

Original Article

Power-Aware RTL Design Techniques for High-Performance ASICs

VenkataKrishna Brahmam Pogadadanda

Senior RTL Design Engineer at AMD-Xilinx, USA.

Received: 07-07-2024

Revised: 07-28-2024

Accepted: 08-05-2024

Published: 08-08-2024

ABSTRACT

The increase in transistor density, frequency, and integration level inside circuits have exposed power consumption issues in high-end ASIC design. Hence, power has become one of the key design objectives together with performance and area of silicon. Power optimization at the RTL level has brought out as the most efficient measure since almost 60% of the chip's power can be saved through early power optimization done at Register Transfer Level (RTL) design phase and at the same time it is cost effective as the redesign time at later stages is totally avoided. In this paper, a variety of power-aware RTL techniques and methodologies have been studied and their implementation in ASIC through which power dissipation both dynamic as well as leakage can be curtailed using clock gating, power gating, operand isolation, multi-threshold voltage design, dynamic voltage and frequency scaling (DVFS), resource sharing, and optimized state machine encoding collectively, have been identified. Besides these, the paper also discusses the compromise between power, performance, and area (PPA) and how proper RTL libraries can aid in energy efficiency without considerably tapping computational throughput. Besides case study assessment, simulation-based power analysis and synthesis-driven optimization have been combined to highlight the practical viability of these power-saving methodologies of the high-performance ASICs. The conclusion drawn from the results of the study was that power consumption in the early right RTL stage can be low significantly even after maintaining timing and functional accuracy. Also, the paper talks about future trends, e.g., AI-based power optimization, adaptive low-power architectures, and new semiconductor technologies which are likely to influence the new generation of energy efficient ASIC design.

KEYWORDS

ASIC Design, RTL Optimization, Low-Power VLSI, Dynamic Power Reduction, Leakage Power, Clock Gating, Power Gating, DVFS, Multi-Vt Design, High-Performance Computing, System-on-Chip (SoC), EDA Tools.

1. INTRODUCTION

1.1. Challenges

Designing power-aware RTL at the ASIC level has become very difficult due to the fast development of semiconductor technology and the complexity of today's ASIC architectures. One of the main problems is dynamic power consumption that is the result of frequent switching activity in digital circuits. Since clock networks receive power continuously and run all over the chip, they are responsible for a large portion of dynamic power dissipation, especially in high-rate designs. Performance requirement increase is synonymous with switching activity, which leads energy consumption to climb and creates thermal stress. Another issue that cannot be ignored is leakage power in deep submicron technologies. Different transistor leakage currents (subthreshold leakage, gate oxide leakage) have grown with transistor scaling and now are a problem even when the circuit is idle. Besides adding to leakage in very small geometry devices, process variations in advanced technology nodes make leakage estimating and power optimizing uncertainties even more critical.

Another considerable challenge in ASIC design is balancing performance and power efficiency. Low-power techniques e.g., clock gating, power gating, and voltage scaling, which reduce power consumption, can sometimes introduce timing penalties, additional hardware complexity, and area overhead. Designers, therefore, need to find timing closure with the help of power budgets that are still acceptable and silicon area that should be at a minimum. Besides, the complexity of modern System-on-Chip (SoC) architectures has created integration challenges because there are multiple voltage domains, heterogeneous processing units, and advanced communication interfaces. Besides being different, systems must be highly reliable and functional. Power modes that have to be validated are increased by standards such as Unified Power Format (UPF) and Common Power Format (CPF) which in turn increase verification overhead and simulation complexity, making RTL power optimization a difficult while still quite necessary task in modern ASIC development. Moreover, power-aware system verification is a tough task.

1.2. Problem Statement

Today's ASICs have to provide incredibly high computing performance, and at the same time, they must meet power and thermal restrictions. Traditional RTL design practices pay a lot of attention to the functionality, timing, and hardware performance, while efficient power management is often left out at the early design stages. Therefore, a lot of ASIC implementations have excessive dynamic and static power consumption issues, which result in higher heat output, lower reliability, and increased running costs. Although the existing low-power design methods are capable of reducing power consumption, very often they bring along timing degradation, increased verification complexity, extra design overhead, and bigger chip area. Besides, the introduction of advanced semiconductor process technologies has exacerbated leakage current problems, and now conventional methods are no longer capable of meeting balanced power-performance-area (PPA) requirements.

Hence, there is a tremendous need for power-aware and optimized RTL design methodologies which can not only reduce the dynamic and leakage power consumption effectively but also maintain system performance, and keep the area overhead minimum. Besides, these methodologies must be capable of scaling with deep submicron technologies and modern heterogeneous SoC architectures. The difficulty is in coming up with RTL-level optimization methods which are both efficient and able to provide power savings at the early stage of design, and at the same time, not cause negative effects on timing closure, functional correctness, or overall design complexity.

1.3. Motivation

The ever-increasing necessity for energy-efficient computing systems has significantly driven research in power-aware ASIC design. New forms of AI accelerators, edge computing devices, mobiles, automobiles, and high-performance servers need great computational power, yet with small energy consumption. To accommodate the increasing transistor count and higher clock speeds, power density and heat dissipation issues have been extensively considered in the design of the current chips. Data centers and cloud computing infrastructures account for huge electrical power consumption which leads to high operation costs related to cooling and also environmental impact. This is why the semiconductor industry is constantly in search of low-power design solutions for the purpose of energy efficiency and cost reduction.

