PES UNIVERSITY
(Established under Karnataka Act No. 16 of 2013)
Electronic City Campus, Bengaluru – 560 100, Karnataka, India

Report on

"Design Methodologies for Sub-Threshold Ultra Low Power Standard Cells"

Submitted by

B K Karthik (PES2201800185)
Pruthvi Raj S (PES2201800426)
Challa Sai Sneha (PES2201800547)
Varun Tewari (PES2201800189)

January - December 2021

Under the guidance of

Dr. Madhura Purnaprajna
Professor, Department of ECE
PES University EC Campus

Prof. Vinay Reddy
Assistant Professor, Department of ECE
PES University EC Campus

FACULTY OF ENGINEERING
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BACHELOR OF TECHNOLOGY

CERTIFICATE

This is to certify that the report titled

"Design Methodologies for Sub-Threshold Ultra Low Power Standard Cells"

is a bonafide work carried out by

B K Karthik (PES2201800185)
Pruthvi Raj S (PES2201800426)
Challa Sai Sneha (PES2201800547)
Varun Tewari (PES2201800189)

In partial fulfillment for the completion of 7th-semester course work in the Program of Study B.Tech in Electronics and Communication Engineering, under rules and regulations of PES University, Bengaluru during the period January – May and June - December. 2021. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report. The report has been approved as it satisfies the 7th-semester academic requirements in respect of Capstone project work.

Signature with date & seal	*Signature with date & seal*	*Signature with date & seal*
Dr. Madhura Purnaprajna	***Prof. Vinay Reddy***	***Dr. Ajey S N R***
Project Guide	*Project Guide*	*Chairperson*

Signature with date & seal
Dr. B K Keshavan
Dean, Faculty of Engg. & Technology

DECLARATION

We, **B K Karthik, Pruthvi Raj S, Challa Sai Sneha, Varun Tewari**, hereby declare that the project report titled, '**Design Methodologies for Sub-Threshold Ultra Low Power Standard Cells'**, is our original work under the guidance of **Dr. Madhura Purnaprajna**, Professor and **Prof. Vinay Reddy**, Assistant Professor, ECE Department and is being submitted in partial fulfillment of the requirements for completion of 7th Semester course work in the Program of Study, B.Tech in Electronics and Communication Engineering.

Date and Place:

Name and Signature of the Candidates:

1.

2.

3.

4.

Abstract

In the present era of high-density and high-speed nanoelectronics, power consumption has been one of the most concerning factors. Hence there is a rapidly growing demand for ultra-low power devices and advanced energy-saving methods for digital integrated circuits. The need for low-power circuits has up to now been limited to a small number of products, but this situation has changed drastically in the last few years, mainly because of the growing need for portability in computer and telecommunication products. In this work, we discuss the motivation, trends, and challenges of operating digital circuits in the sub-threshold regime using UMC's commercial 0.18 μm and 28nm high-performance compact high-K bulk CMOS processes. We synthesize a 32-bit 6-pipeline stage RV32IMFC Chromite RISC-V core. The obtained results are compared with Quentin [4].

Acknowledgments

First, we would like to thank Dr. J Suryaprasad, Hon'ble Vice-Chancellor, PES University in Bengaluru for his continuous and unconditional support. We appreciate the opportunity to utilize industry-standard EDA tools and IPs.

We want to express our gratitude towards our Dr. Ajey S N R, Hon'ble Chairperson, and the entire Department of Electronics and Communication Engineering at the Electronic City Campus.

We would like to express our many thanks to Dr. Madhura Purnaprajna, Professor and Head of the Centre for Heterogeneous and Intelligent Processing Systems at the Electronic City Campus for her continuous support and guidance throughout this entire project.

Finally, we would like to thank Prof. Vinay Reddy for being an extremely knowledgeable and helpful guide throughout the project timeline, who provided both factual assistance and helped boost our morale at critical times.

Contents

Chapter 1

Introduction

The demand for low power consumption is increasing in battery portable battery-operated devices used in modern applications. The power dissipation of an IC can be minimized by reducing the supply voltage below or near the threshold voltage of its transistors, which results in new challenges for the designer [1]. Scaling the voltage brings not only quadratic power savings but also super-linearly reduced leakage currents. This work motivates the use of Near- / Sub- Threshold standard cells for portable applications with a constrained energy budget, exploring methods to further reduce energy consumption [2] [3]. We analyze the performance of our standard cells by synthesizing a 32-bit 6-pipeline stages Chromite RV32IMFC core and compare it with Quentin [4].

1.1 Motivation and Goals

To fulfill the energy requirements of modern mobile and embedded devices with shrinking transistor sizes, energy efficiency has become a critical aspect of designing digital circuits. Voltage scaling techniques to near or sub-threshold voltages have been explored in the recent past [5]. The aim is to take maximum advantage of the quadratic relationship between the dissipated power and the supplied voltage. A thorough investigation of possible design alternatives allows the designer to make an informed decision when evaluating the viability of design candidates. The quality of a digital logic gate is determined by its characteristics, mainly speed, energy efficiency, area, and robustness against noise.

The ultimate goal is the design of an ultra-low power standard cell library in the UMC 28nm high-K HPC bulk CMOS process. To minimize the power dissipation of the gates, a very low supply voltage is applied to keep the transistors in the sub-threshold region. The behavior of the gates in this region is less predictable, and hence extensive and exhaustive testing is required.

