

Cortex-M33-Based Soc Design With Optimized Dynamic Protocol Adaptation For Data Communication

Gundu Ramachandra kumar¹, Budati Anil Kumar², Raenu A/L Kolandaisamy³, Srinivas Bachu⁴, Y Dasaratha Rami Reddy⁵

¹Department of ECE, Koneru Lakshmaiah Education Foundation, Hyderabad, India & Geethanjali College of Engineering and Technology, Hyderabad, India. ramachandrag1418@gmail.com

²Department of ECE, Koneru Lakshmaiah Education Foundation, Hyderabad, India & Adjunct Professor, ICSDI, UCSI University, Kuala Lumpur, Malaysia. anilbudati@gmail.com

³ ICSDI, UCSI University, kuala lumpur, Malaysia. E-mail: raenu@ucsiuniversity.edu.my

⁴Department of ECE, Siddhartha Institute of Technology & Sciences, Hyderabad, India. E-mail: bachusrinivas@gmail.com

⁵Department of CSE, Chaitanya Bharathi Institute of Technology, Proddatur, A.P., India. E-mail: dasradh@gmail.com

Abstract

This paper investigates a dynamic protocol adaptation framework into a System-on-Chip (SoC) framework to enhance power consumption and performance, which contains such elements as processors and communication system (AXI, AHB, APB) inherent in data moving. It is designed by a Cortex-M33 processor and Real-Time Monitoring System to assess the workload and performance measures to determine an ideal selection of communication protocols. The Protocol Adaptation Mechanism dynamically adapts protocol parameters whereas Protocol Configuration Manager helps network movement between different protocols. The inclusion of AXI-to-APB, and AHB-to-APB bridges allow the transfer of data to work efficiently. The SRAM module increases the depth of the performance of memory to 1500-location. The results indicate significant improvement: in the case of the 32-bit processor, where the slice register usage is reduced to 3491 as compared to 4201, the delay is reduced to 4.35 ns as opposed to 7 ns, and the use of power decreases to 10.96 mW according to 18.34 mW. In the case of the 64-bit processor, we will get a reduction in using slice registers (4913 to 3891), improvement in delay (11.45 ns to 7.35 ns) and a decrease in power consumption (17.34 mW to 10.42 mW). The space is also reduced and the throughput is boosted greatly. This framework is verified by the results that were gained in Vivado Design Suite 2018.1 of Zynq 7000 board after full testing, showing greater versatility and effectiveness across different programs.

Keywords: SRAM module SoC, Cortex-M33, AHB-to-APB Bridge, AXI-to-APB Bridge, Real-Time Monitoring System.

1. INTRODUCTION

An affliction is the energy efficiency and performance optimization that can be regarded as a crucial concern in the current paradigm of the embedded systems and System-on-Chip (SoC) architectures. Modern SoCs include numerous components, such as processors, memories, buses, and numerous communication standards like AXI, AHB, and APB, without which the smooth functioning and data transfer in the system will not be possible. With the rising complexity of the SoCs, there is a greater necessity to manage such parts more effectively, particularly when giving importance to the energy requirement, along with the performance. Dynamic protocol adaptation framework is one of the best improvements in this field, as it provides the trade-off between the energy efficiency and performance. The proposed project proposes a new framework to integrate dynamic protocol adaptation in SoC designs with the aim to optimize these two aspects respectively by using the mechanisms of real-time monitoring and configurational adaptability. The main idea in this framework is the intelligent selection and transition among the communication protocols depending on the current requirement of the system hence the efficient transfer of data and usage of resources. This module is meant to support many communication protocols such as AXI-to-APB and AHB-to-APB bridges that ensure data transfer between various subsystems in the SoC. The Real-Time Monitoring System works at monitoring constantly workloads and performance, thus, it provides crucially significant information to the Protocol Adaptation Mechanism. This dynamic mechanism then dynamically configures the protocol operations that keep in line with the perceived conditions of the system. Protocol Configuration Manager helps to maintain

smooth and efficient transition between protocols avoiding interruption and maximizing the overall performance of the system.