Besides that, an important cause for motivation is the requirement for power optimization at the earliest design stage. The RTL level is the most flexible design stage since architectural and coding decisions can still be changed before performing synthesis and physical implementation. Power optimization at this level leads to a drastic decrease in costly redesign iterations during the ASIC development cycle. Besides that, suppliers are power-efficient chips in the semiconductor market by enhancing battery life, reliability, thermal characteristics, and overall system efficiency. Those reasons drive the investigation of powerful RTL-level power optimization methods for the future high-performance ASIC designs.

2. LITERATURE REVIEW

2.1. Overview of Low-Power VLSI Design

The design of low-power VLSI circuits is a focus of ongoing research due to the explosion of portable electronics, AI systems, data centers, and high-performance computing. Initially, semiconductor technologies were concerned only with boosts in speed and increases in transistor density following Moore's Law. The endless scaling allowed billions of transistors to be placed on one chip, which greatly enhanced the computational capability. However, as the sizes of transistors entered the domain of deep submicron technologies, power consumption and generation of heat became two major issues. This finally resulted in a "power crisis" in VLSI systems where higher performance meant excessive energy consumption and thermal problems.

In an attempt to beat these limitations, the research fraternity moved away from a pure performance-based design methodology and partially to a power-performance trade-off. Modern ASICs, therefore, not only require a high computational throughput but that they are energy-efficient, reliable, and with lower need for operational costs. The RTL-level optimization stage has gained more importance because feature choices at the RTL stage have a very strong impact on switching activity, clock behavior, and the total power consumption of the chip. So, low-power RTL methods are adopted in the design of modern ASICs.

2.2. Sources of Power Dissipation in CMOS ASICs

In CMOS ASICs, the major source of power consumption is power dissipation, which consists mainly of dynamic power as well as static power. Dynamic power is generated due to switching activity during a change of logic. When the frequency of operation of the circuit is higher, and a large number of transistors are used, the quantity of switching activity increases, and hence, more energy is consumed. As during the operation of the circuit, the clock networks are constantly toggled, they constitute a significant source of dynamic power. Consequently, one of the main goals of low-power RTL design is to minimize the switching activity that does not contribute to the computation.

On the other hand, static power or leakage power has been a major component especially with advanced semiconductor technologies. Leakage current is mainly due to subthreshold conduction, gate oxide tunneling, and junction leakage. As transistor sizes get smaller and threshold voltages are reduced, leakage currents can increase drastically even when the circuits are not being used. Hence, technology scaling brings new problems of leakage power consumption while keeping the performance of the circuits high. Still, researchers are constantly finding the best techniques to reduce both dynamic and static power of the modern ASIC systems.

2.3. Existing RTL Power Optimization Techniques

Many RTL-oriented power optimization methods have been invented to bring down the power consumption of ASICs. One of the best-known ways, clock gating, is very popular since it turns off the clock signal to parts of the circuit that are not working, thus saving power. Fine-grain clock gating only works at the level of separate registers, whereas coarse-grain clock gating can turn off whole parts of the circuit. Another very effective method is operand isolation which cuts down the switching activity of datapath circuits by preventing unnecessary data transitions.

Power gating methods put to zero the leakage power by isolating the inactive parts of the circuit using sleep transistors. Retention registers are a very typical example of a design element able to retain the content of the register while the circuit goes to the low-power state. Multi-voltage design is the practice of creating voltage islands, each one capable of being at a different supply voltage; thus, it is a strategy to trade-off between performance and power. Dynamic Voltage and Frequency Scaling (DVFS) is a technique that, according to workload, changes the voltage and frequency of the processor, thereby giving better energy efficiency. Multi-Threshold CMOS (MTCMOS) technology is

where transistors having different threshold voltages are used together in order to cut down leakage while at the same time not sacrificing system performance.

Recently, EDA tools like Synopsys Power Compiler, Cadence Genus, and Siemens EDA solutions have been made capable of performing RTL power optimization along with synthesis and verification in a largely automated way. Such tools enable designers to review power consumption at the early stages of the design and make the ASICs more energy-efficient.

Table 1: Small Comparative Table of RTL Power Optimization Techniques

Technique	Main Purpose	Power Type Reduced
Clock Gating	Disable inactive clock paths	Dynamic Power
Operand Isolation	Reduce unnecessary switching	Dynamic Power
Power Gating	Shut down idle blocks	Leakage Power
DVFS	Adjust voltage and frequency	Dynamic & Leakage
Multi-Vt Design	Use multiple threshold voltages	Leakage Power

2.4. Comparative Analysis and Research Gaps

According to the literature, methods like clock gating, DVFS, and power gating are practical measures for power saving in ASICs. Clock gating is a widely adopted power saving technique, which aims at disabling the clock signal to those flip-flops that have no data transition requirement, thereby minimizing the dynamic power consumption. DVFS is a power management technique that adjusts the voltage and frequency depending on the workload, resulting in improved energy efficiency. Power gating, on the other hand, significantly reduces the leakage current by completely switching off the supply voltage to the idle portion of the circuit. On the other hand, these methods come with downsides, e.g. timing overhead, complexity of the design, latency of wake-up, and difficulties in verification, etc.

