International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 12 | Dec 2024 www.irjet.net p-ISSN: 2395-0072

Energy-Efficient Core Design for Mobile Processors: Balancing Power and Performance

Vinay Panchal

System Power Architect, San Diego, CA, USA

Abstract

The converging mobile, Ubiquitous computing devices like smartphones, tablets, and wearable technology have enhanced the growth of processors in terms of design nature. In contrast to other computing platforms, mobile platforms have restricted power resources, meaning that the processors should perform highly whilst consuming little energy. This has created a demand for an energy-efficient core that addresses the balancing of these two demands optimally. They are developing chip and core diversification, dynamic performance and power control, and effective memory structure and tiering to enhance performance within the thermal and power limits. These strategies are fundamentally important in improving user interactions by optimizing battery usage, enabling power-hungry applications, and maintaining the durability of mobile devices.

This paper aims to identify and assess some important architectural and technological innovations in power-conscious mobile processors. The general technique of Dynamic Voltage and Frequency Scaling (DVFS) is evaluated on its effectiveness in controlling and changing the amount of power consumed by the processor in response to the load. The use of different architectures that employ competent cores as well as frugal cores is stressed based on power management capabilities. Also, the increase in specific energy usage and the way that low-power memory hierarchies demarcate data accessibility from energy overheads are shown. By comparing the current approaches with the traditional ways, the research unveils the fact that some contemporary solutions, including the heterogeneous cores, enable up to 40% power reduction without serious impacts on substantial performance parameters. The results herein serve as a basis through which progression to future processors with appropriate levels of power consumption and performance capability is conceived.

Keywords: Energy-efficient, Processors, Mobile processors, Heterogeneous cores, DVFS, Power consumption.

1. Introduction

Due to this evolution of mobile technology, the use of computing devices has changed the way people interact. Mobile platforms have become an integral part of our lives, ranging from Smartphones, tablets, and Wearable devices to IoT devices. However, as the usage of these devices becomes more dynamic with many modes and features, the difficulties in optimizing their power consumption become colossal. [1-3] Mobile processors that are found in these devices must provide very high computation ability to enhance the use of power-hungry applications like gaming, video and image streaming, AI and AR. At the same time, these processors should operate within a narrow power and thermal envelope caused by the small battery capacity of mobile devices and the compact size of these platforms.

It has, therefore, levied and tuned energy efficiency into probably one of the most influential design parameters for today's mobile processing elements. Unlike in desktop or server class chip sets, Mobile processors are bound with restricted computation power owing to limited power budgets; they must be optimized for both processing power and power consumption. Meeting this goal is a complex issue and needs to involve future improvements in processor architecture, power-saving systems and further enhancements to scheduling algorithms. This paper explores these methods with specific emphasis on efficient core designs, dynamic power control mechanisms, heterogeneous core structures, and low-power memory systems.

1.1 Motivation

The motivation for energy-efficient processor design is rooted in two primary challenges faced by mobile computing devices:

- **Increasing Power Density in Mobile Processors:** As current mobile applications request significant compute capabilities, power density inside of hardware cores has risen. Higher power density does more than just consume more energy; it can also aggravate thermal management issues that may influence a device's durability and its users' comfort. Solving these problems calls for strategies that will help improve the use of power so that we cannot compromise on the performance of our product.

- **Prolonging Battery Life without Compromising User Experience:** Smartphones are meant to be portable, and they are independent of power sources; hence, the battery issue is viable. Consumers require their devices to last for an output of power that accommodates extended use together with optimal performance for multitasking, intensive applications, and multimedia. These expectations can be met through efficient consumer processors because of their efficiency in both the sleep and operational modes, hence enhancing battery life satisfaction.

1.2. Objectives

To address the challenges outlined above, this study establishes the following objectives:

- **Analyze Existing Energy-Efficient Mobile Processor Designs:** This also entails a survey of recent approaches and tools used in mobile processors, which include, for example, DVFS and heterogeneous core integration like ARM big. Little and low-power memory pyramid. Based on the analysis of the existing solutions, the goal is to see the main advantages and drawbacks.