Besides the adaptation of protocols, the design urges the use of high-performance SRAM module located at 1500 locations and this is one of the most vital aspects of optimizing data read and write. This module helps data to be efficiently used, which increases the performance of the SoC even more. Random signal generation and thorough test bench properly tests the read and write functions confirming the correctness of the data interfaces. Another checker module is also utilized to present any inconsistency between the slave outputs and the processor inputs, introducing another level of validation. The presented framework is synthesized using Vivado Design Suite 2018.1 and implemented on the Zynq 7000 development board, which proves its applicability in real-life applications. The test bench using deep learning is able to reach the coverage of 100%, with which the design can be cover-verified. The methodology is not only a tool to balance the energy usage and performance, but is also a creative solution to harness modern embedded systems with a view to its flexibility and efficiency problems, which is increasing with SoC with its high adaptability requirements and energy demands.

2. LITERATURE SURVEY

One new system simulator is created to optimize design parameters and does so in order to decrease the rate of collisions. This simulator will have diverse adjustable properties among them being software hardware partitioning operational scheduling and merging memory. A good method of collision detection has to consider only the close objects and avoid unwanted computations. Next, experiment design is a procedure, which consists of systematic development with the view of collecting the best data as per the objectives of an experiment. Orthogonal Matching Pursuit (OMP) algorithm is an algorithm to solve NP-hard problems, connected to sparsity. The simulator was used to predict performance of four other scenarios of the OMP algorithm and this performance was compared to the actual performance results realized using the Zed Board[1] which is the Internet of Things (IoT) requires end nodes that are capable of ultra-low-power, always-on operating modes so that long battery life can be achieved without compromising performance. In the current paper, the potential of mixed analog/digital computing in current designs of deep neural network (DNN) processors is talked about. We offer our all-digital heterogeneous system-on-chip (SoC) an AI-enabled IoT end node realized with 22nm FDX by GlobalFoundries: Marsellus. The open-source, modified, near-threshold (NT) optimized RISC-V processor core available in Marsellus is called RISC-V-NT. During a computation performed, a Fused Multiply-Add (FMA) or Fused Multiply-Add accumulate (FMAC) employs one rounding. Reconfigurable Binary Engine (RBE) is a DNN accelerator, whereby the performance is enhanced by the additional use of Hardware Processing Engine (HWPE). The offered approach caps on internet-based performance checks by means of Process Monitoring Blocks (PMBs), allowing real-time readjustments of transistor start voltages in light of changing circumstances. In DIANA, the inherent trade-offs between power and performance are exploited by mixing the two kinds of cores and combining them in a hybrid SoC tailored to end-to-end neural networks processing and optimizing the shared memory stack. Neural networks enable computers to identify patterns and address advanced problems in such areas as artificial intelligence, machine learning, and deep learning. Even though ALMC cores have potential to provide a gigantic computational parallelism and efficiency, they have come at a price of flexibility and exactness in data flow[3]. The paper presents gem-MARVEL, The 1st unified microarchitecture level fault injection. This dedicated platform is compatible with all the popular Instruction Set architecture (ISAs) CPUs and most domain-specific devices. Gem5-MARVEL relies on the modular design, in which it allows variety of fault injection scenarios depending on the fault model and various system conditions it contains a set of libraries that automatically inject faults and also evaluate the impact of hardware faults under full system operation. Our feasibility of the framework is demonstrated by analyzing the framework on a number of 64-bit CPU ISAs, such as, x86 ARM and RISC-V and a series of domain specific accelerator designs. Artificial intelligence (AI) and extended reality (XR) are similar and dissimilar in their origin and principal aims. Such applications are extremely latency sensitive (typical end to end delay of only 10-20ms) and are power constrained, being required usually to run at only tens of milliwatts average. This amalgamation improves the performance, saves power and optimises semiconductor die volume over the conventional motherboard founded architectures. Access energy of non-volatile memory (NVM) to read the weights of network represents a major factor, which increases overall weight stage power consumption in complex DNNs [5]. The article entitled Dynamic SoC Balance Strategy for Modular Energy Storage