Although quite a number of work done so far mainly covers optimization of power consumption by utilizing only one method, there are very few studies that consider integration of the two or more techniques and also develop an RTL-centric methodology. Besides, only a few works have touched on issues of the scalability for next generation technologies and heterogeneous SoC architectures. Verification, on the other hand, remains to be a major bottleneck as power-aware systems require validation at multiple operating modes and transitions. Hence, the development of scalable RTL methodologies that integrate multiple low-power techniques without loss of performance, area overhead, and verification complexity should attract more research work in the future.

3. PROPOSED METHODOLOGY

3.1. Overview of the Proposed Power-Aware RTL Framework

The methodology outlined here is a power-conscious, one-stop shop for a high-performance ASIC RTL framework. This framework aims at lowering both dynamic and leakage power but at the same time keeping the circuit's timing performance and area efficiency unchanged. This

methodology fits smoothly into a conventional ASIC design flow starting with RTL coding and simulation and going through synthesis, power analysis, verification, and ultimately implementation. In contrast to old-school RTL techniques which mainly concentrate on functionality and timing closure, the new framework sees power optimization as a key design goal right from the first steps of development.

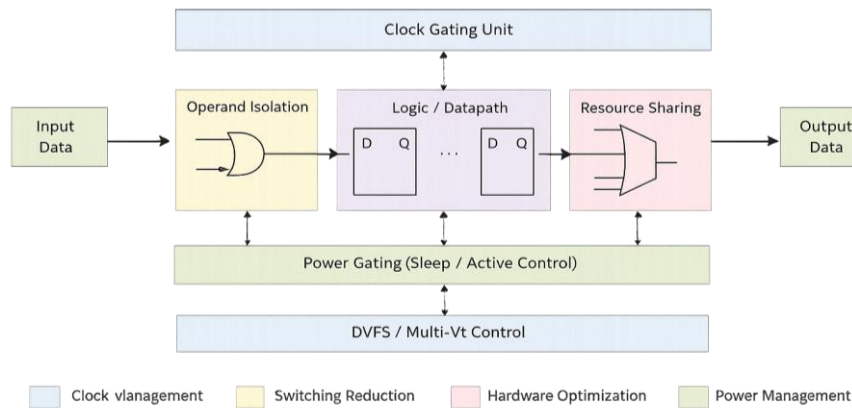


Fig 1 - Proposed Power-Aware RTL Architecture for High-Performance ASICs

The design flow starts with the development of RTL architecture and then focuses on finding the main power-hungry modules and components. Techniques for power saving, such as smart clock gating, operand isolation, FSM optimization, and resource sharing, are subsequently introduced in a structured manner. Once the optimization has been done, the switching activity is taken out from the simulation-generated activity files like SAIF or VCD. Power estimation tools evaluate the power consumption of the design and effectiveness of the techniques implemented and help in finding additional opportunities for optimization. Low-power and functional verification are the last two steps in this framework that helps in reaching the power goals without compromise in correctness and timing behavior of the system. This framework not only saves power at a very early stage but also reduces the number of redesign iterations at the later implementation stages.

3.2. Design Objectives and Optimization Strategy

The main goal of the methodology being put forward is to have an impact on dynamic power which is generated through the switching activity of both the sequential and combinational circuits. Considering that clock networks can be accountable for a large share of the total ASIC power consumption, making changes to the clock so that unnecessary transitions are minimized is the focus of the optimization. Being able to control leakage power in technologies beyond 0.18 microns is another target due to the fact that leakage current is the determining factor for the power consumed during standby.

Besides bringing down the power, the method also seeks to keep up with the timing and area criteria. It is a well-known fact that the logic gates, multiplexers or control circuitry that are used to

implement the low power techniques are almost always detrimental to the timing closure and chip area. For this reason, the rest of the paper will discuss how power optimization is only one of a set of competing objectives, performance and silicon area constraints. Another factor that is taken into account and deemed of high importance is scalability, as nowadays ASICs have multiple processing units, heterogeneous architectures, and complex communication interfaces. The method is capable of being embedded in the design flow of advanced System-on-Chip (SoC) environments at the same time it keeps the flexibility of the design and the simplicity of verification.

3.3. RTL Power Optimization Techniques

The proposed framework integrates a number of RTL-level power optimization techniques. Intelligent clock gating is one of the fundamental techniques that help to reduce switching activity that is not required. Automatically-generated signals work to stop clock propagation to registers or modules that are inactive. The use of conditional register update logic implies that registers will only toggle if data changes are needed. Typically, an RTL implementation contains one or more conditional statements that only update registers when enable signals are active, which decreases the dynamic power consumption.

Operand isolation is another key technique that aims to minimize unnecessary activities in a datapath circuit by controlling redundant signal transitions. The isolation logic separates the signals which prevents redundant transitions at arithmetic and logic units during idle states which, in turn, reduces the switching power. The methodology also includes FSM optimization since state transitions are a major source of switching activity. To reduce transition activity and improve power efficiency, state encoding methods that are more efficient such as Gray encoding and one-hot encoding are utilized. Transition minimization techniques are applied to get rid of unnecessary switching between states.

Resource sharing is one of the ways to cut down hardware redundancy, allowing different operations to share the same arithmetic units such as adders and multipliers. Multiplexer-based control logic makes the selection of the required operation on the fly, resulting in the reduction of the area and power consumption. Also, power-aware pipelining opens up a new opportunity to produce more while having fewer glitches and transient switching activities. The right pipeline stage balancing enables reduction in the activity of combinational paths and leads to energy efficiency that is better overall, without one's system performance being severely impacted.