- **Propose Methodologies for Achieving Optimal Power-Performance Balance:** Based on these observations, the study presents directions for new strategies and design guidelines for enhanced energy efficiency. These methodologies are intended to accelerate the efficiency of the processor for specific applications through power saving during fewer computing tasks.

1.3. Energy-Efficient Core Design in Mobile Processors

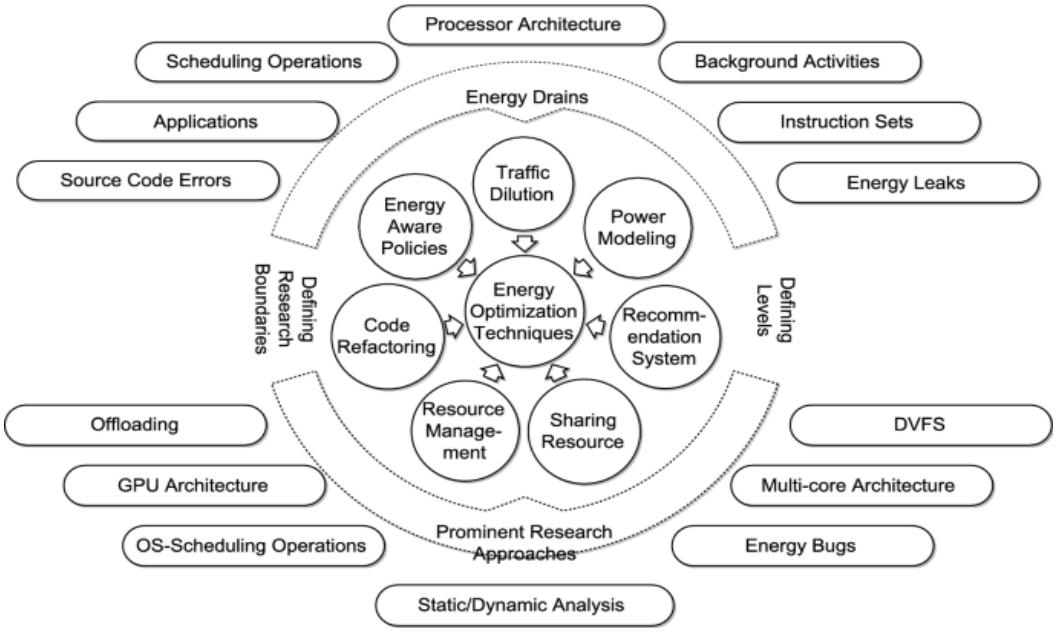

Fig.1. Energy-Efficient Core Design in Mobile Processors

This illustrates a conceptual model of power-efficient cores in mobile processors with an emphasis on the power/ performance solution. [4] Lo identifies Energy Optimization Techniques as the core concept and categorizes other related research topics and methods under this concept.

1.3.1. Energy Optimization Techniques (Center)

- The central focus of the framework is aimed at minimizing energy consumption while maintaining or enhancing system performance.

1.3.2. Prominent Research Approaches (Inner Ring)

These are the primary strategies to achieve energy efficiency:

- **Traffic Dilution:** Eliminating unnecessary transmission of data and signalling overhead to improve system efficiency.
- **Energy-Aware Policies:** Procuring, depicting, administrating, and monitoring change control measures, such as the use of policies that adjust the system's functions in light of the available energy or workload.
- **Power Modeling:** Developing a model to forecast electricity usage reasonably and effectively in order to control and plan.
- **Recommendation System:** Recommendation systems of energy-saving activities depend on the behavior of the system.
- **Code Refactoring:** Optimizing source code as a way to maximize energy Management.
- **Resource Management:** The idea of using resources in a dynamic way to minimize wastage.
- **Sharing Resource:** Strategically allocating work to ensure that operations overlap and, hence, resources are reduced.