1.2 Structure of this Work

In Chapter 1, a short introduction to the topic, the motivation, goals, trends, and challenges are given along with a brief description of recent publications.

Chapter 2 starts with the basic structure of a bulk MOS Field Effect Transistor and expands into CMOS logic gate design. To conclude the chapter, background information on the need for multi-objective optimization is described.

In Chapter 3, we discuss the standard cell library creation methodology in detail along with the toolflow.

Chapter 4 describes the RISC-V instruction set architecture and the structure of the Chromite RISC-V.

Finally, we present our results in Chapter 5 where we demonstrate a detailed performance analysis of the Chromite RISC-V core with Quentin.

1.3 Related Work

The trends and enablement in the recent past have been explored well in [5]. To reduce the dynamic power consumption of integrated circuits and to keep electric fields inside small devices low, supply voltages can be scaled. For a given aspect ratio, the minimum energy point decreases as the transistor activity factor increases. For high switching activity circuits, it is more feasible to operate at the minimum energy point compared to low switching activity circuits [6].

The minimum energy point shifts to a higher level with the stacking of transistors, Therefore the stacking of transistors should be minimal while operating in the sub-threshold region [6]. For many applications in the field of the Internet-of-Things [7], energy is the main limiting factor, while performance plays only a minor role. Hence, there is a need for integrated systems that rigorously trade-off performance against energy consumption.

The need for low power design is doubling the operating frequency every three years with the associated increase in the power dissipation per unit area and the increase in the chip temperature. Dynamic power is greatly reduced, primarily due to the quadratic dependence on supply voltage, while static subthreshold leakage is also much lower, mainly because of the Drain-Induced Barrier Lowering (DIBL) effect [8]. The Challenge of circuit operation in the subthreshold region is the relatively weak current flow resulting in longer delays and lower frequencies, by lowering the supply voltage while a system is in “standby” or low-performance mode avoids delay.

An important advantage of CMOS logic is robustness against voltage and transistor scaling, and its reliable operation at low voltages and minimal transistor size. CMOS logic can also help us achieve robustness and better functionality in the presence of large noise contents due to the presence of a static path.

The processor cores [9] exhibit a minimum energy consumption of 9.94 pJ per clock cycle at 325 mV and 133 kHz. At this voltage, the 2 kb 9T SRAM macro displayed access energies down to 529 fJ. From 38 measured samples at 320 mV, all were functional and the slowest operated at 330 kHz. A level shifter circuit exists that is capable of converting subthreshold to above-threshold signal levels based on Wilson current mirror [14] [15]. The circuit does not have a static current path between the supply rails and therefore offers a reduced static power dissipation.

Level shifter behaves similarly to traditional CMOS logic gates in subthreshold. Subthreshold devices operate at voltages well below the nominal VDD for most processes, level converters are necessary to interface subthreshold circuits to core voltage levels.

Despite the strong motivation, voltage scaling to such low levels is not common in the industry with a few exceptions [10], [11]. The earliest known processor [12] where supply voltage scaling was performed, used a modified pass-transistor logic to perform FFT operations on custom SRAMs with a latch-based register file and multiplex reads to ensure a robust operation at voltages up to 10x lesser than the conventional operating points. A 2-stage microarchitecture [13] demonstrates the usage of shallow pipelines with a high FO4 delay per stage to reduce variability. This processor consumes 2.60 pJ/inst in a commercial 130-nm CMOS process technology. The processor in [16] demonstrates the use of body biasing and gate sizing techniques to mitigate the impact of process variations.

Chapter 2

Background

Transistors are the building blocks of complex logic gates and digital ICs. As they are part of the most fundamental level of the design process, they have a very high impact on the IC's overall performance. The transistors used in this work are bulk-type Metal Oxide Semiconductor Field Effect Transistors (MOSFETs). In this chapter, we discuss the basics of a MOS transistor along with the regions of operation and the optimization methodologies.

2.1 Structure of a MOSFET

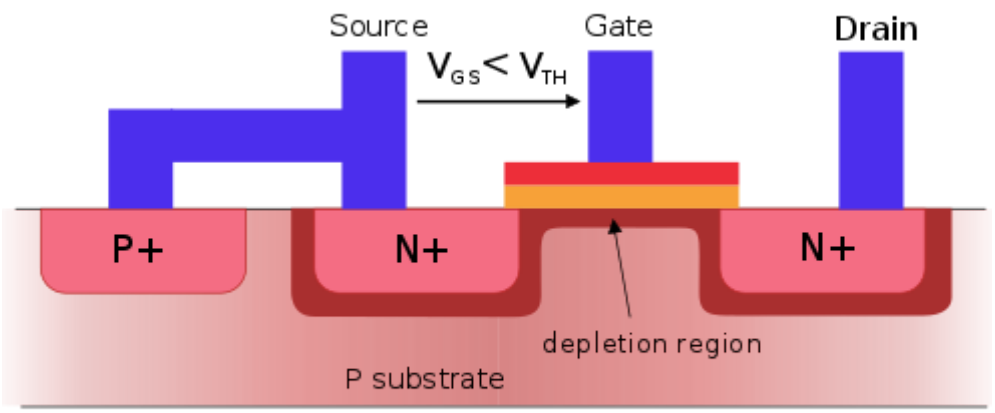

Fig 1: Cross-section of an n-type MOSFET [8]

The Metal Oxide Semiconductor Field Effect Transistor is a type of insulated-gate field-effect transistor that is fabricated by the controlled oxidation of a semiconductor, typically silicon. A MOSFET can be described as a voltage-controlled current source in electrical terms, here it operates as a switch to conduct electricity in charging and discharging circuits.