3.4. Power Management and RTL Analysis Flow

They aimed to organize a power management module that works with power domains primarily. The always-on blocks are the parts of the chip that need to be active all the time, a switchable block can be switched off when the device goes to sleep, thus saving the leakage power. The power domains can be controlled independently and in a very fine way. This kind of power control is something that the complex SoC environment can really use and it can also lead to more efficient use of energy.

So, they have come up with retention techniques which basically store register states in reserve during sleeping modes. For example, if certain registers need to keep the information while their corresponding functional blocks are being powered down, one way to accomplish what you claim is to have special registers for retention. When the system wakes up from the sleep mode, it is much quicker to come to its working state. At the same time, they also made use of adaptive voltage scaling techniques which can change the supply voltage on the fly according to how much work the device is doing. When there is less demand for processing, the operating voltages are reduced, which leads to a drop in both dynamic and static power consumption.

Development of the RTL power estimation process starts after functional simulation that generates switching activity data. Simulation-induced outputs like SAIF and VCD provide the information about which signals have switched and how often throughout the entire design. They are then loaded into power tools for a detailed breakdown of power at the module-level and system-level, respectively. The illumination of primary power hungry spots is the byproduct of the power characterization. As a result, one is not limited to the simple identification of parts of the chip with very high power consumption, but one is open to a whole series of steps of improving it with the help of the RTL development cycle.

3.5. Verification Methodology and Optimization Flow

Verification is critical to the proposed power-aware RTL environment. Power-saving methods often cause functional and timing breakdowns. Functional verification is required at the first level to verify the correctness of the design and to verify that the power-saving methods used do not affect the design fundamentally. Testbenches are developed to verify all the functionalities and also the various switching conditions at different levels of workloads.

Afterwards, low-power verification is performed to check if clock gating behavior, power domain transitions, retention mechanisms, and voltage scaling operations are working correctly. Unified Power Format (UPF)-based verification methodologies are employed to describe power intent and verify correct interaction between different power domains. Verification tools detect errors due to isolation, retention failures, incorrect power transition, and wake-up sequencing issues. This way, the system is safeguarded to operate correctly in all power management scenarios.

The suggested optimization flow is a systematic step-by-step approach. Initially, a switching analysis is run to identify modules with high activity. Then, based on the functionality of a module, the timing requirements, and the power characteristics, different power-saving techniques can be independently implemented. Power analysis and synthesis are performed after each optimization phase to both evaluate the achieved benefits and identify areas for potential improvements. This decision-making process that is continuously carried out effectively reduces power consumption without compromising timing performance, functional correctness, and area efficiency of modern high-performance ASIC designs.

4. CASE STUDY

4.1. Introduction to the Case Study

To test the efficiency of a power-aware RTL approach, a study on the AI accelerator block was taken up. Modern applications of AI accelerators include machine learning, image processing, autonomous systems and edge computing devices. Such high-performance systems need to be power and thermally efficient at the same time. The prolonged arithmetic processing and heavy data transfer activities are the reasons why AI accelerator architectures consume a lot of dynamic and leakage power and hence their suitability to analyze RTL-level power optimization techniques.

The chosen accelerator unit consists of arithmetic processing units, control logic, memory interfaces and data communication modules. The design executes matrix multiplication and parallel data processing which are the characteristic operations of neural network computations. Due to the fact that these operations cause a lot of switching and high-frequency clocking, the design serves as an excellent test for the power saving strategies. The case study is a performance comparison between the baseline RTL implementation and the optimized power-aware RTL.

4.2. Baseline RTL Design and Functional Specifications

The basic RTL architecture was implemented in a standard ASIC design flow and initially it did not have any advanced low-power optimization techniques incorporated. The design contained multiple ALUs, register banks, control FSMs, and pipelined datapath units to increase compute throughput. It was powered by a single clock distribution network, and all modules operated at standard voltage levels.

The functional requirements of the accelerator were: parallel execution ability, high-speed data transfer, minimal latency in operations, and continuous workload support. The design aimed to deliver top-notch performance for matrix computations, ensuring proper timing closure even at high frequencies. But, the initial version showed quite high switching activities in clock networks, arithmetic units, and state transition logic which led to an escalation in dynamic power consumption. On top of that, leakage power is very significant during the idle phase since there are so many integrated transistors.

The baseline performance goals were to keep the system stable at the highest frequencies, maintain comfortable timing margins, and ensure that the functionality is correct. The first design is also used as a benchmark when it comes to checking out the effects of the power-aware RTL optimization techniques proposed later.

4.3. Implementation Environment and Experimental Setup

Case study implementation and analysis are conducted in a high tech semiconductor node such as 14nm CMOS technology, which is a typical example of modern ASIC manufacturing processes where issues with leakage power and heating become very critical. Therefore, this specific

technology node is selected. Industry leading Electronic Design Automation (EDA) tools are used in the entire implementation process to ensure that the analysis and optimization are realistic.

RTL synthesis and power optimization are done by using Synopsys Power Compiler and Cadence Genus synthesis tools. Functional simulation and switching activity analysis are carried out in simulation environments producing SAIF and VCD activity files. Besides that, timing analysis and verification are done to ensure that low-power changes do not lead to performance degradation or loss in functional correctness.

The experiments involve various benchmark workloads representing different usage scenarios of the AI accelerator. These workloads are low-utilization, moderate-utilization, and peak-performance conditions to observe power behavior changes due to varying switching activities. Among other performance metrics, dynamic power, leakage power, timing delay, area overhead, and throughput are measured both pre- and post-optimization. This experimental structure allows a thorough evaluation of the suggested method under practical ASIC operating conditions.