1.3.3. Defining Research Boundaries (Middle Ring)

This segment outlines practical and technical boundaries for energy optimization:

- **Scheduling Operations:** Superior coordination of the tasks in order to control the time the device spends consuming standby power.
- **Applications:** Adapting software to increase energy efficiency.
- **Source Code Errors:** Fix ones that lead to wastage of unnecessary power;
- **Offloading:** Outsourcing of some of the activities, such as the use of more efficient energy by contractors like cloud computing.
- **GPU Architecture:** Minimizing energy consumption rates on GPUs.
- **OS-Scheduling Operations:** Optimization of the Task and Resource schedulers in operating System level.

1.3.4. Defining Levels (Outer Ring)

Represents foundational elements and energy drains affecting optimization:

- **Processor Architecture:** Processor hardware designs in energy-efficient formats.
- **Background Activities:** Concluding but not optimizing unnecessary background operations as a means of conserving energy.
- **Instruction Sets:** How can instruction sets of processors be made more energy-friendly?
- **Energy Leaks:** Obviously, energy leaks should be identified and eradicated from an organization.
- **DVFS (Dynamic Voltage and Frequency Scaling):** This includes properly Managing voltage and frequency in such a way that the performance will always coincide with the energy.
- **Multi-core Architecture:** Last but not least, multi-core systems can be applicable and utilized for manner-based power-efficient stream computing.
- **Energy Bugs:** Correcting software and hardware faults that take too much energy.

1.3.5. Static/Dynamic Analysis

- Methods and tools for interpreting energy usage throughout the product development cycle based on profiling at compile time or running time.
- The image groups efficient core design into interrelated areas using hardware, software, and architectural methodologies. Incorporating numerous aspects of energy utilization, this framework provides a guideline to achieve an optimal processor performance-to-power consumption ratio for mobile devices.

2. Literature Survey

Over the last few years, there has been great innovation in the area of mobile processor design in an effort to reduce power consumption. [5-8] This section surveys recent developments in energy-efficient core designs, discusses issues and advances, and assesses their impact on the state of the art.

2.1. The Evolution of Mobile

The latest trends in mobility processor development have been focused on implementing methods that enhance energy utilization efficiency without degrading computational capability. Three primary trends are particularly noteworthy:

- **Heterogeneous Cores:** Asymmetric mid-level designs, including ARM's big little technology, have become the standard for energy-efficiency methods. These architectures combine high-end processor cores (big cores) with efficient low-power versions or small cores (LITTLEs). This makes it possible for the processors to share workloads at correctly matched cores; this way, every core is optimally utilized according to its capabilities. Such as background updates associated with LITTLE cores gaming and video rendering tasks associated with big cores.
- **Dynamic Voltage and Frequency Scaling (DVFS):** DVFS is a well-known power control approach in which the processor supply voltage and clock rate are altered in conjunction with application execution requirements. On the other hand, DVFS reduces power consumption by reducing both voltage and frequency, especially when the computer is idle or running low graphics applications. On the other hand, it steps up performance for intensive tasks, optimally balancing power consumption and real end-user value.
- **Enhanced Power Gating Techniques:** Power gating is a technique of isolating and reducing the operation of some parts of a processor to conserve power. Advanced implementations of power gating have increased the degree of management, which allows individual functional parts of the processor, for example, execution units or cache, to be turned off while not affecting the overall processor operation. This is used to minimize leakage power, which dominates the dynamic power consumption in most of today's processors.