The voltage of the covered gate determines the electrical conductivity of the device; this ability to change conductivity with the amount of applied voltage is used for switching electronic signals [8]. The base material for MOSFETs is the semiconductor silicon (Si) which is a member of Group IV in the periodic table of elements. With its four valence electrons, Si can form strong covalent Si-Si bonds with four surrounding Si atoms, leading to the formation of a 3-dimensional monocrystal structure. Si, as a substrate or wafer, has a lot of advantages due to the small mass, good thermal conductivity, the low cost, and is widely used in microchip production because of these features. The conductivity of a pure Si crystal is very low since all of the four valence electrons are required to form the bonds. To improve the conductivity of the Si crystal impurities, called dopants, are introduced into the crystal structure which replaces Si atoms inside the crystal grid.

After the formation of four covalent bonds with the surrounding Si atoms, dopants from Group V of the periodic table of elements (e.g. Phosphorus, Arsenic) have one excess electron left that can act as a free charge carrier inside the grid. Since this results in a net negative charge, it is called an n-type semiconductor. If the silicon is doped with elements from Group III (e.g. Boron), i.e. with only three valence electrons, the dopant borrows an electron from one of its neighbors creating a positively charged hole. This hole can propagate through the crystal, similar to free electrons, leading also to conductivity. This is called a p-type semiconductor.

The operating principle of a transistor, e.g. MOSFET, is based on the diode effect of a semiconductor. Depending on the types of dopants used, different kinds of MOSFETs with varying functionality can be built.

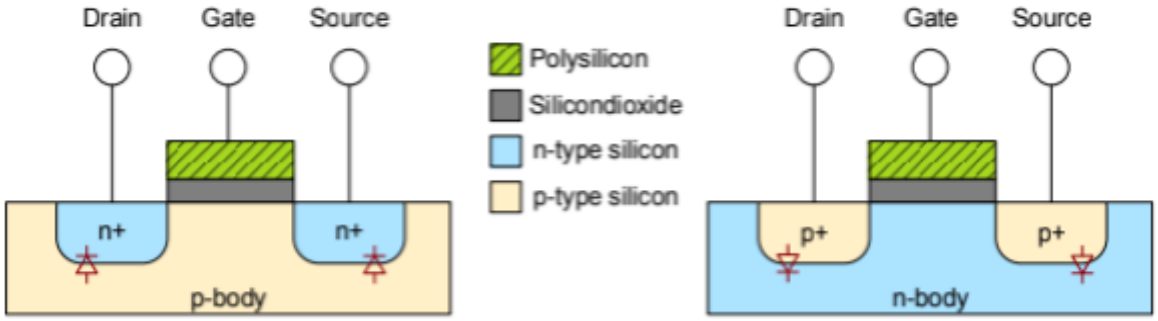

Fig 2: Cross-section of nFET (l.) and pFET (r.)

For the nFET, the body is p-type silicon with two areas of n-type silicon called source and drain. The source is generally connected to the ground (GND). The gate stack, which consists of an insulating layer of SiO_2 and the conducting gate terminal made of polysilicon, is placed between the source and drain. Besides source, drain, and gate the fourth terminal connects the p-type body to a low potential, usually GND. Hence, the p-n junctions from the body to source and drain remain reverse-biased.

The transistor functions as an on/off switch whose state is controlled by the gate. If no voltage is applied to the gate (a logical 0 in an IC) no electric field is created between the body and the gate. Hence, the p-n junctions remain reverse-biased so that no current flows between source and drain and the transistor is off. Increasing the gate voltage Vgs creates an electric field from the gate to the

body that attracts electrons to the gate. When Vgs exceeds the threshold voltage Vt, enough electrons are gathered beneath the gate to form a conducting channel. Current can flow from drain to source, turning the transistor on.

For the pMOS transistor, the source and drain are made from p-type silicon while the body is n-type silicon. The pMOS body is held at a high potential, usually, the supply voltage Vdd, to avoid forward-biasing the p-n junctions between source/drain and body as well as the p-n junction between the bodies of nFET and pFET in proximity. The source is usually connected to Vdd.

As long as the gate is also at Vdd, the diodes between body and source and drain stop any current flowing from the source and drain. The transistor is in the off state. When the gate voltage is lowered an electric field between the body (at Vdd) and gate forms. If Vgs is lowered enough, by the threshold Vt, the field becomes so strong that it attracts enough holes to the area beneath the gate and a conducting channel forms. Electric charge can flow from source to drain and the transistor is turned on.

2.2 Regions of Operation

The threshold voltage Vt is not a constant but instead related to dynamic factors, like the source-body voltage Vsb of the transistor.[1] The latter dependency can be explained by the increasing amount of charge required to invert the channel if a body potential is applied. This phenomenon is also termed the body effect.