Systems proposes the integrated State of Charge (SoC)-based droop control technique. The offered control methodology improves the effectiveness of SoC balancing and removes existing deviations. The LBC line provides average SoC value and the SoC is the basis of the droop coefficient that are controlled at the local control layer. Also, a lightweight and a high-performance automated detector of abnormal activity is created to meet unjustified usage of on-chip SoC resources. These methods are illustrated using case study of a real SoC in connected autonomous vehicles (CAVs) in the highly unpredictable conditions. security measures would fail to address this aspect in a heterogeneous SoC. It might be intriguing to discuss in more details the details of how security is integrated with existing SoC designs, what sort of anomalies it is capable of detecting, and how operations related to false positives are carried out. Moreover, the ability to get an overview of computational overhead and scalability of the solution may be important regarding its implementation into large-scale, diverse SoC settings [8]. The possibility of NV-LDPS coding the omission of space telecommand connection utilisation by a RISE-V softvert-core processor together with a vector co-processor. Altogether, your proposal is one of the feasible solutions to the space telecommand link applications based on the reconfigurable hardware usage in order to complete the necessary tasks and remain flexible and less costly. It will depend on how much effort is dedicated to performance of the content and how much effort is dedicated to versatility in order to be successfully implemented [9]. There is strong relevancy and timeliness in droop control based on IoT of energy storage systems (ESUs) in and between microgrids. It may also be advantageously applied to explain how your IoT-based droop control manages possible issues of delays in communications, data security, and compatibility with the existing system. Also, it might be beneficial to present some performance values and compare with the traditional processes, so as to demonstrate the benefits of what you are doing [10]. Your activity related to the creation of an SoC-based solution aimed at signal processing in impulse radio ultrawideband (UWB) is rather topical and innovative. Microgrid also assists in elevating the energy marker by forming an environment of restricted energy production and transportation. These are especially useful in the field because of their capability to offer high-resolution information in the time domain.

It is timely and sophisticated that in [12] proposal to co-design of MPSoC (Multiprocessor System on Chip) architecture using ant colony optimization (ACO), some privacy and security issues were addressed. An ant colony optimization (ACO) is a meta-heuristic algorithm based on the behaviour of foraging ants. It is applied to answer problems given in the form of locating optimum paths through graphs. In it, a nest of fake ants coordinate their efforts to search various directions and build upon solutions over time, resembling the process of synchronization and discovery of effective tracks, exhibited by genuine ants. The definition of data security is often recognized as a specific set of protection toward the significance of such data not being accessed and stolen by non-authorized parties. The modularity of MPSoC with the use of system-on-chip (SoC) technology, any number of the subsystems (or all of them) can be integrated in a single device. The optic of co-designing MPSoC architectures that you use by employing the ant colony optimization techniques is fair as it covers both performance and major issues relating to privacy and security issues when handling sensitive information. The Ultrascale MPSoC design has the type of scalable process 32 to 64 bits and has the support of virtualization and amalgam of soft and hard engines. Highly real-time road segmentation to support autonomous driving in the virtual reality is improved through Road Net-RT architecture. Although CNNs are efficient in processing visual data, they are also becoming more complex, which negatively affects real-time performance and hence the autonomous driving app presents a such challenge.