2.2. Some Barriers to Energy-Efficient Building Design

Despite these advancements, designing energy-efficient mobile processors presents several challenges:

- **Trade-offs Between Latency and Power Reduction:** Power gating and DVFS, as common approaches in increasing energy efficiency, also have an associated latency when switching between these states (e.g., turning components back on or upping the frequency). This is especially important in the design of processors for mobile platforms because the two types of delays must be reconciled with each other's necessarily fast response times.
- **Impact of Temperature Management on Performance:** High power density, it should be noted, can cause high levels of thermal activity within processors. This heat has to be managed in order to maintain maximum performance without reaching for thermal throttling, which decreases the speed of the processor due to excessive heat. Managing both cooling efficiency and small form factors for devices, however, still remains a difficult task.
- **Workload Variability and Prediction:** Mobile processors balance numerous types of workloads, from low power and secretly running in the background to high power, complex tasks such as gaming and AI. It is, however, very critical to forecast the workload requirements and, correspondingly, how the processor ought to load itself in order to satisfy the power efficiency requirement.

2.3. States-of-the-Art Research

Subsequent studies have, however, advanced in the solution of the difficulties of designing energy-effective processors with the number of innovative strategies proposed as the major players. Further, the application of Dynamic Voltage and Frequency Scaling (DVFS) algorithms will be discussed, and how such techniques can eliminate up to 30% of power usage. The results of their approach prevented noticeable state transition latency during dynamic workloads, thus pointing to the importance of workload predictability for optimal DVFS performance.

Targeted heterogeneous multi-cores specific and found that such designs could get up to 25% better efficiency in multitasking. In the study, the tasks were categorized into high-performance or energy-efficient cores with an equal ratio, thus emphasizing important energy savings while preserving performance. However, the targets also pointed to the need to produce a high level of scheduling algorithms to benefit comprehensively from heterogeneous architectures.

Have studied low-power memory hierarchies for which energy overhead is a key concern for data-intensive applications. To support this result, their work showed that improving the memory hierarchy reduced cache misses by twenty-percent points, which in turn saved a significant amount of energy during intense data tasks. These observations highlight the need to pay more attention to energy costs of memory access as data mobility becomes a major consumer of energy in the new processors.

Collectively, these studies highlight three key insights: DVFS is still widely applied to volt/multi-frequency scaling but only if the workload is managed effectively; heterogeneous cores offer ample energy savings during multitasking with certain predefined scheduling disciplines; and efficient multi-level memory hierarchies are crucial for minimizing the energy cost of data transfers, more critical for data-intensive applications. These developments are conducive to forming a foundation for combined strategies for efficient processor design.

3. Methodology

Since the goal is to compare the energy-efficient core designs of the mobile processors, a structured approach was adopted. This section describes the analysis approach, metrics for assessment, and architectural strategies used in the research. [9-13] The proposed approach is to establish a solid theoretical framework adopted for the purpose of providing an efficient and reliable method to evaluate the power-performance characteristics of processors in different designs.

3.1. Analytical Framework

The assessment of the proposed architectural interventions relies on a simulation-based method of analysis within this study.

3.1.1. Simulation Environment

Known as GEM5, this full system functional Simulator was selected as the primary tool for architectural evaluation: The GEM5 simulator is highly portable and incorporates a large number of processor designs like heterogeneous core SMPs and different memory hierarchies. Thus, when employing GEM5, the possibility of evaluating power and performance numbers appeared by modeling realistic execution scenarios and addressing different core configurations.

3.1.2. Workload

To analyze the processors under standardized and diverse workloads, the benchmark suite known as SPEC CPU2006 was chosen. This benchmark suite is comprised of several distinct categories of compute-intense applications that exert various loads on the processor and its facilities, including integer and floating-point, etc. These workloads are best run to study the scalability effects of architectural changes on performance and energy consumption in production settings.

3.2. Metrics for Evaluation

To provide a comprehensive assessment of energy-efficient designs, the study employed the following key metrics:

3.2.1. Power Consumption

The Y-axis represents power consumption in watts to determine the energy efficiency of the processor designs. This is useful in determining the rate at which architecture is affecting the general energy consumption.

