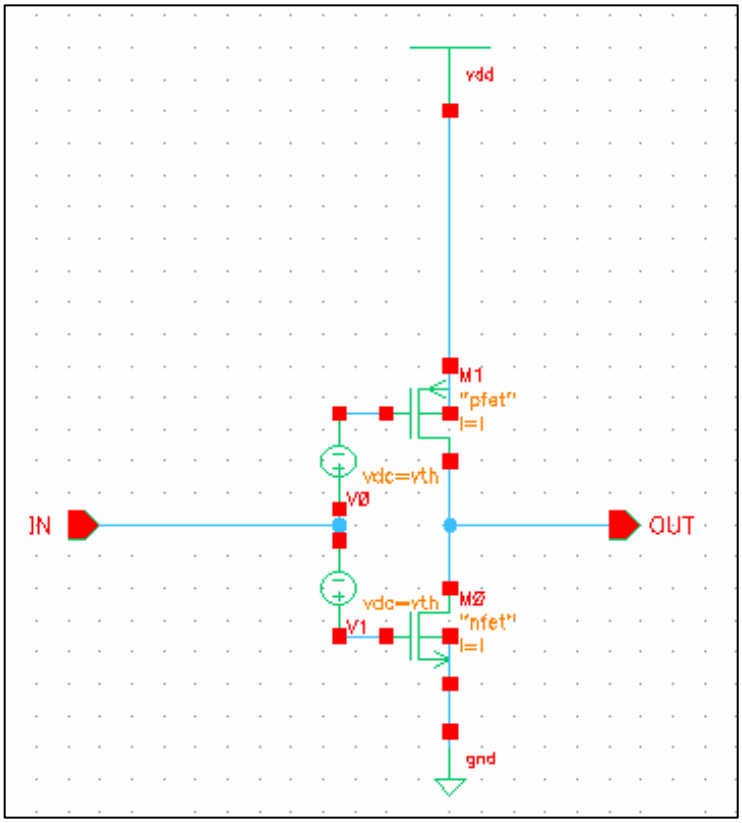


Figure 4.15: FinFET Inverter

Table 4.4: Tfin Optimized Values for Inverter

Transistor	T_{fin}
M1	8nm
M0	10nm

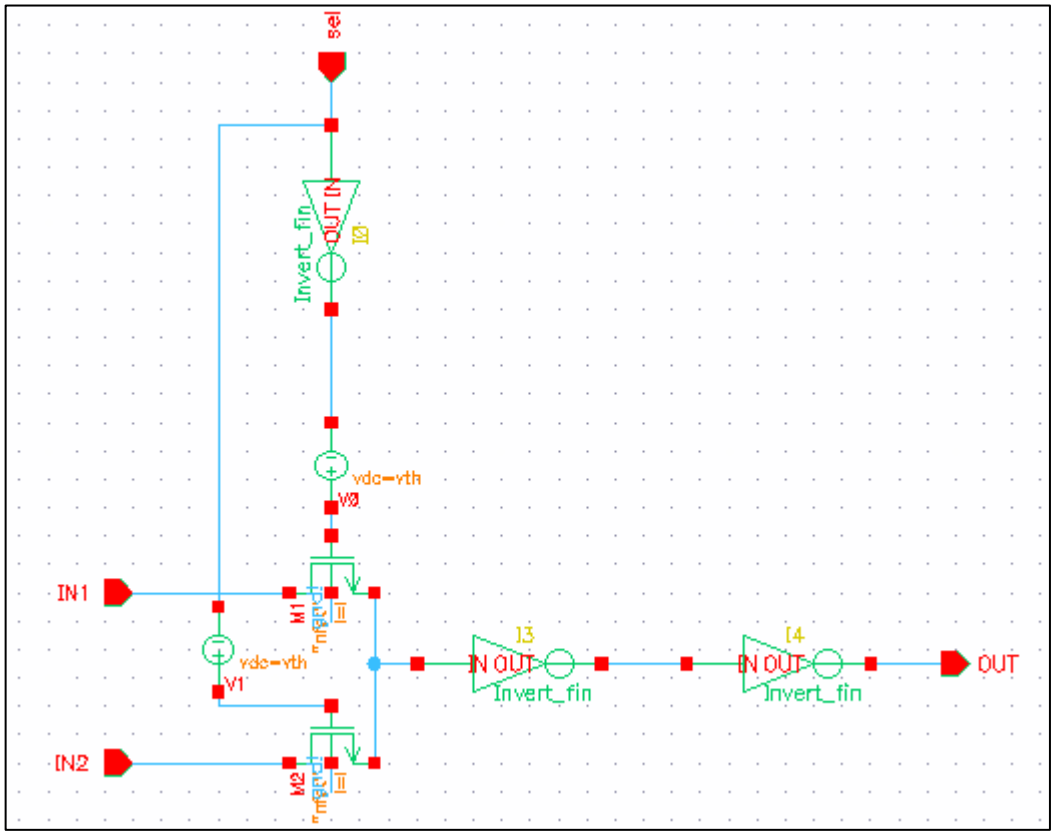


Figure 4.16: FinFET 2-to-1 multiplexer

Table 4.5: Tfin Optimized Values for 2-to-1 Multiplexer

Transistor	T_{fin}
M1	10nm
M2	8nm

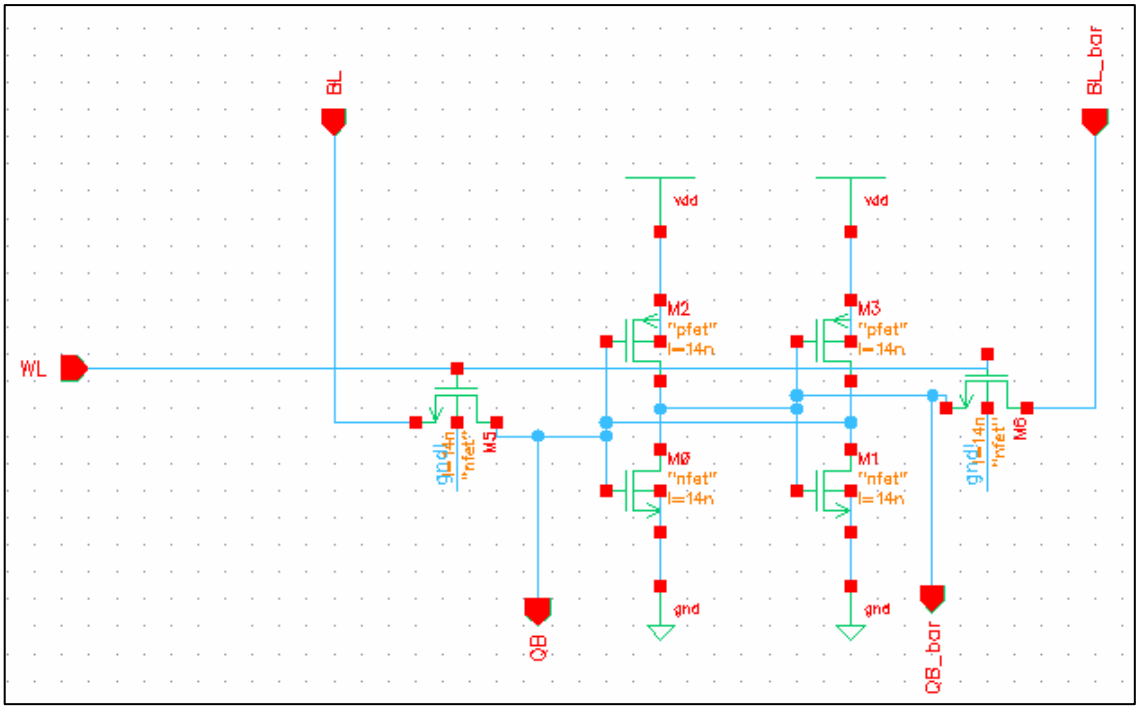


Figure 4.17: FinFET 6T SRAM cell

Table 4.6: Tfin Optimized Values for 6T SRAM

Transistor	T_{fin}
M0, M1 (Driver)	8nm
M2, M3 (Load)	10nm
M5, M6 (Access)	14nm

4.5. Conclusion

Leakage power is evaluated under threshold voltage and temperature variations, using PTM models for 14nm technology, for 2-Bit adder benchmark implemented by FinFET-Based FPGA cluster. Three solutions investigated and implemented in order to control the leakage power including transistor stacking, minimum leakage vector, and gate sizing through fin thickness optimization. The maximum improvement for both leakage power and leakage power variation, with the maximum delay overhead, are reported in Table 4.7 for the three solutions.

Table 4.7: Maximum Improvements and Delay Overhead for the Three Solutions

Solution	Leakage Power	Leakage Power Variation	Delay Overhead	Area Overhead
Transistor Stacking	65%	46%	4.54%	17.2%
MLV	52%	29%	5.6%	5.384%
Gate Sizing	70%	49%	5%	--

Discussion and Conclusions

In this research, the performance of FinFET-based FPGA cluster is evaluated with technology scaling. The cluster is then configured to be 2-Bit adder benchmark. The impact of V_{th} and temperature variations on basic performance metrics such as average power, delay, and Power-Delay-Product is reported. The simulation results, launched using PTM models for technology nodes 20nm down to 7nm, demonstrate that the variations of both the average power and the PDP with V_{th} increase with technology scaling. For the delay variation with V_{th} variation, it is not following a certain trend with the technology scaling. Design constraints are outlined for each performance metric for all the technology nodes included in this study in order to give the designers some useful design insights.

Leakage power is evaluated as well under V_{th} variations and temperature variations, using PTM models for 14nm technology, for the same 2-Bit adder benchmark used in the technology scaling study. Three solutions are implemented to control the leakage power, including transistor stacking, minimum leakage vector (MLV), and gate sizing by optimizing the fin thickness T_{fin}. The maximum improvement achieved for leakage power and leakage power variation, with the maximum delay overhead, are reported, showing that the gate sizing solution offers the largest improvements for both the average leakage power and leakage power variations.

Finally, our work suggests two future work directions: building an FPGA tile that utilizes the cluster designed and evaluated in our work, with the associated routing channels and Inter-cluster routing to be considered in simulations, and also extending the leakage power evaluation study for future FPGA technology nodes (e.g. 10nm and 7nm).