In DIII-D tokamak where a number of large electronic biologicals have been found in one SoC-based instrument, space and complexities involved are improved. Automatization using SoC reduces human participation, the diagnostics is more accurate. The analog driver generates non linear sweep of 0-20V, and Data Acquisition System (DAQ) runs in-phase (I) and quadrature (Q) components which guarantees that the data acquisition is high-speed and high-resolution. The system design should be attended to make it well integrated and performing. The validation and rigorous testing along with performance evaluation on the integrated instrument is required to prove that all the needs of performance and accuracy have been met. The ARDI provides the potential of modernization and simplification of the process of the reflectometry diagnosis in the lab of the tokamak. A combination of several elements in one instrument that is based on SoC would respond to several challenges concerning the space, manual handling of the instrument, and the complexity of a system by improving the functionalities, the accuracy of the system, and the ability to configure them remotely [14]. Detecting, locating, and tracking of the static, and solving

these key challenges on the topic of concerns in the tokamak lab environment is represented in a radar in [15] which is a type of radar called a fast chirp frequency-modulated continuous-wave (FMCW). Space requirements The rigidity and bulkiness of the current setup requires much lab space which (is) in short supply in tokamak labs. FMCW reflectometry uses single continuous wave with its frequency modulated with time. Manual intervention to alter control parameters, which is required with Automation and Remote Configurability, may create inconsistencies and wastages. Consolidation of Components in corporation All components required to set up the system can be incorporated into a single and compact unit that will ease the complexity posed by individual and bulkier machines to handle. Make the parts of the compact FMCW reflectometry instrument compatible and best suited to perform. The compact instrument should also undergo testing so as to ascertain that it has realised the required FMCW reflectometry performance specification. Write an intuitive user interface which allows remote configuration and monitoring. With automation and remote configurability incorporated into a simplified design, your solution increases efficiency, accuracy and flexibility, which would make it a highly useful tool to perform modern tokamak experiments and diagnostics [15]. The Intelligent Reflecting Surface (IRS) technology has the potential of delivering massive impact on next-generation communication systems with the potential to enhance signal integrity and network performance. But, as you mentioned, an IRS integration does bring further issues, most notably in phase-shift optimization, and this issue may affect the overall latency of the system. The multiple small elements that comprise the IRS are reconfigurable, passive, low cost reflectors that represent a metasurface building block. IRS can be applied especially in the situation where there is low signal coverage, it can be interfered, and direct communication by line of sight is not evident. An Intelligent Reflecting Surface (IRS) in reflection mode reflects signals between the access point and the client causing phase shifts that affects the performance of the system. One of the major problems is wherein accurate optimization of the phase shifts must be achieved; however, at a reasonable time to calculate the optimal values of the phase shifts. IRS is a new technology of hardware that increases the coverage of signals and also saves on energy at a minimal cost of deployment. These obstacles can be suppressed with the application of efficient algorithms, real-time processing, adaptive hardware-accelerated strategies, and a fantastic opportunity to exploit the benefits of the IRS technology. In [17], increased energy and area efficiency of edge-based machine learning (ML) systems have various novelty elements and deliver outstanding performance figures. non-volatile weight storage using 2-MB of so-called Magneto resistive Random Access Memory (MRAM). This implementation makes use of MRAM to store weights, a feature that is associated with non-volatile memory where data is not lost incase there is no access to power. CNN Loop ordering Gerdr AB optimization of the order of operations, in the CNN loop, will also help to minimize the power to be used in the memory access and computation, which will end up saving power. Decreases power usage with the integration of MRAM, memory optimization with IAMEM, and a better neutron loop ordering, your design also consumes much less power, which is of importance to energy-limited edge devices. Complete their running tests to confirm the performance parameter and make sure all conditions are met by the design in different operation situations. Your proposal on the edge ML using the SoC system with MRAM-backed non-volatile storage, IAMEM Buffer optimization, and CNN loop scheduling represent an incredible progress in energy- and area-efficient design. Your approach to meeting the demands of next-generation edge devices is effective based on the performance metrics, such as enhanced efficiency of Harris corner detection as well as the CNN task. This design portrays Edge devices, the data is created here and forms a major basis of more innovation and development [17].