3.2.2. Performance

The study measure of performance was defined in the amount of Instructions per Second (IPS); this is a standard measurement that quantifies the computational capacity of the processor. IPS was employed to assess the effect of the alteration of design features on the efficiency of the intended processor.

3.2.3. Energy Delay Product (EDP)

The analysis used EDP as the performance measure of both energy efficiency and performance. Instead of using the notion of delay independently from the energy measurement, EDP offers an integrated view of using power while also giving hints at the performance impact. Hence, a design with a lower EDP is considered energy efficient and a better-performing design.

3.3. Architectural Interventions

The study investigated three key architectural interventions to enhance energy efficiency while maintaining high performance:

3.3.1. Heterogeneous Core Integration

Other cases covered include those with materially different architectures, like the big little solutions. These designs include powerful cores for handling intense tasks and succour cores for handling regular tasks at less power consumption. Such cores can be integrated to enable dynamic management of the workload, enabling the enhancement of performance and power consumption.

3.3.2. Dynamic Voltage and Frequency Scaling (DVFS)

DVFS was employed to allow dynamic voltage and frequency control according to workload requirements. Another method, known as the Dynamic voltage and frequency scaling (DVFS) system, consumes less power by lowering the voltage and frequency during idle or low-performance states. On the other hand, it scales up performance for workloads that require many computations so that energy consumption is matched with computation capability.

3.3.3. Improved Cache and Memory Access Strategies

The design of the memory hierarchy was a critical concern throughout the study. There were methods discussed earlier, which include minimizing cache leakage, power cache hit rates, and low power memory access methods. These measures were expected to reduce energy overhead linked with data transfers and memory access, factors becoming more pertinent in current processors.