References

1. T.-C. Chen, "Where is CMOS going: trendy hype versus real technology," in *Proceedings of the International Solid-State Circuits Conference ISSCC*, pp. 22-28, 2006.
2. S. R. Nassif, "Modeling and analysis of manufacturing variations," in *Proceedings of IEEE Custom Integrated Circuits conference*, pp. 223-228, 2001.
3. H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: variability characterization and modeling for 65- to 90-nm processes," in *Proceedings of IEEE Custom Integrated Circuits conference*, pp. 593-599, 2005.
4. B. Wong, A. Mittal, Y. Cao, and G. W. Starr, *Nano-CMOS Circuit and Physical Design*. Wiley-Interscience, 2004.
5. The International Technology Roadmap for Semiconductors (ITRS). [Online]. Available: http://public.itrs.net
6. J. Tschanz, K. Bowman, and V. De, "Variation-tolerant circuits: circuit solutions and techniques," in *DAC 'o5: Proceedings of the 42nd annual conference on Design automation*, pp. 762-763, 2005.
7. D. Frank, R. Dennard, E. Nowak, P. Solomon, Y. Taur, and H. S. Wong, "Device scaling limits of Si MOSFETs and their application dependencies," *Proc. IEEE*, col. 90, no. 3, pp. 259-288, Mar 2001.
8. J. A. Croon, W. Sansen, and H. E. Maes, *Matching Properties of Deep Sub-Micron MOS Transistors*. Springer, 2005.
9. S. Sapatnekar, "*Timing," New York: Springer-Verlag*, 2004.
10. F. N. Najm, "On the need for statistical timing analysis," in *DAC '05: Proceedings of the 42nd annual conference on Design automation*, pp. 764-765, 2005.
11. A. Agrawal, D. Blaauw, and V. Zolotov, "Statistical timing analysis for intra-die process variations with spatial correlations," in *ICCAD '03: Proceedings of the 2003 IEEE/ACM international conference on Computer-aided design*, pp. 900-907, 2003.
12. A. Chandrakasan, W. J. Bowhill, and F. Fox, "Design of high performance microprocessor circuits," *IEEE Press*, 2001.
13. A. Srivastava, D. Sylvester, and D. Blaauw, *Statistical analysis and optimization for VLSI: Timing and Power (Series on Integrated Circuits and Systems)*. Springer, 2005.
14. S. Bhunia, and S. Mukhopadhyay, "Low-power variation-tolerant design in nanometer silicon," *Springer*, 2011.
15. Y. Taur and T. H. Ning, "Fundamentals of modern VLSI devices," *New York, NY, USA: Cambridge University Press*, 1998.
16. T. Mizuno, J. Okumtura, and A. Toriumi, "Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFETs," *IEEE Transactions on Electron Devices*, vol. 41, pp. 2216-2221, November 1994.

17. M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits (JSSC),* vol. 24, pp. 1433-1439, October 1989.

18. B. Razavi, "Design of analog CMOS integrated circuits," McGraw-Hill, 2000.

19. J. Luo, S. Sinha, Q. Su, J. Kawa, and C. Chiang, "An IC manufacturing yield model considering intra-die variations," *Proceedings of the IEEEE Design Automation conference (DAC'06),* pp. 749-754, 2006.