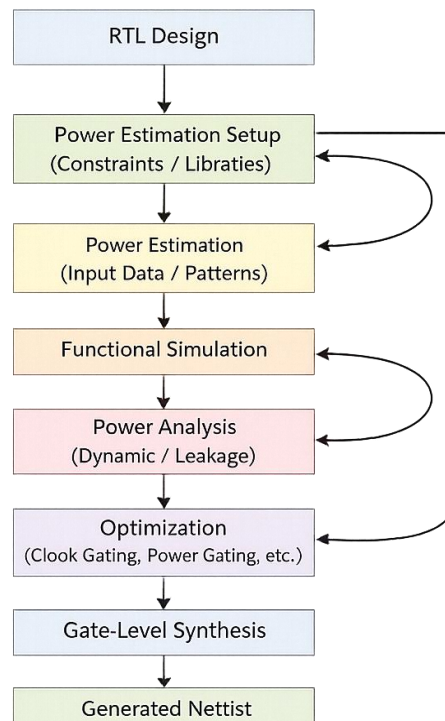


Fig 2 - Simulation-Based Power Optimization Flow for ASIC Design

4.4. Applied Power Optimization Techniques

Several power optimization techniques at RTL level have been applied jointly on the initial architecture to achieve the reduction of total power consumption. One of such an implementation is to intelligently gate the clock which results in disabling the propagation of clocks to inactive registers as well as processing modules in case of idleness. Conditional enable logic is introduced in sequential

circuits so as to keep changes in the clock network to a minimum by limiting unnecessary switching activities.

Operand isolation methods are used in arithmetic datapaths in order to limit and suppress extra transitions in the case of processing units, which are not performing calculations, being idle. Isolation logic stops undesirable signal propagation in multipliers, adders as well as data routing modules and thus leads to power saving in terms of dynamic power dissipation. Besides, FSM has been optimized by cutting down the number of unnecessary state transitions and getting rid of the excess state switching activity within control logic by means of the application of efficient state encoding schemes.

Multi-voltage design is realized by partitioning the system into different power domains where different supply voltages are used depending on performance needs. Processing units which are highly critical work at a higher voltage level for optimization of performance whereas less important modules use lower voltages so as to save energy. Selective power gating of idle modules during low-activity periods is one of the ways to reduce leakage in addition. Collectively, these methods bring about significant improvements in power efficiency at the overall level without making any sacrifices in system functionality and timing performance.

4.5. Power-Aware RTL Architecture and Analysis

The optimized power-aware RTL architecture introduces dedicated power management logic, multiple voltage domains, and enhanced control circuitry to enable low-power operation. The modified architecture disconnects the always-on blocks from the switchable power domains, so that the inactive processing units can be put into the low-power sleep mode during the idle periods. Retention registers are used to keep the state of the main registers stable when the different blocks are temporarily powered down.

The optimized architecture leads to a remarkable decline in the switching of datapath and control modules as compared to the baseline implementation. Clock gating is the method that changes the clock path so as to minimize the number of clock transitions, operand isolation and FSM optimization are the two methods that even further reduce the internal switching activity. Besides, the presence of multiple voltage domains and power gating leads to a drastic decrease in the leakage power, without having any detrimental effect on the overall computational throughput.

The conducted experiments reveal the fact that the suggested approach strongly decreases both dynamic and static power dissipations while keeping the timing closure and the area overhead within the limit. The findings confirm that the optimization of RTL at the early stages is one of the most potent ways of increasing the energy efficiency of the contemporary high-performing ASIC designs. In addition, the case study demonstrates the ability of the proposed framework to be expanded and used for future AI accelerators and next-generation heterogeneous SoC architectures.

5. RESULTS AND DISCUSSION

5.1. Power Analysis Results

The implementation of the proposed power-aware RTL methodologies has brought about a great improvement in the level of power efficiency of ASICs as a whole. Close examination of the changes made in the design of the baseline architecture and the optimized one revealed a drastic drop in the consumption of dynamic power attributable to the diminished switching activities in the clock networks, the datapath modules as well as the control logic. Using intelligent clock gating, quite simply, inactive registers and functional units are no longer given unnecessary clock transitions during periods of idleness. In a similar manner, operand isolation has greatly reduced unnecessary switching activities in arithmetic blocks, particularly in low-utilization operating conditions.

There was a noticeable decrease in leakage power as well following the provision of power gating and multi-voltage techniques. In sleep mode, the modules which are not in use are physically isolated from the power supply using sleep transistors, resulting in a significant decrease in the leakage current during standby. Retention registers help to maintain important state information while at the same time, they allow modules to safely enter low-power states. The final design showcased that not only does it use less energy during active periods of time but it is also capable of saving power while idle, which in essence is a proof of the effectiveness of coupling a number of RTL-level optimization techniques within one single framework.

5.2. Timing Performance and Area Evaluation

Timing analysis was carried out to see the impact of low-power optimization techniques on the system performance. After analyzing the critical path, we realized that the optimized design can still successfully finish timing with only a slightly increased delay due to the usage of more gating logic and power management circuitry. The good placement of clock gating logic minimized timing losses, whereas upgrading of the data path design was effective in reducing the unnecessary combinational transitions and glitch propagation. The end design was capable of reaching the required operational frequency for AI accelerator workloads comfortably.