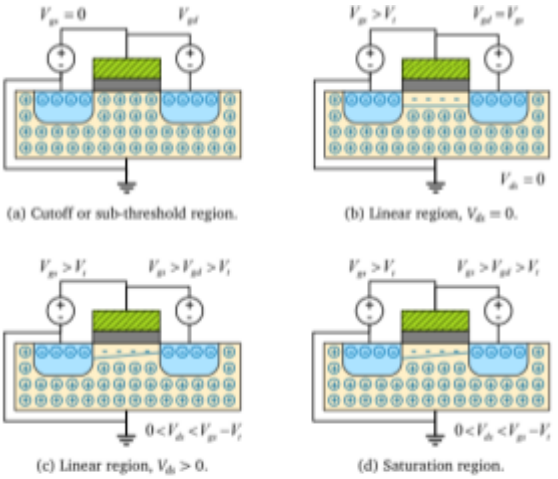

Fig 3: Regions of Operation

The drain-source voltage Vds also affects Vt in the form of the drain-induced barrier lowering

[1] The variation of the threshold voltage may also depend upon physical factors such as temperature and process variations. These have been discussed in the coming sections of this chapter.

(DIBL) due to the electric field that is created between source and drain. Turning a transistor on or off requires charging or discharging the gate where the amount of charge contributes to the power consumption of the transistor and also influences the time it takes to charge or discharge the gate. In consequence, it is important to consider the capacitance of a transistor in more detail.

The terminals of a MOSFET are capable of storing electric charge and thus act as a capacitor when a potential difference between two terminals occurs. The most obvious is the gate capacitor Cg that can be modeled as a simple plate capacitor with plate size W · L where W is the width and L is the length of the gate. While the gate capacitor is required for the function of the transistor the other capacitors are not. Yet, their capacitance influences the performance of the transistor. As a consequence, they are called parasitic capacitors. Another aspect of energy consumption is the unwanted current in a transistor, called leakage. In a transistor below the threshold voltage, a small amount of charge flows between the source and drain. It is caused by the thermal emission of carriers. In this state, which is also called weak inversion, the drain current is described by

$$I_{sub} = \mu_0 C_{ox} \frac{W}{L} \left(\frac{C_{dm}}{C_{ox}} - 1 \right) (v_T)^2 e^{(V_g - V_t)/m v_T} (1 - e^{-V_{ds}/v_T})$$

Hence, a linear dependency between channel length L and width W is observed. Another source of the leakage is the gate leakage Igate caused by electrons tunneling through the gate oxide layer. This is related to the type and thickness of the oxide layer, but also linearly dependent on the gate width W. The final leakage type is called junction-leakage Ijunc. It happens at the p-n junctions between drain/source and the body as even reverse-biased diodes still conduct small amounts of current. Most of the leakage happens where the dopant is strongest, i.e. at the junction towards the gate.

2.3 CMOS Logic Gates

For logic gate design, a transistor can be regarded as a switch. If the switch is on a conducting connection between its source and drain is established. The on and off states of the switch are controlled by its gate input. pFETs are low active, being on when the gate has a low input signal (logical 0). nFETs transistors behave the

opposite way, being on when the gate is high (logical 1). These simple switches allow the creation of more complex logic gates.

CMOS logic gates are made from a combination of pMOS and nMOS transistors. The transistors are placed complementary in the circuit. This ensures that either the nFET or the pFET is turned off while its counterpart is active. The pFETs form the pull-up network while the nFETs form the pull-down network. Further logic gates can be created by combining pull-up and pull-down networks of parallel or series of transistors.

To transition from the logic design to a physical layout, a set of masks has to be created using a layout program. These masks specify the size and position of the different transistor parts, e.g. n-well, poly-silicon, n- and p-type diffusion zones, metal, and contact layers. The designer has to follow design rules specified by the manufacturer that depend on the production process. They constrain for example the sizes of and distances between specific parts. Generally, smaller scales require more complex design rules making it difficult to adapt a layout to a new manufacturing process.

2.4 Multi-Objective Optimisation

As seen earlier, the process of designing a CMOS standard cell requires balancing many different design parameters like transistor dimensions, voltages, and topologies. The set of possible combinations of these design parameters is called the design space (DS). The process of examining the DS to find favorable sets of parameters (candidates) is called design space exploration (DSE).

The dimension n of the design space is the number of parameters that can be adjusted by the designer and the DS grows exponentially with n. Usually, it is the designer's choice which parameters should be investigated for the DSE and which should be set to a promising constant value to reduce the size of the DS. To find the best design candidates for a given application, it is necessary to compare as many configurations as feasible.

2.5 Formulation of the Optimisation Problem

To compare the quality of the candidates, a metric is used that considers the

timings, energy, robustness, and area of the standard cell. As each of these attributes is sought to be optimized, it is called a MOO problem that can be formulated as

$$\min_{x \in S} f(x)$$

where f: S ⊂ Rn → Rm is a function that takes the n-dimensional vector of design parameters from DS S and returns a three-dimensional vector of design objectives. Usually, there is no best single candidate but the best set of candidates called the Pareto set P. The corresponding set function outcomes to the Pareto set are called the Pareto front. The members of the Pareto set are all non-dominated candidates with domination defined as:

$$\text{For } a,b \in Rm: a \prec b \text{ (a dominates b)} \Leftrightarrow ai \leq bi, \forall i = 1..m \text{ and } a \neq b$$

As each objective is supposed to be minimized in this equation, for some objectives, such as NM, higher values are preferable. This definition problem can be circumvented by optimizing the negative NM instead. With the formal definition of domination, the Pareto set can be described as:

$$P = \{p \in S | \nexists x \in S : f(x) \prec f(p)\}$$

A DSE for CMOS standard cells is a major challenge considering the size of the search space and requiring simulation and analysis of every configuration. To increase the efficiency of the search for the Pareto set a good strategy is required. This can be achieved by the use of an optimization algorithm that provides a heuristic to find the Pareto set or a close approximation of it. The objective of an OA is to achieve good quality results while being resource-efficient (run time and memory usage). The quality of the results is determined both by the distance of the found Pareto front to the optimal Pareto front and the diversity of the solutions (distribution in search and object space).