20. W. Liu, "MOSFET models for SPICE simulation including BSIM3v3 and BSIM4," *John Wiley & Sons, Inc,* 2001.

21. M. H. Abu-Rahma, and M. Anis, "A statistical design-oriented delay variation model accounting for within-die variations," *IEEE Transactions on Computer-aided Design (TCAD) of Integrated Circuits and Systems*, vol. 27, pp. 1983-1995, November 2009.

22. A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in Decananometer and Nanometer-scale MOSFETs," *IEEE Transactions on Electron Devices*, vol. 50, pp. 1837-1852, September 2003.

23. Y. Cheng and C. Hu, *MOSFET Modeling and BSIM User Guide*. Kluwer Academic Publishers, 1990.

24. M. Popovich, A. V. Mezhiba, and E. G. Friedman, "Power distribution networks with on-chip decoupling capacitors," *Springer*, 2008.

25. S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and micro-architecture," *Proceedings of the IEEE Design Automation Conference (DAC'03)*, pp. 338-342, 2003.

26. K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicron CMOS circuits," *Proceedings of IEEE*, vol. 91, pp. 305-327, February 2003.

27. Z. Chen, M. Johnson, L. Wei, and K. Roy, "Estimation of standby leakage power in CMOS circuits considering accurate modeling of transistors stacks," *Proceedings of the IEEE International Symposium on Low Power Electronics and Design (ISLPSD'98)*, pp. 239-244, 1998.

28. S. Narendra, V. De, S. Borkar, D. Antoniadis, and A. Chandrakasan, "Full-Chip sub-threshold leakage power prediction model for sub-0.18um CMOS," *IEEE Journal of Solid-State Circuits (JSSC),* vol. 39, pp. 501-510, February 2004.

29. S. Narendra, V. De, S. Borkar, D. Antoniadis, and A. Chandrakasan, "Full-Chip sub-threshold leakage power prediction model for sub-0.18um CMOS," *Proceedings of the International Symposium on Low Power Electronics and Design (ISLPED'02)*, pp. 19-23, 2002.

30. S. Nassif, "Waiting for the post-CMOS godot," *Keynote Speaker Slides in the Great Lakes Symposium on VLSI (GLSVLSI'11)*, pp. 1-40, 2011.

31. S. Borkar, T. Karnik, and V. De, "Design and reliability challenges in nanometer technologies," *Proceedings of the IEEE Design Automation Conference (DAC'04),* pp. 75-75, 2004.

32. S. H. Choi, B. C. Paul, and K. Roy, "Novel sizing algorithm for yield improvement under process variation in nanometer technology," *Proceedings of the IEEE Annual Design Automation Conference (DAC'04)*, pp. 454-459, 2004.

33. A. Agarwal, K. Chopra, and D. Blaauw, "Statistical timing based optimization using gate sizing," *Proceedings of the IEEE Conference on Design, Automation, and Test in Europe (DATE'05),* pp. 400-405, 2005.

34. H. Fukui, M. Hamaguchi, H. Yoshimura, H. Oyamatsu, F. Matsuoka, T. Noguchi, T. Hiaro, H. Abe, S. Onoda, T. Yamakawa, T. Wakasa, and T. Kamiya, "Comprehensive study on layout dependence of soft errors in CMOS latch circuits and its scaling trend for 65nm technology node and beyond," *Digest of Technical Papers in VLSI Circuits Symposium*, pp. 222-223, 2005.

35. S. M. Jahinuzzaman, M. Sharifkhani, and M. Sachdev, "Investigation of process impact on soft errors susceptibility of nanometric SRAMs using a compact critical charge model," *Proceedings of the IEEE International Symposiums on Quality Electronic Design (ISQED'08),* pp. 207-212, 2008.

36. T. Chen and S. Naffziger, "Comparison of Adaptive Body Bias (ABB) and Adaptive Supply Voltage (ASV) for improving delay and leakage under the presence of process variation," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems*, vol. 11, pp. 888-899, October 2003.

37. B. H. Calhoun and A. P. Chandrakasan, "Static noise margin variation for sub-threshold SRAM in 65nm CMOS," *IEEE Journal of Solid-State Circuits (JSSC),* vol. 41, pp. 1673-1679, July 2006.

38. F. Frustaci, P. Corsonello, S. Perri, and G. Cocorullo, "High-performance noise tolerant circuit techniques for CMOS dynamic logic," *IET Journal of Circuits, Devices, and Systems,* vol. 2, pp. 537-548, December 2008.

39. K. Bowman, S. Duvall, and J. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for Gigascale integration," *IEEE Journal of Solid-State Circuits (JSSC),* vol. 37, pp. 183-190, February 2002.

40. K. Bowman, S. Duvall, and J. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution," *Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC'01)*, pp. 278-279, 2001.

41. K. Bowman and J. Meindl, "Impact of within-die parameter fluctuations of future maximum clock frequency distributions," *Proceedings of the IEEE Conference on Custom Integrated Circuits (CICC'01),* pp. 229-232, 2001.

42. A. Datta, S. Bhunia, S. Mukhopadhyay, N. Banerjee, and K. Roy, "Statistical modeling of pipeline delay and design of pipeline under process variation to enhance yield in sub-100nm technologies," in *DATE'05: Proceedings of the conference on Design, Automation and Test in Europe*, pp. 926-931, 2005.

43. B. Zahiri, "Structured ASICs: Opportunities and challenges," in *Proceedings of Intl. Conf. on Computer Design*, pp. 404-409, 2003.

44. R. R. Taylor and H. Schmit, "Creating a power-aware structured ASIC," in *Proceedings of Intl. Symp. On Low Power Electronics and Design*, pp. 74-77, 2004.

45. K. J. Han, N. Chan, S. Kim, B. Leung, V. Hecht, B. Cronquist, D. Shum, A. Tilke, L. Pescini, M. Stiftinger, and R. Kakoschke, "Flash-based Field Programmable Gate Array technology with deep trench isolation," in *Proceedings of IEEE Custom Integrated Circuits Conf.*, pp. 89-91, 2007.

46. S. D. Brown, "An overview of technology, architecture and CAD tools for programmable logic devices," in *Proceedings of IEEE Custom Integrated Circuits Conf.*, pp. 69-76, 1994.

47. J. Greene, E. Hamdy, and S. Beal, "Antifuse Field Programmable Gate Arrays," in *Proceedings of IEEE*, vol. 81, no. 7, pp. 1042-1056, July 1993.

48. H. Hassan, M. Anis, and M. Elmasry, "Leakage-aware placement for FPGAs," in *Proc. ACM International Symposium on FPGAs*, pp. 267 (abstract – poster), 2005.