The need to develop a low-caste IoT SoC embodied in an open standard instruction set architecture (ISA), [18] based on known reductions in instruction set computer (RISE) addresses the following needs, which have an increased priority in information technology industry. The progress of information technology is IoT, which is a giant step toward connecting the world of digital technology and the world of ordinary things to make life more productive and equal. IoT has already become a potent source of enhancing operational efficiency, decision making processes, overall productivity, and data management processes of SoCs that are capable of performing different tasks and maintaining a low cost in these processes. RISC the common ISAs allow designers to start off with the same basics ISA and customize their device to meet the demands of applications of embedded design which helps in flavouring low-cost and vendor-efficient SoCs. The versatile use allows carrying out of image acquisition and barcode recognition extending the list of applicants of the IoT devices and making them more diverse and perform intricate operations. Real-time processing means that SoC is capable of processing in real time, especially in scanning barcodes to

satisfy the need of different applications in terms of real-time performance. Feature expansion looks into future enhancements or installing of additional features on the SoC such that it might give it a greater scope of capabilities or better performance. The combination of open-source architecture with cost-effective design support guarantees meeting the critical demands of the IoT market and enhancing the studying of SoC development. The suggested SoC chip has a high possibility of increasing the capabilities and the prices of IoT devices making their use diffusing and constructive in the field [18].

Dilithium selection algorithms: Algorithms announced today are the first completed standard published by NIST pertaining to post-quantum cryptography (PQC) standardization themes with regards to positively demonstrating its security and promising it as a quantum-resistant solution. FPGAs support parallel processing, and this allows the implementation of cryptographic operations to be performed at an increased speed as many tasks are run at the same time. Latency refers to the amount of time needed to accomplish one cryptographic procedure. Cross-platform testing that goes as far as the platform and compares the outcomes can give an in-depth overview of how well dilithium works in different environments. Efficient mapping of lattice-based dilithium cryptographic algorithm on an FPGA SoC platform is important progress in the assessment of the applicability and efficiency of the system in the context of post-quantum cryptography. Using the flexibility and multiprocessing power of the FPGA technology, your contribution makes an important contribution in offering valuable dilithium a lattice-based digital signing scheme to protect data against quantum computing attacks. This is towards the wide objective of pursuing post-quantum cryptographic standards and maintaining strong security against emerging quantum attacks [19].

Automatic clock gating (ACG) is an elaborate method of cutting down on dynamic power dissipation in clock distribution systems by adding a control in the form of automation to a clock gating mechanism. Classical use of clock gating manually or statically disables the clock to idle blocks in order to conserve power. In this respect, ACG improves this by abstracting the graph data structure as a union of nodes interconnected by edges. With the increase in complexity of the digital design and power limitation being more stringent, automated methods such as ACG will be used in power management techniques. Further improvement in the control mechanisms as well as integration methods will subsequently improve the effectiveness of ACG.

Automatic Clock Gating (ACG) is another dramatic breakthrough in clock-gating process and modelling of the global clock distribution network as graph. The dynamic and efficient power management introduced here due to adding control mechanism to the arcs on the graph results in a decrease in dynamic power dissipation and increase in system performance. The application of ACG is a process that requires thorough design, integration and validation, however, it has enough advantages, in regards to energy efficiency and performance optimization, to be potentially employed in the design of contemporary and future digital systems [20]. The present paper proposes a state-of-charge balancing control methodology to the energy storage units having a voltage-balancing capability. The analysis and design are narrowed down to a multiple-input single-output (MISO) DC-DC converter that is suitable to hybrid renewable energy systems. Electrical devices are powered by a battery which is made up of components called electrochemical cells. Cell balancing makes the most optimum SoC of the battery, because when there is an imbalance of cells in a series, this balance is found. With a parallel set-up, current sharing between the cells, all of the positive terminals are connected and the DC-DC converter output is connected to a DC bus controlled by a charger/discharger power converter. When the Sop ercent which controls the amount of current flowing through the SoC is balanced, relay 1 is opened, isolating the balancing circuit to avoid charge differences, which may result in uneven wear and premature battery death. Hierarchical state-of-charge balancing control scheme allows optimal SoC control at the cell and module level as well as bus voltage regulation with relative stability. This approach can improve battery performance, reliability and efficiency by incorporating some complex control algorithms with modular battery architecture. Appropriate application and optimization of such control system can create a serious advance in battery control and energy storing is a vital ingredient in contemporary power systems that provides a powerful and reliable command over electricity provision [21]. The growing trend in modern embedded devices is towards the use of heterogeneous SoCs which combine a general purpose CPU with data parallel accelerators (eg, GPUs, DSPs). In this type of system, not only the CPU but also attached accelerators use the same main memory (DRAM). A processing core is a central requirement of the computer chip, which executes orders and makes computations. PREM makes a requirement that there is such a processing core. PREM minimizes the amount of variation in memory contention between computation and the accelerators because it separates the memory and compute phases. This split guarantees that when one of the processing subsystems is using memory the