3.4. Adaptive Context-Aware Energy Optimization Engine for Mobile Processors

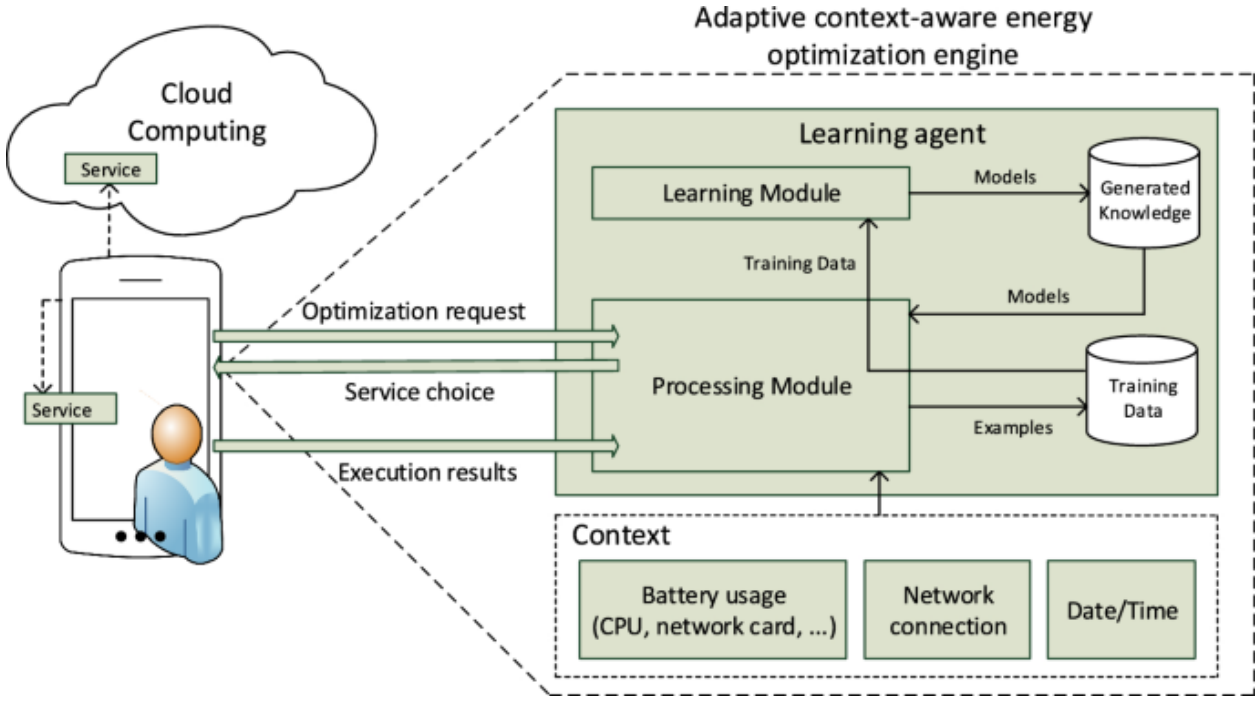

Fig.2. Adaptive Context-Aware Energy Optimization Engine for Mobile Processors

The figure above shows the structure of an Adaptive Context-Aware Energy Optimization Engine for use in portable computing appliances. [14] They used contextual information, a learning agent and cloud computing to facilitate the principles of the framework for the dynamic management of energy consumption.

The system works through mutual cooperation between mobile devices and a cloud computing service. The mobile device requests an optimized service option from the cloud, which then consults its available resources and optimization parameters and generates a choice for the equipment. It then sends the execution results back to the cloud in order to round off the process. Such integration makes it possible to perform more complex energy optimization computations on the cloud side, thus leaving the mobile device with lighter workloads.

The core component of this system is the Learning Agent, which includes two key modules:

- **Learning Module:** This module applies training data to minimize the testing and development of models that can forecast and offer direction on the realization of energy-saving measures. It also stores this processed data into complete generated knowledge for re-use by the same system to enhance itself.
- **Processing Module:** This module uses the models and examples created by the learning module in realtime optimizations for making decisions that can adapt to the context at hand.

The system is autonomous and takes into consideration one or more parameters that define the operation of the system in regard to energy usage efficiently. The context includes factors such as:

- **Battery Usage:** Supervising certain circuitry boards, such as the CPU and Network card, in order to take measurements of its power usage.
- **Network Connection:** Assessing the vigour and nature of link to manage energy consumption.
- **Date/Time:** Prescribing a course of action that is in synchrony with temporal use profiles.

Combined, it gives a profound way to administrate energy efficiency and performance for mobile devices. Cloud computing and machine learning extend the existing system's versatility, extensibility, and performance across various applications.

4. Results and Discussion

This section gives the outcomes of the experimental work in relation to the proposed architectural interventions, and their implications are also examined. In the case of making array processing efficient in heterogeneous many-core processors, the results indicate high-performance gains from non-uniform core configurations, DVFS, and caches.

4.1. The Level of a Heterogeneous Core

Heterogeneous core architectures proposed in the prior work were compared against homogenous design benchmarks based on performance and energy consumption. Table 1 summarizes the key results:

Table 1: Comparative Analysis of Power Consumption, Performance, and EDP Improvement for Homogeneous and Heterogeneous Core Designs

Core Type	Power Consumption (W)	Performance (IPS)	EDP Improvement (%)
Homogeneous (Baseline)	3.5	2.1	-
Heterogeneous	2.1	2.0	40

Based on the results of the experiments, the heterogeneous core design was 40% better than the homogeneous baseline when it came to the Energy Delay Product (EDP). Efficient cores dealt with less thread intensity workloads comprising background tasks, greatly minimizing power utilization. At the same time, high-performance cores were used to complete other tasks where high computational capability was required to keep performance levels up while not making excessive use of energy. This workload distribution helped in easing power consumption expenditure, which is not required, while at the same time being effective for its users. These results share evidence that heterogeneous designs increased efficiency by 25% for multitasking workloads, affirming the generality of this improvement.