49. E. Ahmed, J. Rose, "The effect of LUT and cluster size on deep-submicron FPGA performance and density," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems* – Special section on the 2002 *International Symposium on Low-Power Electronics and Design (ISLPED)*

50. Monther Abusultan, Sunil P. Khatri, "A comparison of FinFET based FPGA LUT designs," *Proceedings of the 24th edition of the Great Lakes Symposium on Very Large Scale Integration conference (GLSVLSI'14)*, pp. 353-358.

51. K. J. Kuhn, "CMOS scaling for the 22nm node and beyond: Device physics and technology," in *Proceedings of the International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA'11)*, pp. 1-2, April 2011.

52. C. Hu, "Gate oxide scaling limits and projection," in *Proceedings of the IEEE International Electron Devices Meeting*, pp. 319-322, December 1996.

53. Y.-C. Yeo, T.-J. King, and C. Hu, "MOSFET gate leakage modeling and selection guide for alternative gate dielectrics based on leakage considerations," *IEEE Transactions on Electron Devices*, vol. 50, no. 4, pp. 1027-1035, 2003.

54. J. Chen, T. Y. Chan, I. C. Chen, P. K. Ko, and C. Hu, "Subbreakdown drain leakage current in MOSFET," *Electron device letters*, vol. 8, no. 11, pp. 515-517, 1987.

55. T. Skotnicki, J. A. Hutchby, T.-J. King, H.-S. P. Wong, and F. Boeuf, "The end of CMOS scaling: toward the introduction of new materials and structural changes to improve MOSFET performance," *IEEE Circuits and Devices Magazine*, vol. 21, no. 1, pp. 16-26, 2005.

56. H.-S. P. Wong, D. J. Franks, and P. M. Solomon, "Device design considerations for double-gate, ground-plane, and single-gated ultra-thin SOI MOSFET's at the 25nm channel length generation," in *Proceedings of the IEEE International Electron Devices Meeting (IEDM'98)*, pp. 407-410, San Francisco, Calif, USA, December 1998.

57. P. M. Solomon, K. W. Guarini, Y. Zhang et al., "Two gates are better than one," *IEEE Circuits and Devices Magazine*, vol. 19, no. 1, pp. 48-62, 2003.

58. K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFET's," *IEEE Transactions on Electron Devices*, vol. 40, no. 12, pp. 2326-2329, 1993.

59. E. J. Nowak, I. Aller, T. Ludwig et al., "Turning silicon on its edge [double gate CMOS/FinFET technology]," *IEEE Circuits and Devices Magazine,* vol. 20, no. 1, pp. 20-31, 2004.

60. D. Hisamoto, W.-C. Lee, J. Kedzierski et al., "FinFET—a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Transactions on Electron Devices*, vol. 47, no. 12, pp. 2320–2325, 2000.

61. B. Yu, L. Chang, S. Ahmed et al., "FinFET scaling to 10 nm gate length," in *Proceedings of the IEEE International Devices Meeting (IEDM '02)*, pp. 251–254, San Francisco, Calif, USA, December 2002.

62. S. Tang, L. Chang, N. Lindert et al., "FinFET—a quasiplanar double-gate MOSFET," in *Proceedings of the International of Solid-State Circuits Conference,* pp. 118–119, February 2001.

63. M. Guillorn, J. Chang, A. Bryant et al., "FinFET performance advantage at 22 nm: an AC perspective," in *Proceedings of the Symposium on VLSI Technology Digest of Technical Papers (VLSIT '08)*, pp. 12–13, June 2008.

64. F.-L. Yang, D.-H. Lee, H.-Y. Chen et al., "5nm-gate nanowire FinFET," in *Proceedings of the Symposium on VLSI Technology—Digest of Technical Papers*, pp. 196–197, June 2004.

65. X. Huang, W.-C. Lee, C. Kuo et al., "Sub 50-nm FinFET: PMOS," in *Proceedings of the IEEE International Devices Meeting (IEDM '99)*, pp. 67–70, Washington, DC, USA, December 1999.

66. J.-P. Colinge, *FinFETs and Other Multi-Gate Transistors*, Springer, New York, NY, USA, 2008.

67. T.-J. King, "FinFETs for nanoscale CMOS digital integrated circuits," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '05)*, pp. 207–210, November 2005.

68. J. B. Chang, M. Guillorn, P. M. Solomon et al., "Scaling of SOI FinFETs down to fin width of 4 nm for the 10 nm technology node," in *Proceedings of the Symposium on VLSI Technology, Systems and Applications (VLSIT '11)*, pp. 12–13, June 2011.

69. C. Auth, "22-nm fully-depleted tri-gate CMOS transistors," in *Proceedings of the IEEE Custom Integrated Circuits Conference (CICC '12)*, pp. 1–6, San Jose, Calif, USA, September 2012.

70. C.-H. Lin, J. Chang, M. Guillorn, A. Bryant, P. Oldiges, and W. Haensch, "Non-planar device architecture for 15 nm node: FinFET or trigate?" in *Proceedings of the IEEE International Silicon on Insulator Conference (SOI '10)*, pp. 1–2, October 2010.

71. K. Lee, T. An, S. Joo, K.-W. Kwon, and S. Kim, "Modeling of parasitic fringing capacitance in multifin trigate FinFETs," *IEEE Transactions on Electron Devices,* vol. 60, no. 5, pp. 1786–1789, 2013.

72. M. Alioto, "Comparative evaluation of layout density in 3T, 4T, and MT FinFET standard cells," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems*, vol. 19, no. 5, pp. 751–762, 2011.

73. N. Collaert, M. Demand, I. Ferain et al., "Tall triple-gate devices with TiN/HfO_2 gate stack," in *Proceedings of the Symposium on VLSI Technology*, pp. 108–109, June 2005.

74. T.-S. Park, H. J. Cho, J. D. Choe et al., "Characteristics of the full CMOS SRAM cell using body-tied TG MOSFETs (Bulk FinFETS)," *IEEE Transactions on Electron Devices*, vol. 53, no. 3, pp. 481–487, 2006.

75. H. Kawasaki, K. Okano, A. Kaneko et al., "Embedded bulk FinFET SRAM cell technology with planar FET peripheral circuit for hp32 nm node and beyond," in *Proceedings of the Symposium on VLSI Technology (VLSIT '06*), pp. 70–71, June 2006.

76. S. Y. Kim and J. H. Lee, "Hot carrier-induced degradation in bulk FinFETs," *IEEE Electron Device Letters*, vol. 26, no. 8, pp. 566–568, 2005.

77. J. Markoff, "Intel increases transistor speed by building upward," http://www.nytimes.com/2011/05/05/science/05chip.html, May 2011.

78. J. Kedzierski, D. M. Fried, E. J. Nowak et al., "High-performance symmetric-gate and CMOS-compatible Vt asymmetric-gate FinFET devices," in *Proceedings of the IEEE International Electron Devices Meeting (IEDM '01)*, pp. 437–440, December 2001.