A very popular approach is the use of EAs that mimic processes of biological evolution. Here, a population Pt of design candidates is evaluated by a fitness metric, the best candidates are chosen for reproduction in the selection phase and a new generation Pt+1 is created by crossover and mutation.

While many implementations of Pareto calculations already exist, the specific requirements for this project justified a new implementation. Other Pareto calculations usually restrict the search to two Pareto objectives (POs), while logic

gate design often requires more than two optimization targets. Therefore, a generic approach to Pareto front calculation is implemented, where the number of Pareto objectives can be defined without restrictions. Further, a common PO in logic gate design is noise margin, where a high value is preferred instead of a minimum value. To accommodate this fact, the implementation allows both minimizing and maximizing each PO by defining a flag that indicates a low or high desired value. This allows a more intuitive way of presenting the Pareto objectives (instead of, for example, minimizing the negative of the PO).

2.6 Optimisation Algorithms

As one of the first genetic algorithms to find Pareto-optimal solutions for a multi-objective optimization problem, the Non-dominated Sorting Genetic Algorithm (NSGA) was proposed. The NSGA-II improves on the main disadvantages of the NSGA, which are the high computational complexity, the lack of elitism, and the requirement of specifying a sharing parameter.

The primary selection criterion for the NSGA-II, based on the non-domination of the population, is called fast non-dominated sort. For each candidate p, the number of candidates, by which it is dominated, is calculated as the non-domination count np. Additionally, a set of candidates that p dominates is kept as Sp. Each candidate is then sorted into non-domination fronts Fi, based on its non-domination count np = i. To minimize the complexity of this sorting process, firstly only the candidates with np = 0 are selected and put into front F0. Then the algorithm iterates through the candidates from the domination list Sp of each member of F0 and reduces their non-dominated count by 1. When a candidate's non-dominated count reaches 0, it is sorted into the subsequent front. This process is repeated for all fronts Fi. Calculating non-domination count and domination sets for N candidates with M objectives requires O(MN2) comparisons.

Afterward, each candidate's nomination list is visited, containing at most N - 1 members, which results in a complexity of O(N2). Thus, the overall complexity of the fast non-dominated sort is O(MN2). Due to higher precision in the fitness assignment, an archive that prevents outer solutions to be lost, and a density estimation as a secondary selection criterion, its solutions are more diverse than its predecessor's.

2.7 Machine Learning Techniques

Simulation-based Genetic Algorithms function by stochastically sampling an initial population and mutating the best children to produce offspring to then simulate and sample from again. Traditionally, these methods are sample inefficient, and not guaranteed to converge because of stochasticity. In addition, they require re-starting the algorithm from scratch if any change is made to the goal. Learning-based tools use machine learning methods to solve multi-objective problems.

Reinforcement Learning (RL) is a machine learning technique known to solve complex tasks in many systems. Specifically, it consists of an agent that iterates in an environment using a trial and error process that mimics learning in humans. It is a simulation-in-loop method, having the ability to verify outputs. At each environment step, the RL agent, which contains a neural network, observes the state of the environment and takes an action based on what it knows. The environment then returns a new state that is used to calculate the reward for taking that particular action. The agent iterates through a trajectory of multiple environment steps, accumulating the rewards at each step until the goal is met or a predetermined maximum number of steps is reached. After running multiple trajectories the neural network is updated to maximize the expected accumulated reward via policy gradient.

Chapter 3

Standard Cell Library

A standard cell is a group of transistors and interconnect structures that provide a boolean logic function (e.g., AND, OR, XOR, XNOR, inverters) or a storage function (flip flop or latch). The simplest cells are direct representations of the elemental NAND, NOR, and XOR boolean function, although cells of much greater complexity are commonly used (such as a 2-bit full-adder, or muxed D-input flip flop.) The cell's boolean logic function is called its logical view: functional behavior is captured in the form of a truth table or Boolean algebra equation (for combinational logic), or a state transition table (for sequential logic).

3.1 Standard Cell Design

Simply put, a standard cell is a group of transistors and interconnect structures that provide a boolean logic or a storage function as an off-the-shelf component for semi-custom design flows. Conventional standard cells may be logic gates such as the NAND, NOR, and so on, along with storage cells such as flip-flops and latches. It is up to the designer to provide complex standard cells such as full adders and mixed flip-flops which would take more than one conventional standard cell to implement leading to slower operation and a larger area overhead.

A standard cell used in digital logic circuits could be optimized for multiple combinations of parameters such as area, delay, and power. A standard cell may also be characterized for corners such as TT, FS, or SF. Depending on the type of circuit, a particular corner may be selected to improve the performance of the circuit.

3.2 Operating Points

A standard cell is characterized in a single process, voltage, and temperature setting termed as a PVT corner. Our standard cells designed using UMC's 28nm HPC Commercial process are characterized to operate at a power supply of 180 mV at 25 degrees centigrade in the TT corners on silicon. For mixed-mode operation, process variations and temperature deflections are modeled as random variables to perform a statistical Monte-Carlo analysis to understand the extremes of functionality as analog circuits.