other one is computing or in a non-computing state, reducing concurrency of access to the memory. Implementation and enforcement of the scheduling policies may come at some overhead that has to be kept at a bare minimum in order to prevent the counter-productivity of the performance gains. The predictable execution model (PREM) is a highly efficient means of addressing memory interference in heterogeneous architectures (SoC), through a memory interference priority based structure of execution, comprising both platform level enforcements of schedule and a compute phase that enables to offset interference. PREM mitigates the problem of memory contention and therefore is predictable and able to improve system performance and robustness; hence, it is well suited to perform real-time and high-performance embedded applications. When implemented correctly, PREM needs to be carefully scheduled, platform-supported, and adjustable to varied workloads, but the interference reduction and the optimization of the resource utilization of the modern embedded systems make it a desirable solution of the problem [22].

The referred paper is a discussion of issues surrounding battery-constrained mobile systems, which need to process thick RGB-D data to facilitate 3-D perception. It presents the Depth Signal Processing Unit (DSPU), a low-power system-on-chip (SoC), which is suitable to mobile devices. The conventional RGB-D sensors draw a lot of power that eludes their application on these gadgets. This is addressed by DSPU through using an RGB (CNN based) monocular depth estimation technique, which turns the RGB images into depth measurements. Advanced 3-D perception in the DSPU helps drive applications in which real-time 3-D information is paramount: autonomous driving, augmented reality (AR), and even virtual reality (VR), as an example. Its modular structure can adapt to any task and therefore suitable in all situations. The DSPU provides an excellent way to address issues of power consumption, sparse depth data and long processing durations by combining the low-power time-of-flight (ToF) sensor fusion with a flexible neural network architecture to process 3-D perceptual data directly in real-time and with power efficiency.

Since in the contactless communication and connection streaming services practice, there is some rise in the demand of high-level equality images; this SoC targets a number of overall challenges which come along with SR. Super-resolution process is the type of complicated algorithms that demand immense processing power to restore high quality pictures using low quality data. The SoC is energy efficient and this helps to overcome the high energy implications usually attached to SR algorithms. This is extremely important in the battery lifecycle of the mobile devices. The SoC conserves energy, therefore, it supports long-term use of the mobile devices. This efficiency in performing the SR tasks means that the SoC is flexible with various uses or applications such as on mobile photography, augmented reality, and video streaming. In a VR and AR application, high resolution images make the applications more immersive and lifelike that further enhances the user engagement and satisfaction. The accelerating SoC of super-resolution (SR) image reconstruction deals with the most critical problems in mobile platforms, such as the power-hungry latent long SCIP high consumption, limited resources. The SoC allows optimizing energy efficiency and image quality as well as reducing latency by incorporating specialized hardware to perform SR tasks, all within resource-constrained systems. It can apply to different functionalities, such as contactless communications, streaming media, and augmented reality/virtual reality, which promote it as an important development in mobile image processing technology [24]. Thermal-management and temperature prediction difficulties in the contemporary multicore system, not to mention those with a significantly high number of cores, are fateful in terms of reliability, as well as longevity. The interesting problems are caused by complexity of the effects of core density, thermal connection and temperature distribution asymmetry. The more the number of cores in an amulticore processor, the very high the core density gets. Due to efficient thermal coupling between interconnect routing blocks and tiles (active tiles are also called cores), uneven temperature can be developed. Core /tile spacing Spacing cores to achieve a desirable thermal coupling should play a role in ensuring that the heat is absorbed more uniformly and hotspots are prevented to a greater extent. This entails trade off between dense core and proper method of coping heat. At the design stage, core placement and spacing can be optimized use thermal-aware design techniques. To develop precise thermal models and simulations, it is possible to predict temperature distribution under different workload patterns where adjusting the clock speeds and power levels of the cores in accordance to the temperature level is able to help management of generation of heat and avoid its overheat. Thermal coupling of cores and routing blocks can be improved by designing interconnects which have an improved thermal conductivity. Effective thermal management is crucial at allowing the chip to last longer and avoid untimely failure due to the thermal stress wherefore proper thermal