79. L. Mathew, M. Sadd, B. E. White, and et al, "FinFET with isolated n+ and p+ gate regions strapped with metal and polysilicon," in *Proceedings of the IEEE International SOI Conference Proceedings*, pp. 109–110, October 2003.

80. A. N. Bhoj and N. K. Jha, "Design of logic gates and flip-flops in high-performance FinFET technology," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems*, vol. 21, no. 11, pp. 1975–1988, 2013.

81. S. Xiong and J. Bokor, "Sensitivity of double-gate and FinFETdevices to process variations," *IEEE Transactions on Electron Devices*, vol. 50, no. 11, pp. 2255–2261, 2003.

82. X. Wang, A. R. Brown, B. Cheng, and A. Asenov, "Statistical variability and reliability in nanoscale FinFETs," in *Proceedings of the IEEE International Electron Devices Meeting (IEDM '11)*, pp. 541–544, Washington, DC, USA, December 2011.

83. E. Baravelli, L. de Marchi, and N. Speciale, "VDD scalability of FinFET SRAMs: robustness of different design options against LER-induced variations," *Solid-State Electronics*, vol. 54, no. 9, pp. 909–918, 2010.

84. P. Mishra, A. N. Bhoj, and N. K. Jha, "Die-level leakage power analysis of FinFET circuits considering process variations," in *Proceedings of the 11th International Symposium on Quality Electronic Design (ISQED '10)*, pp. 347–355, March 2010.

85. T. Matsukawa, S. O'uchi, K. Endo et al., "Comprehensive analysis of variability sources of FinFET characteristics," in *Proceedings of the Symposium on VLSI Technology (VLSIT '09)*, pp. 118–119, Honolulu, Hawaii, USA, June 2009.

86. S. Chaudhuri and N. K. Jha, "3D vs. 2D analysis of FinFET logic gates under process variations," in *Proceedings of the 29th IEEE International Conference on Computer Design (ICCD '11)*, pp. 435–436, Amherst, Mass, USA, November 2011.

87. S. M. Chaudhuri and N. K. Jha, "3D vs. 2D device simulation of FinFET logic gates under PVT variations," *ACM Journal on Emerging Technologies in Computing Systems,* vol. 10, no. 3, 2014.

88. J. H. Choi, J. Murthy, and K. Roy, "The effect of process variation on device temperature in FinFET circuits," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '07),* pp. 747–751, November 2007.

89. Predictive Technology Model (PTM). http://ptm.asu.edu/.

90. M. Mohie, H. Mostafa, H. A. H. Fahmy, Y. Ismail, and H. Abdelhamid, "Performance evaluation of FinFET-Based FPGA cluster under threshold voltage variation," *New Circuits and Systems Conference (NEWCAS), 2015 IEEE 13th International*, pp. 1-4, June 2015.

91. P. Mishra, A. Bhoj, and N. Jha, "Die-level leakage power analysis of FinFET circuits considering process variations," in *Proceedings of the 11th International Symposium on Quality Electronic Design (ISQED),* 2011.

92. Jeon, Heung Jun; Kim, Yong-Bin; and Choi, Minsu, "Standby Leakage Power Reduction Technique for Nanoscale CMOS VLSI Systems". Faculty Research & Creative Works. Paper 1065, 2010.

93. F. Li, Y. Lin, L. He, and J. Cong, "Low-power FPGA using pre-defined Dual-Vdd/Dual-Vt fabrics," in Proceedings of ACM Intl. Symp. On Field Programmable Gate Arrays, pp. 42-50, 2004.

94. N. Sirisantana, L. Wei, and K. Roy, "High-performance low-power CMOS circuits using multiple channel length and multiple oxide thickness," in *Proceedings of the International Conference on Computer Design*, pp. 227-232, 2000.

95. S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, "I-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS," in *Proceedings of IEEE Journal of Solid-State Circuits*, vol. 30, pp. 847-854, August 1995.

96. S. Mutoh, S. Shigematsu, Y. Matsuya, H. Fukuda, T. Kaneko, and J. Yamada, "A 1-V multithreshold-voltage CMOS digital signal processor for mobile phone application," in *Proceedings of IEEE Journal of Solid-State Circuits*, vol. 31, no. 11, pp. 1795-1802, November 1996.

97. S. Shigematsu, S. Mutoh, Y. Matsuya, Y. Tanabe, and J. Yamada, "A 1-V high-speed MTCMOS circuit scheme for power-down application circuits," in *Proceedings of IEEE Journal of Solid-State Circuits,* vol. 32, no. 6, pp. 861-869, June 1997.

98. J. Kao, S. Narendra, and A. Chandrakasan, "MTCMOS hierarchical sizing based on mutual exclusive discharge patterns," in *Proceedings of Design Automation Conference,* vol. 19, no. 19, pp. 495-500, June 1998.

99. L. Wei, Z. Chen, M. Johnson, K. Roy, and V. De, "Design and optimization of low voltage high performance dual threshold CMOS circuits," in *Proceedings of Design Automation Conference,* vol. 19, no. 19, pp. 489-494, June 1998.

100. P. Pant, R. Roy, and A. Chattejee, "Dual-threshold voltage assignment with transistor sizing for low power CMOS circuits," in *Proceedings of IEEE on Very Large Scale Integration (VLSI) Systems*, vol. 9, no. 2, pp. 390-394, April 2001.

101. M. Kethar and S. Sapatnekar, "Standby power optimization via transistor sizing and dual threshold voltage assignment," in *IEEE/ACM International Conference on Computer Aided Design (ICCAD 2002)*, vol. 10, no. 14, pp. 375-378, November 2002.

102. A. Keshavarzi, S. Ma, S. Narendra, B. Bloechel, K. Mistry, T. Ghani, S. Borkar, and V. De, "Effectiveness of reverse body bias for leakage control in scaled dual Ct CMOS ICs," in *Proceedings of the 2001 International Symposium for Low Power Electronics and Design*, pp. 207-211, 2001.

103. S. Martin, K Flautner, T. Mudge, and D. Blaauw, "Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads", in *Proceedings of IEEE/ACM on International Conference for Computer Aided Design (ICCAD'02),* vol. 10, no. 14, pp. 721-725, November 2002.

104. L. Yan, J. Luo, and N. Jha, "Joint dynamic voltage scaling and adaptive body biasing for heterogeneous distributed real-time embedded systems," in *Proceedings of IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol. 24, no. 7, pp. 1030-1041, July 2005.

105. J. Tschanz, J. Kao, S. Narendra, R. Nair, D. Antoniadis, A. Chandrakasan, and V. De, "Adaptive body bias for reducing impacts of die-ti-die and within-die parameter variations on microprocessor frequency and leakage," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 11, pp. 1396-1402, November 2002.