3.3 Library Characterization

Several steps are required to design a custom standard cell library in a specific technology (in this case UMC 28nm Bulk CMOS). Each standard cell is designed and optimized considering the constraints of the target technology. A schematic of the standard cell is initially drawn in a CAD schematic editor (Cadence Virtuoso), where its logic function, structure, and design parameters are defined. The schematic is then extended to include a test bench to begin simulation (for instance Cadence ADE) of the standard cell. In simulations, the characteristics of the

gate are calculated with the given design parameters and subsequently evaluated. A Monte Carlo (MC) simulation is performed that models the variance of the design parameters in the final fabrication of the cell. Based on the evaluation of the simulations, the schematic can be changed to better fit the constraints. When the design passes the simulation and MC analysis, the cell layout is created in a layout tool (Cadence Virtuoso Layout L) and extracted (Mentor Calibre, Cadence Assura, Cadence PVS). The cell layout is then analyzed with an MC simulation where again deviations of the layout caused by the production process are examined. Once an error-free layout has been drawn for a standard cell, we extract the post-layout netlist along with parasitic elements and proceed to perform a back annotation.

Finally, a library characterization tool creates the library files from the layouts and simulation files. The characterization of the cells provides all necessary information to the place and route (P&R) tool on the properties of the cells.

Chapter 4

Performance Analysis

RISC-V is an open standard instruction set architecture (ISA) based on established reduced instruction set computer (RISC) principles. In this chapter, we discuss the performance of a 32-bit Arithmetic Logic Unit and a 6-stage pipelined RV32IMFC Chromite RISC-V Core.

4.1 32 Bit Arithmetic and Logic Unit

Our Arithmetic and Logic Unit is designed using the Verilog hardware description language. Upon Synthesis with Faraday's IPs and our subthreshold standard cells, we can draw the following results about the power and timing information. We use Cadence Genus with clock and power gating to obtain these results. The synthesis was area-unaware and the cells were optimized for minimum power dissipation.

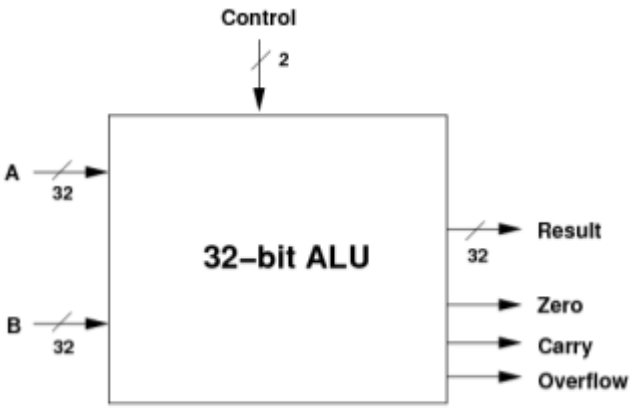

Fig 4: 32 Bit Arithmetic and Logic Unit

Sub-Threshold Cells		Super-Threshold Cells	
Power		Power	
Parameter	Value (nW)	Parameter	Value (mW)
Leakage	1.27 uW	Leakage	1.329 mW
Internal	3.92 uW	Internal	5.722 mW
Switching	7.60 uW	Switching	17.262 mW
Total	**12.79 uW**	Total	**22.985 mW**
Timing		Timing	
Parameter	Value	Parameter	Value
Critical Path	**22.26 ns**	Critical Path	**84.27 ps**

4.2 RV32IMFC Chromite RISC-V Core

RISC-V is a free and open ISA enabling a new era of processor innovation through open standard collaboration. In this section, we discuss the structure of a Chromite RISC-V core. The Chromite RISC-V core incorporates a 6-stage pipelining. The core pipeline consists of the following stages:

1. Program Counter Generation Stage (PC Gen)
2. Instruction Fetch Stage (IFS)
3. Decode Stage
4. Execution Stage
5. Memory Stage
6. Write Back Stage

The figure describes the different stages in the core's pipeline. Upon generation of the hardware description, we perform a constrained synthesis using Cadence Genus Synthesis Solutions. The translation is based on the supervisor spec defined in the RISC-V Privileged Spec. The Chromite Core Generator provided by InCore Semiconductors [17] was used with the given configuration.

```
ISA: RV32IMFC
iepoch_size: 2
depoch_size: 1
dtvec_base: 256
s_extension:
 mode: sv32
 itlb_size: 4
 dtlb_size: 4
 asid_width: 9
pmp:
 enable: true
 entries: 4
 granularity: 8
m_extension:
 mul_stages: 1
 div_stages: 32
branch_predictor:
 instantiate: True
 predictor: gshare
 on_reset: enable
 btb_depth: 32
 bht_depth: 512
 history_len: 8
 history_bits: 5
 ras_depth: 8
icache_configuration:
 instantiate: true
 on_reset: enable
 sets: 64
 word_size: 4
 block_size: 16
 ways: 4
 fb_size: 4
 replacement: RR
 ecc_enable: false
 one_hot_select: false
dcache_configuration:
 instantiate: true
 on_reset: enable
 sets: 64
 word_size: 4
 block_size: 16
 ways: 4
 fb_size: 8
 sb_size: 2
 replacement: RR
 ecc_enable: false
 one_hot_select: false
 rwports: 1
reset_pc: 4096
physical_addr_size: 32
bus_protocol: AXI4
fpu_trap: false
debugger_support: false
no_of_triggers: 0
csr_configuration:
 structure: daisy
 counters_in_grp4: 7
 counters_in_grp5: 7
 counters_in_grp6: 7
 counters_in_grp7: 8
bsc_compile_options:
 test_memory_size: 33554432
 assertions: true
 trace_dump: true
 compile_target: 'sim'
 suppress_warnings: ["all"]
 verilog_dir: build/hw/verilog
 build_dir: build/hw/intermediate
 top_module: mkTbSoc
 top_file: TbSoc.bsv
 top_dir: test_soc
 open_ocd: false
verilator_configuration:
 coverage: none
 trace: false
 threads: 1
 verbosity: true
 sim_speed: fast
 out_dir: bin
```