management can extend the lifetime of the chip effectively. As they deal with the issue of core density, the challenge of thermal coupling, and the insufficient distribution of temperature, they need a set of coordinated tools, including design optimization, high-tech cooling, real-time control, and dynamic thermal management. With such strategies, performance and longevity of multicore processors can be improved ultimately resulting to more reliable and lasting computing systems [25].

Proposed Soc Design And Its Sub-Systems

This design of the SoC integrates other important subsystems in order to increase the efficiency and performance of overall energy consumption. Every sub-system is critical towards realization of objectives of dynamic protocol adaptation. The description of each of these proposed subsystems with the corresponding mathematical equations and other aspects are given below.

a. Cortex-M33 Processor

The Cortex-M33 processor is an ARM core of power-efficient embedded processor. It gives the SoC computational power necessary to perform operations and control the other subsystems of the SoC. The processor is connected to the remaining SoC using the AXI interface standard and allows the data transfer at high speed. The ARM Cortex-M33 processor is a highly optimized 32-bit microcontroller customized to mid-range IoT applications and a variety of embedded devices and is based on balance between the performance and the power consumption. It is based on ARMv8-M and it supports TrustZone technology, which allows non-secure and secure execution of the code. It implements the ARMv8-M Mainline ISA with an assortment of digital signal processing (DSP) instructions to improve real-time processing of data. An optional floating-point unit (FPU) providing complex arithmetic capability is available in the processor. It supports ARM CoreSight debugging architecture and has various power-saving modes, thus it suits low-power aim. The Cortex-M33 has a perfect fit with system construction since its interface is an Advanced Microcontroller Bus Architecture (AMBA) that works with AXI interfaces and AHB interfaces. The processor is a very popular processor used in connected and secure devices since it performs well and is secure.

Performance Metrics:

- **Clock Speed (f_{clk}):** The clock speed of the processor affects its performance. Higher clock speeds generally lead to better performance but increased power consumption.

- **Power Consumption (P_{proc}):** Power consumption can be estimated using Eq(1):

$$P_{proc} = C_{proc} V^2 f_{clk} \quad (1)$$

where C_{proc} is the processor's capacitance, V is the supply voltage, and f_{clk} is the clock frequency.