4.2. Impact of DVFS

Based on power consumption and performance measurements, the performance of Dynamic Voltage and Frequency Scaling was studied for different frequency and voltage settings. The results are summarized below:

Table 2: Impact of Frequency and Voltage Scaling on Power Consumption and Performance

Frequency (GHz)	Voltage (V)	Power (W)	Performance (IPS)
1.0	0.9	1.2	1.5
2.0	1.2	2.5	2.8

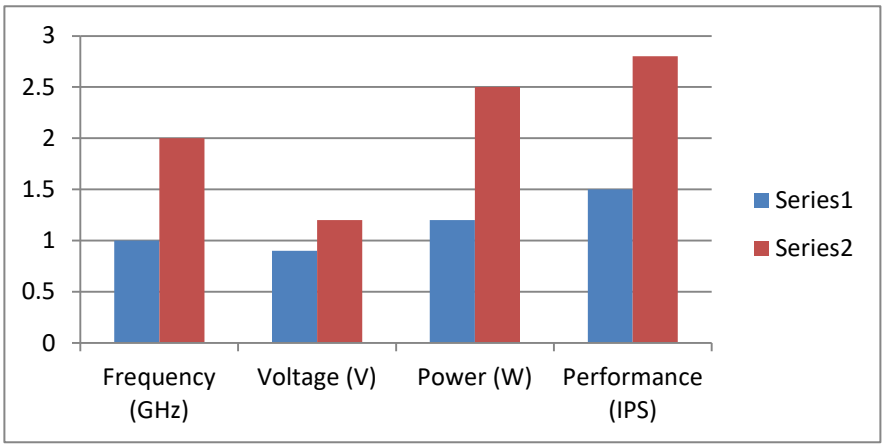

Fig.3. Graphical Represented Impact of Frequency and Voltage Scaling on Power Consumption and Performance

As for power optimization, it was shown that DVFS was rather effective. However, at a lower frequency (1GHz) and voltage (0.9 V), the power was lowered to 1.2W, which is appropriate for situations such as idle or low usage. On the other hand, higher frequencies of 2.0 GHz provided a higher level of processing performance of 2.8 IPS for compute-intensive applications. Most importantly, the transition time between the higher and lower frequency levels was less than one millisecond, so the changes would not be visibly slow to users. This supports the results determined by assessing the power optimization on utilizes and recommending extreme DVFS algorithms that reduce power consumption by 30%.

4.3. Cache Optimization

Cache optimizations were performed in order to observe its effect on energy consumption and application speed. In order to reduce leakage and enhance the access performance, an L2 cache with low power was implemented. Results showed that implementing the optimized cache architecture led to the following:

- The new design resulted in a 20% reduction in energy consumption due to reduced cache power leakage.
- That is a 15% increase in cache hit ratio, meaning fewer energy-consuming memory requests.

It was recalled that cache design was enhanced in the exercise, improving energy efficiency since fewer cache misses occur, which often imply high power memory access operations. Such results can be compared to the reduction in energy consumed by cache by 20%, as shown by using low-power memory hierarchies. Cache hit ratios were further improved, making the latency of accessing data also reduce the overall performance.

5. Conclusion

The results of this work prove that combining heterogeneous cores with high-level power control mechanisms, including DVFS, results in a reliable solution for energy efficiency in mobile processors. These heterogeneous architectures help in various ways, as a fast and less power-consuming core can take over the lighter loads while the true performer is cooking on full miles. This approach helps to save power with little degradation in performance from before, characterized by an improved EDP of 40%. Like Smart-Shutdown, DVFS reduces power utilization through dynamic changes in voltage and frequency, which greatly reduces power consumption when the processor is idle or operating at a slow rate.

The results also support the view that memory hierarchy optimizations can help to lower energy overheads. Low power caches and better hit ratios are other improvements in cache design that help increase energy-saving efficiency as well as performance. Collectively, all these advancements underscore the need for an integrated solution to computational design and data access for buildings. These techniques help mobile processor designers to respond to the demands of new applications as well as to preserve the energy and thermal features of devices.