The Chromite test SoC has the following hierarchy of modules:

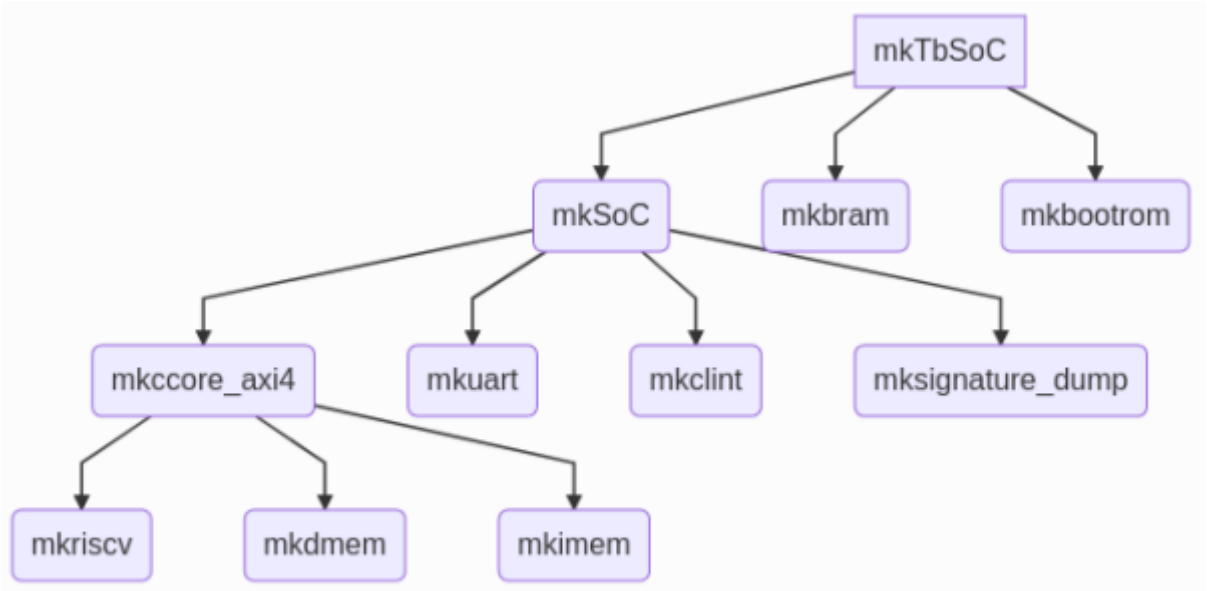

Fig 5: Structure of the Test SoC

The module *mkriscv* contains the 5-stages of the core pipeline including the execution and only the interface to the memory subsystem, whereas *mkcore_axi4* contains the *mkdmem* or the data memory subsystem that includes the data-cache and data-tlbs and the *mkimem* or the instruction memory subsystem that includes the instruction-cache and the instruction-tlbs.

The mkccore_axi4 contains the above modules and the integrations across them. Also provides 3 AXI-4 interfaces to be connected to the crossbar fabric. To evaluate our standard cells, we only synthesize the *mkriscv* subsystem as we don't have memory cells in the library. This would lead to the memories being represented as an array of registers, and would provide severely inaccurate results.

The performance of our core upon synthesizing *mkriscv* has been tabulated below.

Parameter	Chromite (Custom Sub-Vt)	Chromite (Custom Sub-Vt)	Chromite (Faraday Super-Vt)	Chromite (Skywater Super-Vt)	Quentin [4]
ISA	RV32IMFC	RV32IMFC	RV32IMFC	RV32IMFC	RV32IMFC
Process technology	UMC 28nm Bulk	UMC 28nm Bulk	Faraday IP 55nm	Skywater 130nm Bulk	GF 22nm FDX
Voltage range (VDD)	0.18 V	0.3 V	1.32 V	1.8 V	0.5 V - 0.8 V
Voltage range (Body Bias)	-	-	-	-	0.0 V - 1.4 V
Frequency range (without BB)	1.21 MHz	3.33 MHz	200 MHz	50 MHz	32 KHz - 670 MHz
Frequency range (with BB)	-	-			32 KHz - 938 MHz
Power range	64.65 uW	296.89 uW	52.58 mW	55.03 mW	300 uW - 66.2 mW
Power density (per MHz)	73.58 uW/MHz @ 0.18 V, 1.21 MHz	89.04 uW/MHz @ 0.3 V, 3.33 MHz	262.9 uW/MHz @ 1.32 V, 200 MHz	1.60 mW/MHz @ 1.8 V, 50 MHz	8.7 uW / MHz @ 0.52 V, 187 MHz
Cell Count	142038	109045	49156	50667	-
Max CoreMarks per MHz	2.9	2.9	2.9	2.9	-
CoreMarks (score)	3.5	9.66	580	145	-

Chapter 5

Conclusion and Future Scope

In this work, we discussed the motivation, trends, and challenges in subthreshold circuit design. An ultra-low-power standard cell library in the UMC 180nm bulk and 28nm high-K HPC bulk CMOS process was designed. All cells have transistors operating in the subthreshold regime. The standard cell library has 31 cells at the time of submission and can be extended as per requirements. All cells were fully characterized at the transistor level for timing, area, and delay. A liberty file and layer exchange format were generated for all the cells.