106. J. Tschanz, N. S. Kim, S. Dighe, J. Howard, G. Ruhl, S. Vagna, S. Narendra, Y. Hoskote, H. Wilson, C. Lam, M. Shuman, C. Tokunaga, D. Somasekhar, S. Tang, D. Finan, T. Karnik, N. Borkar, N. Kurd, and V. De, "Adaptive frequency and biasing techniques for tolerance to dynamic temperature-voltage variations and aging," *IEEE International Journal of Solid-State Circuits (ISSCC 2007)*, vol. 11, no. 15, pp. 292-604, February 2007.

107. G. Gammie, A. Wang, M. Chau, S. Gururajaro, R. Pitts, F. Jumel, S. Engel, P. Royannez, R. Lagerquist, H. Mair, J. Vaccani, G. Baldwin, K. Heragu, R. Mandal, M. Clinton, D. Arden, and U. Ko, "A 45nm 3.5G baseband-and-multimedia application processor using adaptive body-bias and ultra-low-power techniques," *IEEE International Solid-State Conference (ISSCC 2008)*, vol. 3, no. 7, pp. 258-611, February 2008.

108. S. Kulkarni, D. Sylvester, and D. Blaauw, "Design-time optimization of post-silicon tuned circuits using adaptive body bias," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 3, pp. 481-494, March 2008.

109. H. Mostafa, M. Anis, and M. Elmasry, "A novel low area overhead direct adaptive body bias (D-ABB) circuit for die-to-die and within-die variations compensation," *IEEE Transactions on Very Large Scale Integration (VLSI)*, vol. 19, no. 10, pp. 1848-1860, October 2011.

110. J. W. Chun, "Methodology for standby leakage power reduction in nanometer-scale CMOS circuits," *Syracuse University*, December 2012.

111. Mark Bohr, "14nm process technology: opening new horizons"

112. S. Srinivasan, A. Gayasen, N. Vijaykrishnan, and T. Tuan, "Leakage control in FPGA routing fabric," in *Proceedings of the 2005 Asia and South Pacific Design Automation Conference (ASP-DAC '05),* pp. 661-664, 2005.

113. G. Saini, A. K. Rana, "Physical scaling limits of FinFET structure: A simulation study," in *Proceedings of International Journal of VLSI and Communication Systems (VLSICS),* vol. 2, no. 1, March 2011.

114. J. Kayalieros et al., "Tri-gate transistor architecture with high-k gate dielectrics, metal gates and strain engineering," in *TVLSI*, pp. 50-51, 2006.

115. C. Wu et al., "High performance 22/20nm FinFET CMOS devices with advanced high-k/metal gate scheme," in *IEDM*, vol. 27, no. 1, pp. 1-4, December 2010.

116. C.-Y. Chang et al., "A 25nm gate length FinFET transistor module for 32nm node," in *IEDM*, pp. 1-4., December 2009.

117. T. Yamashita et al., "Sub-25nm FinFET with advanced fin formation and short channel effect engineering," in *TVLSI*, pp. 14-15, June 2011.

118. S. Sinha, G. Yeric, V. Chandra, B. Cline, and Y. Cao, "Exploring sub-20nm FinFET design with predictive technology models," in *Proceedings of Design Automation Conference (DAC '12) 49th ACM/EDAC/IEEE*, pp. 283-288, June 2012.

119. S. Sinha, B. Cline, G. Yeric, V. Chandra, and Y. Cao, "Design benchmarking to 7nm with FinFET predictive technology models," in Proceedings of ISLPED, pp. 15-20, 2012.

120. Z. Jaksic, R. Canal, "Effects of FinFET technology scaling on 3T and 3T1D cell performance under process and environmental variations," *3rd workshop on Resilient Architectures in Conjunction with the 45th Annual IEEE/ACM International Symposium on Microarchitecture*, Vancouver, December 2012.

121. E. Amat, C. G. Almudever, N. Aymerich, R. Canal, and A. Rubio, "Impact of FinFET technology introduction in the 3T1D-DRAM memory cell," IEEE Transactions on Device and Materials Reliability, vol. 13, no. 1, pp. 287-292, March 2013.

122. O. Abdelkader, H. Mostafa, H. Abdelhamid, A. Soliman, "Impact of technology scaling on the maximum energy point for FinFET based flip flops," IEEE International Conference on Electronics, Circuits, and Systems (ICECS 2015), pp. 462-465, 2015.

Appendix A: PTM models

Predictive Technology Model for Multi Gate (PTM-MG) [89] model cards for sub-22nm multi-gate transistors have been developed based on MOSFET scaling theory, the 2011 ITRS roadmap and early stage silicon data from published results [5]. PTM-MG used the published results from leading foundries such as TSMC, IBM, and Intel [114-117] to extract the fitting PTM parameters such as sub-threshold slope (S) and DIBL. However, PTM-MG models do not have complete information about the fabricated devices in [114-117]. They are introduced by fine tuning both primary parameters (such as gate length, fin pitch, fin thickness, and fin height) and secondary parameters (gate work function (Φ_m), DIBL coefficient, source-drain channel coupling, and channel doping) [117] to match both the on-current and the off-current of these published results. Table A.1 lists the key technology parameters as reported by ITRS. Table A.2 lists the verifications results of PTM-MG models with published measurements results from renowned leading foundries [118]. Some studies have discussed the analysis of PTM models on some circuits with technology scaling [118-121]. A simulation study using PTM models for ring oscillator and basic logic gates is discussed [119].

For future technology nodes (beyond 14nm), PTM-MG models are developed using ITRS as a reference. The off-current for 14nm technology node and below is expected to be (I_{OFF}=0.01nA/um for LSTP and 100nA/um for HP) according to ITRS trends [5]. PTM-MG models normalized per effective width (W_{eff}) for a constant off-current (I_{OFF}=0.1nA/um for LSTP and 100nA/um for HP).

The difference between both PTM off-current and ITRS off-current impact on transmission gate flip-flop (TG-FF) basic metrics, including average power, delay, and Power-Delay Product, is evaluated and plotted in Figure A.1 to A.3. These figures show that the simulation results using nominal PTM-MG parameters exhibit slight deviation from fabricated devices with ITRS off-current. For example, 7nm PTM TG-FF has 5% deviation in the average power from similar device with ITRS off-current.