When comparing the frequencies of the baseline and optimized versions, there proved to be only a minor reduction in performance even though the latter implemented aggressive power-saving measures. This clearly shows that methods of lowering power consumption may be combined in an efficient way without having a major impact on throughput or computational efficiency. Nevertheless, some area overhead was caused by the need for retention registers, isolation cells, sleep transistors, and extra control logic. Despite the fact that the optimized design took up a somewhat larger silicon area than the baseline version, the increment was still well within the typical limits of the industry, especially considering the very significant power savings we achieved. Hence, the trade-off between area and power efficiency is generally seen as favorable for high-performance ASIC applications.

5.3. Comparative Results and Performance Analysis

A closer look at the baseline and optimized architectures through a thorough comparison unveils the upgrades brought by the suggested approach. Because of the clock gating and switching activity suppression methods, the dynamic power consumption took the largest hit (downfall). Power gating and multi-voltage implementation also made a very significant impact in reducing leakage power. In terms of timing performance, only a slight decrease in the maximum operating frequency was noticed whereas, the area overhead registered a moderate rise due to extra low-power circuitry.

Table 2: Comparative Power, Performance, and Area (PPA) Analysis of Baseline and Optimized ASIC Designs

Metric	Baseline Design	Optimized Design	Improvement
Dynamic Power	520 mW	360 mW	30.7% Reduction
Leakage Power	140 μ W	82 μ W	41.4% Reduction
Frequency	950 MHz	920 MHz	Minimal Impact
Area	5.2 mm ²	5.5 mm ²	5.7% Increase

The results show that the optimized RTL framework is able to achieve the right balance between power efficiency and performance/area requirements. Low active and standby power levels in conjunction imply that the integration of various optimization techniques is more effective than depending on a single low-power strategy.

5.4. Graphical Interpretation and Trade-Off Analysis

This gives a very clear understanding of improvements we achieve by our new design when we plot the experimental data. Take dynamic power consumption charts as example, they show a significant reduction in the dynamic power that processors consume during the times they are directly working. On the contrary, leakage standby plots indicate that the component of leakage power when the system is in idle has decreased significantly. Timing difference figures suggest the revamped architecture maintains a performance level almost equal to the original implementation. Area utilization results also show a small increase in chip area owing to the addition of extra low-power management circuits.

Studying the energy-delay product indicates that the new architecture not only enhances energy efficiency but also keeps computational speed at a good level. In fact, the biggest dynamic power savings were recorded in intelligent clock gating since clock networks accounted for a major share of switching activity. In terms of leakage reduction, power gating was very effective, especially when the circuits had lower activity levels. Operand isolation coupled with FSM optimization helped in reducing switching activity in the datapath as well.

Trade-off assessment sheds light on the fact that a very strong power optimization may bring in extra logic overhead and minor timing penalties. Still, the power savings gained significantly exceed these drawbacks in the majority of the new ASIC designs where energy efficiency and thermal management are major factors.

5.5. Scalability and Industrial Applicability

The power-aware RTL methodology proposed in this paper is capable of scaling efficiently with new semiconductor technologies as well as complex System-on-Chip (SoC) architectures. As the transistor size gets smaller in process technologies, leakage power and thermal issues will be a big challenge. However, the optimization methods utilized will still work effectively in advanced nodes as they tackle switching activity and standby leakage at the same time. So, the methodology can even be modified to cater to the development of high-performance AI accelerators, networking processors, automotive controllers, and data-center ASICs in the future.

Looking at it from an industrial point of view, the suggested framework fits well with the design requirements of the modern semiconductor industry where power efficiency is a major differentiating factor. The manufacturers of mobile processors, edge AI devices, and cloud computing hardware are in need of low-power architectures that can sustain high computational throughput within the constraints of limited power budgets. The RTL-centric optimization technique is particularly beneficial as it allows for power reduction at an early phase, thereby minimizing expensive redesign work during the physical implementation phase.

The paper illustrates the fact that combining clock gating, operand isolation, power gating, FSM optimization, and multi-voltage methods into a single RTL framework not only makes the deployment of ASICs feasible but also provides a scalable solution in the industry. These findings show how crucial power-aware RTL methodologies are in optimizing power-performance-area in a balanced manner for the next generation semiconductor systems.

6. CONCLUSION AND FUTURE SCOPE

6.1. Conclusion

This work presents an extensive investigation into power-aware Register Transfer Level (RTL) design techniques for High-Performance Application-Specific Integrated Circuits (ASICs). It mainly targeted lower power consumption at the same time keeping timing performance and area efficiency. The research also looked at the main power drain of CMOS ASICs, specifically dynamic and leakage powers. Besides, it pointed out the increasing role of low-power design techniques in present-day semiconductor devices. The paper presented various RTL optimization techniques such as Smart Clock Gating, Operand Isolation, FSM Optimization, Resource Sharing, Power Gating and Multi-Voltage Designs. These were then incorporated in a single power-aware RTL framework.

The suggested approach led to notable enhancement of the total ASIC energy efficiency as proved through the in-depth case study with a reference to an AI accelerator hardware. The AI hardware was used to demonstrate the ability of the proposed method to achieve substantial power savings, both dynamic and leakage, while sacrificing the frequency of operation insignificantly and leaving the area overhead reasonably high in the first instance. It has been found that the optimization of RTL at an early stage offers more freedom and leads to less expensive redesigns at the time of ASIC implementation. It was further underlined by the work that the trade-offs of power,

performance, and area limitations are indispensable components of any design using present-day semiconductor technologies. In short, the introduced framework was validated as a capable tool for designing scalable, energy-efficient, and high-performance ASICs that fit well AI computing, networking, mobile systems, and data centers.