To analyze the performance, A 32-bit 6-stage pipelined RV32IMFC Chromite RISC-V core was synthesized and the parameters were compared to Quentin [4]. We found the core to be approximately 2.4x slower than Quentin but the core was also approximately 435x more power efficient.[2] Upon considering the caches, the core turns out to be upto 17x power efficient per MHz.

The future scope for this work is limitless. For instance, user-preference-based multi-objective optimization of transistor sizing may be a good contribution. This work motivates the use of sub-threshold circuits where we have applications with constrained energy budgets. Super-threshold circuits may be used only where there is a need for high performance / high bandwidth in the circuitry. We plan to develop and synthesize multi-core RISC-V processors for domain-specific applications and ultra-low power FPGA Fabrics for domain-specific acceleration.

Upon making some changes to the microarchitecture of the Chromite core, we will be able to evaluate the performance of a 32/64 bit in order 6 stage pipelined RISC-V microprocessor for Ultra Low Power Applications.

[2] These values were obtained upon synthesizing mkriscv as the top module which excludes caches. As the Quentin SoC contains memories, the comparison would only be accurate when memories are included.

Bibliography

[1] S. Fisher, A. Teman, D. Vaysman, A. Gertsman, O. Yadid-Pecht and A. Fish, "Digital subthreshold logic design - motivation and challenges," 2008 IEEE 25th Convention of Electrical and Electronics Engineers in Israel, 2008, pp. 702-706, doi: 10.1109/EEEI.2008.4736624.

[2] Liu, Bo & Pineda de Gyvez, José & Ashouei, Maryam. (2013). Sub-Threshold Standard Cell Sizing Methodology and Library Comparison. Journal of Low Power Electronics and Applications. 3. 233-249. 10.3390/jlpea3030233.

[3] M. Vohrmann, S. Chatterjee, S. Lütkemeier, T. Jungeblut, M. Porrmann and U. Rückert, "A 65 nm standard cell library for ultra low-power applications," 2015 European Conference on Circuit Theory and Design (ECCTD), 2015, pp. 1-4, doi: 10.1109/ECCTD.2015.7300041.

[4] P. D. Schiavone, D. Rossi, A. Pullini, A. Di Mauro, F. Conti and L. Benini, "Quentin: an Ultra-Low-Power PULPissimo SoC in 22nm FDX," 2018 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S), 2018, pp. 1-3, doi: 10.1109/S3S.2018.8640145.

[5] K. Singh and J. Pineda de Gyvez, "Twenty Years of Near/Sub-Threshold Design Trends and Enablement," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 1, pp. 5-11, Jan. 2021, doi: 10.1109/TCSII.2020.3040970.

[6] B. Dokic and A. Pajkanovic, "Low power CMOS sub-threshold circuits," 2013 36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2013, pp. 60-65.

[7] M. L. de Oliveira, F. S. Torres and K. da Silva Fernandes, "Cell Library Design for Ultra-Low Power Internet-of-Things Applications," 2018 3rd International Symposium on Instrumentation Systems, Circuits and Transducers (INSCIT), 2018, pp. 1-6, doi: 10.1109/INSCIT.2018.8546716.

[8] Behzad Razavi. 2000. Design of Analog CMOS Integrated Circuits (1st. ed.). McGraw-Hill, Inc., USA.

[9] S. Luetkemeier, T. Jungeblut, M. Porrmann and U. Rueckert, "A 200mV 32b subthreshold processor with adaptive supply voltage control," 2012 IEEE International Solid-State Circuits Conference, 2012, pp. 484-486, doi: 10.1109/ISSCC.2012.6177101.

[10] (Oct. 2020). Ambiq Micro, Inc. [Online]. Available: http://www.ambiq.com/

[11] (Oct. 2020). Everactive. [Online]. Available: https://everactive.com/

[12] A. Wang and A Chandrakasan, "A 180-mV Subthreshold FFT Processor Using a Minimum Energy Design Methodology" IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 40, NO. 1, JANUARY 2005

[13] B. Zhai et al. "A 2.60pJ/Inst Subthreshold Sensor Processor for Optimal Energy Efficiency" 2006 Symposium on VLSI Circuits Digest of Technical Papers IEEE

[14] S. Lütkemeier and U. Rückert, "A Subthreshold to Above-Threshold Level Shifter Comprising a Wilson Current Mirror," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 57, no. 9, pp. 721-724, Sept. 2010, doi: 10.1109/TCSII.2010.2056110.

[15] J. Parimala, K. Priyanka, L. S. Kaumudi, K. S. Pande and N. S. Murty, "Subthreshold voltage to supply voltage level shifter using modified revised wilson current mirror," 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2015, pp. 1-5, doi: 10.1109/ICCIC.2015.7435749.

[16] S. Hanson, M. Seok, D. Sylvester and D. Blaauw, "Nanometer Device Scaling in Subthreshold Logic and SRAM," in IEEE Transactions on Electron Devices, vol. 55, no. 1, pp. 175-185, Jan. 2008, doi: 10.1109/TED.2007.911033.

[17] Chromite Core Generator, InCore Semiconductors India[3]

[3] We use the Chromite Core Generator to generate the RISC-V core that has been synthesized for evaluating the performance of our standard cells.