Table A.1: Key Technology Parameters

Year	2012	2014	2016	2018	2020
Technology	20nm	16nm	14nm	10nm	7nm
M1 Pitch (nm)	64	48	38	30	24
Lg (nm)	24	21	18	14	11
V_{DD} (V)	0.9	0.85	0.8	0.75	0.7
T_{FIN} (nm)	15	12	10	8	6.5
H_{FIN} (nm)	28	26	23	21	18
W_{EFF}(nm)	71	64	54	51	42.5
Fin Pitch (nm)	60	42	32	28	22
3D Factor	1.2	1.524	1.75	1.78	1.75

Table A.2: PTM-MG Verification

Data source	[16]	[20]	[19]	[21]
Foundry	Intel	TSMC	TSMC	IBM
Lg (nm)	40	25	24	25
V_{DD} (V)	1.1	1	1	1
EOT (nm)	1.2	1.1	1.2	1.15
T_{FIN} (nm)	25	15	15	10
H_{FIN} (nm)	29	30	30	30
R_{DS} (Ω - µm)	194	220	244	262
I_{on} (µA/µm)	1395	1300	1200	1300
PTM-MG I_{on}	1385	1330	1214	1264
I_{off} (nA/µm)	139	41	100	100
PTM-MG I_{off}	139	43	100	100
Worst Case Error	0.7%	4.87%	1.16%	2.77%

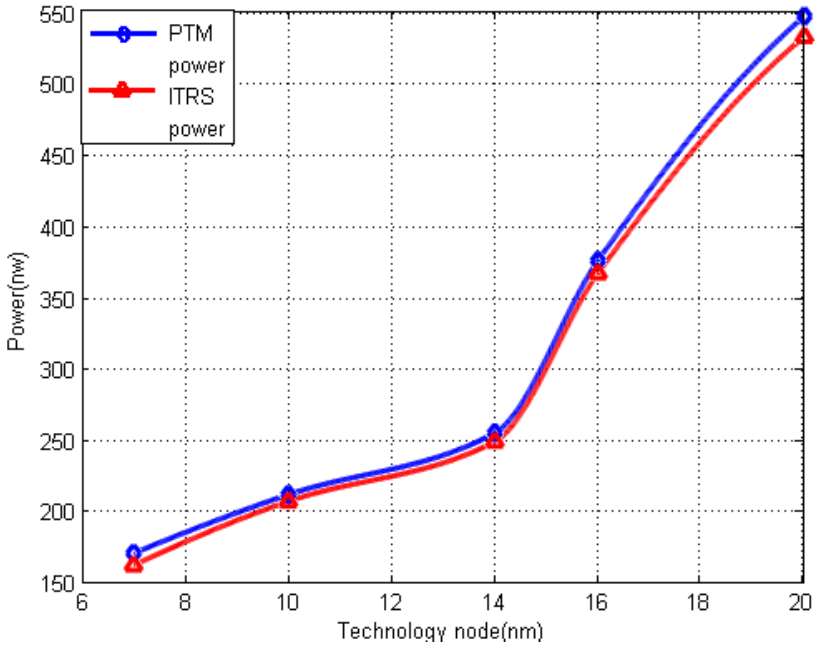


Figure A.1: The difference between ITRS off-current and PTM off-current impact on TG-FF power [122]

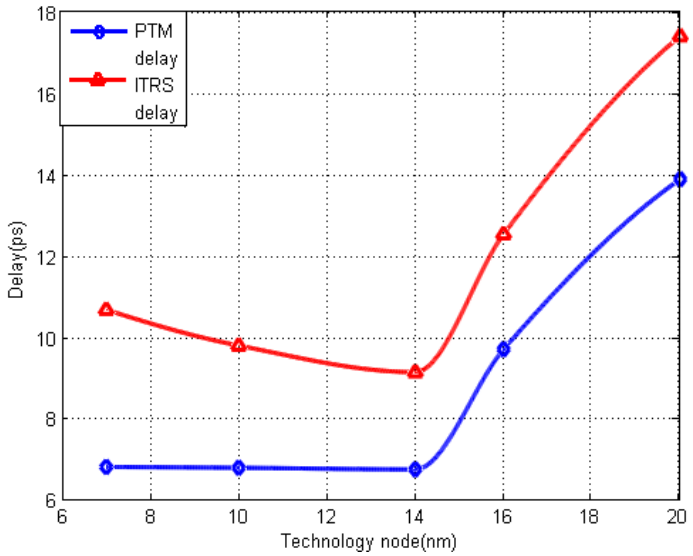


Figure A.2: The difference between ITRS off-current and PTM off-current impact on TG-FF delay [122]

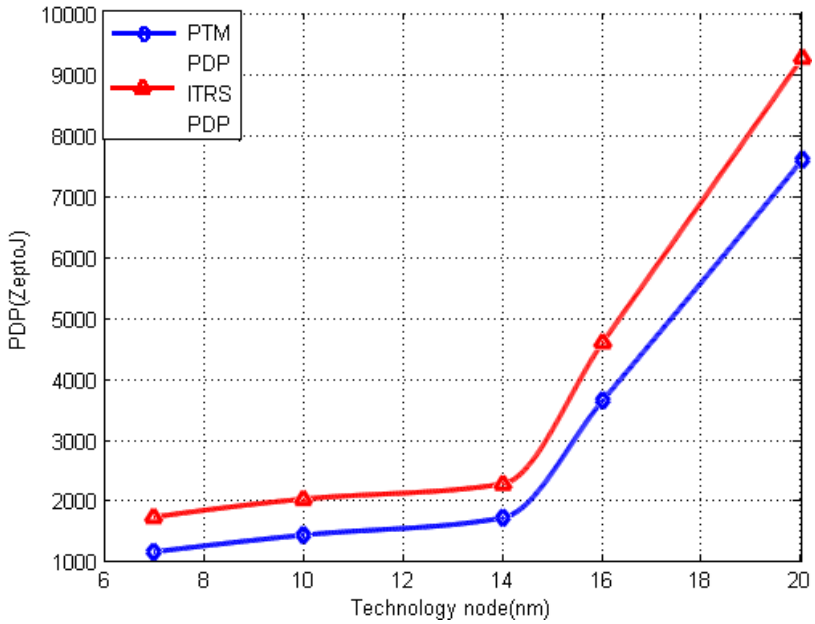


Figure A.3: The difference between ITRS off-current and PTM off-current impact on TG-FF PDP [122]

ملخص الرسالة

كنتيجة للتصغير المستمر لتكنولوجيا ال CMOS تجاه النظام ما دون الميكرون العميق, مصممى الدوائر الرقمية يواجهون تغييرات متزايدة فى صورة تغييرات فى التصنيع أو تغييرات بيئية. يتم تصنيف تلك التغييرات إلى تغييرات بين الشرائح و بعضها أو تغييرات فى نفس الشريحة. العمل البحثى المقدم فى هذه الرسالة يهدف إلى تقييم آداء دوائر المصفوفات المنطقية القابلة للبرمجة بإستخدام تكنولوجيا ال FinFET مع التصغير التكنولوجى بدءا من تكنولوجيا ٢٠ نانومتر إلى ٧ نانومتر فى وجود تغييرات فى جهد الحد و الممثلة للتغييرات بين الشرائح و بعضها بإستخدام نماذج التكنولوجيا التنبؤية التابعة لجامعة Berkley, مع طرح لسلوك بعض مقاييس الآداء المختلفة مثل القدرة المتوسطة, الزمن, و الطاقة. يتم إقتراح بعض التوصيات و النصائح للمصممين بهدف الوصول إلى نسبة كفاءة تصنيع ٩٩,٨٧٪.