6.2. Future Scope

One avenue for research arising from the present work is AI-assisted RTL optimization where machine learning methods would be capable of independently helping in the synthesis and power optimization decision processes. Adaptive runtime power management would be another topic of research that results in self-tuning, power-aware ASIC architectures that change their power profiles dynamically in response to workload changes. Besides, near-threshold computing has been one of the most attractive directions for ultra-low-power operation in energy-limited devices.

With the continued miniaturization of semiconductor technologies towards 3nm and smaller nodes, issues such as leakage current, thermal reliability, and process variation will pose a challenge requiring the implementation of advanced low-power techniques. Besides, future ASICs are anticipated to be based on chiplet heterogeneous architectures giving rise to the evolution of power-conscious interconnects and decentralized power management systems. These new trends will be essential in directing the development of low-power ASIC design methodologies in the future.

REFERENCES

- [1] Ahuja, S. (2010). High level power estimation and reduction techniques for power aware hardware design.
- [2] Muslim, F. B., Qamar, A., & Lavagno, L. (2015, September). Low power methodology for an ASIC design flow based on high-level synthesis. In *2015 23rd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)* (pp. 11-15). IEEE.
- [3] Suryadevara, S. S. K., & Shaik, K. (2023). Real-Time Anomaly Detection and Attack Mitigation for Cloud-Based Content Delivery Paths Using AI. *International Journal of Emerging Research in Engineering and Technology*, 4(1), 175-185. <https://doi.org/10.63282/3050-922X.IJERET-V4I1P119>
- [4] Vppalapati, M. (2023). When Identity Decisions Throttle Data Movement. *International Journal of Emerging Research in Engineering and Technology*, 4(3), 160-170. <https://doi.org/10.63282/3050-922X.IJERET-V4I3P117>
- [5] Ahuja, S., Lakshminarayana, A., & Shukla, S. K. (2011). *Low power design with high-level power estimation and power-aware synthesis*. Springer Science & Business Media.
- [6] Muppaneni, R. K. (2023). Low-Code Revolution: How Power Platform Extends Dynamics 365 Capabilities. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 4(3), 162-171. <https://doi.org/10.63282/3050-9262.IJAIDSML-V4I3P119>
- [7] Takkalapally, D. (2023). HoloSearchAI: AI-Driven Latency Optimization Framework for Distributed Search Systems. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(3), 217-227. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I3P122>
- [8] Crone, A., & Chidolue, G. (2007, September). Functional verification of low power designs at RTL. In *International Workshop on Power and Timing Modeling, Optimization and Simulation* (pp. 288-299). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [9] Vanapalli, Satyendra Kumar. "Reducing Financial Fraud Using Machine Learning and CRM Data Models." *International Journal of Machine Learning and Predictive Analytics* 6.2 (2023): 01-20.
- [10] Shiramalla, R. (2022). Predictive Record Assignment Engine in Salesforce using LWC and Einstein AI. *International Journal of AI, BigData, Computational and Management Studies*, 3(3), 147-159. <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V3I3P117>