نظرا لأن القدرة المسربة ذات فاعلية شديدة فى التكنولوجيا االمتطورة, تم دراسة القدرة المسربة و تغييراتها لتكنولوجيا ١٤ نانومتر فى وجود تغييرات فى جهد الحد و فى درجة الحرارة. توضح نتائج الدراسة طبيعة العلاقة العيارية–اللوغارتمية بين القدرة المسربة و جهد الحد و العلاقة الأسية مع درجة الحرارة.

تم تنفيذ بعض الحلول التى تهدف للتحكم فى القدرة المسربة فى وجود تغييرات فى جهد الحد و التى تشمل تراكم الترانزيستورات, المتجه المنتج لأقل قدرة تسريب, و تحجيم البوابات المنطقية. تلك الحلول أوضحت تحسن ملحوظ فى كل من القدرة المسربة و تغييرات القدرة المسربة, و لكنهم أيضا تسببت فى زيادات طفيفة للزمن و المساحة و التى تم ذكرهم و مقارنتهم.

تم برمجة دوائر المصفوفات المنطقية القابلة للبرمجة المتى تم بناءها لتمثل مؤشر دائرة الجامع ل ٢ بت فى كل من دراسة تصغير التكنولوجيا و القدرة المسربة. تم إستخدام أدوات Cadence Virtuoso و ADE-GXL فى بناء دوائر المصفوفات المنطقية القابلة للبرمجة و محاكاتها على التوالى.

مهندس: محمد محيى الدين محمد على حسن

تاريخ الميلاد: ١٩٩٠١٣١٨

الجنسية: مصرى

تاريخ التسجيل: ٢٠١١١٩١١٧

تاريخ المنح: ..../..../.....

القسم: الإلكترونيات و الإتصالات الكهربية

الدرجة: ماجستير العلوم

المشرفون:

أ.د. حسام على حسن فهمى

د. حسن مصطفى حسن مصطفى

الممتحنون:

أ.د..... (الممتحن الخارجي)

أ.د..... (الممتحن الداخلي)

أ.د. حسام على حسن فهمى (المشرف الرئيسي)

د. حسن مصطفى (مشرف)

عنوان الرسالة:

تصميم دوائر مصفوفات البوابات القابلة للبرمجة (FPGA) بإستخدام تكنولوجيا FinFET بأبعاد أصغر من ٢٢ نانومتر بكفائة عالية ضد عيوب التصنيع

الكلمات الدالة:

التصميم ضد عيوب التصنيع، تغييرات التصنيع، تكنولوجيا FinFET، دوائر مصفوفات البوابات المنطقية، تصغير التكنولوجيا، القدرة المسربة

ملخص الرسالة:

فى هذه الرسالة نقوم بتقييم آداء دوائر مصفوفات البوابات المنطقية بإستخدام تكنولوجيا FinFET مع التصغير التكنولوجى من ٢٠ نانومتر إلى ٧ نانومتر, مع تغييرات جهد الحد الممثلة للتغييرات بين الشرائح الإلكترونية و بعضها, و ملاحظة سلوك بعض مقاييس الآداء المختلفة مثل القدرة المتوسطة, الزمن, و الطاقة. يتم إقتراح بعض التوصيات و النصائح للمصممين بهدف الوصول إلى نسبة كفاءة تصنيع ٩٩،٨٧٪. أيضا يتم دراسة القدرة المسربة لتكنولوجيا ١٤ نانومتر فى حالة وجود تغييرات فى جهد الحد و درجة الحرارة. تم أيضا تنفيذ بعض الحلول للتحكم فى القدرة المسربة مثل تراكم الترانزيستورات, المتجه المنتج لأقل قدرة تسريب, و تحجيم البوابات المنطقية.

تصميم دوائر مصفوفات البوابات القابلة للبرمجة (FPGA) بإستخدام تكنولوجيا FinFET بأبعاد أصغر من ٢٢ نانومتر بكفائة عالية ضد عيوب التصنيع

اعداد
محمد محيى الدين محمد على حسن

رسالة مقدمة إلى كلية الهندسة - جامعة القاهرة
كجزء من متطلبات الحصول على درجة
ماجستير العلوم
في
الإلكترونيات و الإتصالات الكهربية

يعتمد من لجنة الممتحنين:

الاستاذ الدكتور:	الممتحن الخارجي
الاستاذ الدكتور:	الممتحن الداخلي
الاستاذ الدكتور: حسام على حسن فهمى	المشرف الرئيسى
الدكتور: حسن مصطفى	مشرف

كليــة الهندســة - جامعــة القاهــرة
الجيـزة - جمهوريـة مصـر العربيــة
٢٠١٦

تصميم دوائر مصفوفات البوابات القابلة للبرمجة (FPGA) بإستخدام تكنولوجيا FinFET بأبعاد أصغر من ٢٢ نانومتر بكفائة عالية ضد عيوب التصنيع

اعداد

محمد محيى الدين محمد على حسن

رسالة مقدمة إلى كلية الهندسة - جامعة القاهرة

كجزء من متطلبات الحصول على درجة

ماجستير العلوم

في

الإلكترونيات و الإتصالات الكهربية

تحت اشراف

أ.د. حسام على حسن فهمى

أستاذ

قسم الإلكترونيات و الإتصالات الكهربية

كلية الهندسة , جامعة القاهرة

د. حسن مصطفى حسن مصطفى

مدرس

قسم الإلكترونيات و الإتصالات الكهربية

كلية الهندسة , جامعة القاهرة

كليــة الهندســة - جامعــة القاهــرة

الجيـزة - جمهوريـة مصـر العربيــة

٢٠١٦

تصميم دوائر مصفوفات البوابات القابلة للبرمجة (FPGA) بإستخدام تكنولوجيا FinFET بأبعاد أصغر من ٢٢ نانومتر بكفائة عالية ضد عيوب التصنيع

اعداد

محمد محيى الدين محمد على حسن

رسالة مقدمة إلى كلية الهندسة - جامعة القاهرة
كجزء من متطلبات الحصول على درجة
ماجستير العلوم
في
الإلكترونيات و الإتصالات الكهربية

كليــة الهندســة - جامعــة القاهــرة
الجيـزة - جمهوريـة مصـر العربيــة
٢٠١٦