- [11] Mazzoni, L. (2003). Power-aware design for embedded systems. *Electronics Systems and Software*, 1(5), 12-17.
- [12] Srigadde, B. R. (2023). The Hidden Gem: Lightning Headless Component. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(1), 244-254. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I1P125>
- [13] Parakala, A. (2023). Citizen-Facing Automation: Chatbots and Self-Service in Public Services. *International Journal of AI, BigData, Computational and Management Studies*, 4(4), 108-118. <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V4I4P112>
- [14] Pedram, M., & Rabaey, J. M. (Eds.). (2002). *Power aware design methodologies*. Springer Science & Business Media.
- [15] Allenki, S. S. (2023). Applying Cloud Security Best Practices in Regulated Environments. *American International Journal of Computer Science and Technology*, 5(3), 48-60. <https://doi.org/10.63282/3117-5481/AIJCSIT-V5I3P105>
- [16] Paim, G., Ferreira, R. S., Rocha, L. M., da Costa, E. A., Alves, T. G., & Bampi, S. (2017, August). A power-predictive environment for fast and power-aware ASIC-based FIR filter design. In *Proceedings of the 30th Symposium on Integrated Circuits and Systems Design: Chip on the Sands* (pp. 168-173).
- [17] Katangoori, S., & Katangoori, A. (2025). Data-Centric AI in the Era of Large Volumes: Improving Model Outcomes through Data Quality Engineering. *American International Journal of Computer Science and Technology*, 7(4), 129-141. <https://doi.org/10.63282/3117-5481/AIJCSIT-V7I4P112>
- [18] Muppaneni, R. K. (2023). AI-Driven Forecasting in Dynamics 365 Sales: What Businesses Need to Know. *International Journal of AI, BigData, Computational and Management Studies*, 4(1), 168-176. <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V4I1P117>
- [19] Narayanaswamy, R. P. (2010). *Design of a Power-aware Dataflow Processor Architecture* (Doctoral dissertation, Virginia Tech).
- [20] Shiramalla, R. (2023). Optimizing Cross-Platform Enterprise Integrations Using Workato: A Case Study of Salesforce and Oracle SaaS Applications. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(1), 232-243. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I1P124>
- [21] Kumar Doodala, A. N., Thatraju, S., & Kankanala, V. (2023). Post- Pandemic QA evolution in Healthcare IT. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(2), 223-232. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I2P122>
- [22] Dick, R. P., Shang, L., & Jha, N. K. (2006). Power-aware architectural synthesis. In *The VLSI Handbook*. CRC Press.
- [23] Gaddam, R. R. (2022). Cost-Aware Autoscaling for Batch vs. Online Inference. *International Journal of Emerging Trends in Computer Science and Information Technology*, 3(4), 134-143. <https://doi.org/10.63282/3050-9246.IJETCSIT-V3I4P113>
- [24] Vppalapati, M., & Talasila, P. K. . (2023). Unobservable Performance: Storage Failures That Leave No Metrics Behind. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 4(4), 177-188. <https://doi.org/10.63282/3050-9262.IJAIDSML-V4I4P120>
- [25] Kommuru, Madhurima, and Appala Nooka Kumar Doodala. "Performance Bottlenecks Necks in Data Heavy Python Applications." *International Journal of Applied Data Science & Modern Computing* 6.2 (2023): 01-19.
- [26] Mohamed, C., Soukaina, M., & El Abidine, A. I. Z. (2019). Power aware setup timing optimization in physical design of ASICs. *Microelectronics Journal*, 83, 147-154.
- [27] Suryadevara, S. S. K., & Nakirikanti, S. (2023). Privacy-Preserving Personalization Using Federated Learning in AEM . *International Journal of AI, BigData, Computational and Management Studies*, 4(4), 190-199. <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V4I4P119>
- [28] Takkalapally, D., & Takkellapally, M. R. (2023). GC-TuneHFT: AI-Based Garbage Collection Optimization in High-Frequency Trading Environments. *American International Journal of Computer Science and Technology*, 5(6), 25-37. <https://doi.org/10.63282/3117-5481/AIJCSIT-V5I6P103>
- [29] Chentouf, M., Stevmelin, F., & Ismaili, Z. E. A. A. (2021). Power-aware hold optimization for ASIC physical synthesis. *Integration*, 76, 13-24. Vanapalli, Satyendra Kumar. "Data-Driven Decision Making: Integrating AI Tools with CRM Platforms." *International Journal of Applied Data Science & Modern Computing* 6.1 (2023): 01-21.
- [30] Muppaneni, K., & Vejella, M. (2023). Security and Data Privacy in Redux Stores. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 4(4), 153-162. <https://doi.org/10.63282/3050-9262.IJAIDSML-V4I4P117>
- [31] Srigadde, B. R. (2023). Creating Object Quick Actions with Lightning Web Components. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 4(2), 167-180. <https://doi.org/10.63282/3050-9262.IJAIDSML-V4I2P118>
- [32] Taraate, V. (2021). ASIC Design and Synthesis. *Springer Nature*.

- [33] Allenki, S. S. (2023). Reducing Security Vulnerabilities with Encryption, IAM, and Regular Audits. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(1), 265-275. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I1P127>
- [34] Parakala, A. (2023). Vendor Highlights - IoT, AI, and Process Mining. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(4), 135-146. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I4P115>
- [35] HA, Q. Y. L., NGUYEN, T. T., Dac-Binh, H. A., & Minh-Tam, N. L. (2019, July). High Performance and Security Design for Cryptosystem Using Simultaneous Multiple Hardware Threads and Power Aware Technique. In *2019 International Conference on System Science and Engineering (ICSSE)* (pp. 302-307). IEEE.
- [36] Katangoori, S., & Katangoori, A. (2023). Intelligent ETL Orchestration with Reinforcement Learning and Bayesian Optimization. *International Journal of Emerging Research in Engineering and Technology*, 4(4), 208-219. <https://doi.org/10.63282/3050-922X.IJERET-V4I4P123>
- [37] Kommuru, Madhurima. "Concurrency is hard: Java vs Python pitfalls you can't ignore." *International Journal of Modern Research in Science & Engineering* 5.2 (2022): 01-18.
- [38] Muppaneni, K. (2023). Virtual DOM vs Real DOM: Performance Benchmarks. *International Journal of AI, BigData, Computational and Management Studies*, 4(4), 180-189. <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V4I4P118>
- [39] Donno, M., Macii, E., & Mazzoni, L. (2004, April). Power-aware clock tree planning. In *Proceedings of the 2004 international symposium on Physical design* (pp. 138-147).
- [40] Kumar Doodala, A. N. (2023). Offline-First Android Architecture for waste management in low connectivity zones. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(1), 201-209. <https://doi.org/10.63282/3050-9246.IJETCSIT-V4I1P121>
- [41] Gaddam, R. R. (2022). Advanced Data & Model Drift Detection at Scale. *International Journal of AI, BigData, Computational and Management Studies*, 3(2), 124-136. <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V3I2P113>
- [42] Chakraborty, A., & Pan, D. Z. (2010, August). PASAP: Power aware structured ASIC placement. In *Proceedings of the 16th ACM/IEEE international symposium on Low power electronics and design* (pp. 395-400).
- [43] Taluri, R. (2022). Cloud Data Engineering Strategies for Large-Scale Financial Data Integration and Intelligent Corporate Performance Reporting. *International Journal of Emerging Research in Engineering and Technology*, 3(4), 176-188. <https://doi.org/10.63282/3050-922X.IJERET-V3I4P119>