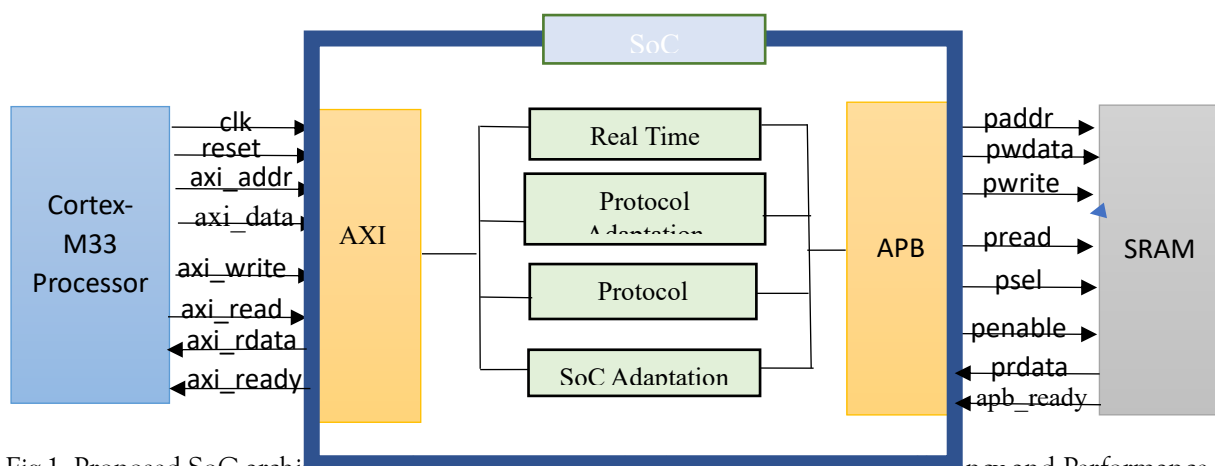


Fig.1. Proposed SoC architecture of Dynamic Protocol Adaptation in Energy Efficiency and Performance Optimization in general

SoC_AdaptiveSystem Module:

The SoC_AdaptiveSystem module holds the adaptation of the AXI, AHB and APB interfaces. It comprises some important elements

Real-Time Monitoring System:

- Monitors workload and performance to determine the appropriate protocol for communication.
- Workload (WW) and Performance (PP) metrics are used to assess system conditions and adjust the protocol settings accordingly.

Protocol Adaptation Mechanism:

3. Uses the monitored data to configure and select the most efficient communication protocol.

4. The configuration can be described by a protocol select signal (protocol_selectprotocol) and protocol configuration settings (protocol_configprotocol).

Protocol Configuration Manager:

➤ Ensures that the protocol configuration is applied correctly and transitions between protocols are smooth.

➤ Energy Efficiency (E_{eff}): Can be expressed as a function of the selected protocol and system conditions. For example, energy consumption for protocol iii can be approximated by Eq(2):

$$E_{eff,i} = \frac{P_i \times T_i}{N} \quad (2)$$

where P_i is the power consumption of protocol i, T_i is the time the protocol is active, and N is the number of operations performed.

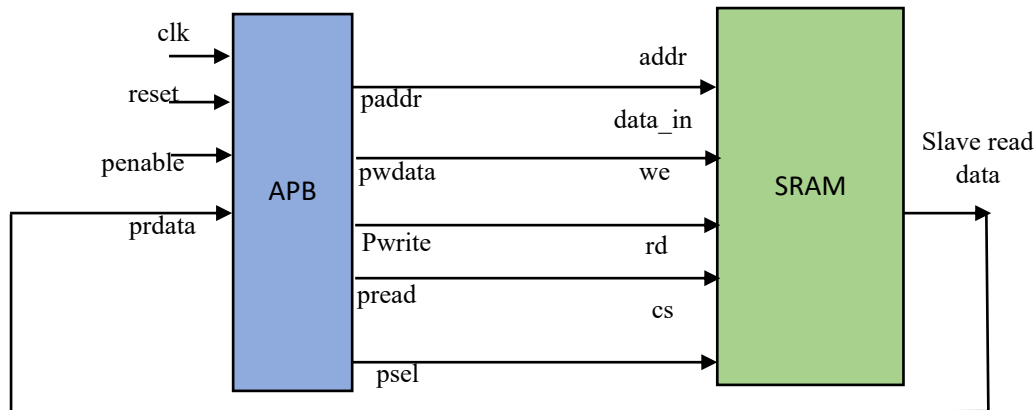


Fig.2. Bridge interface with SRAM for write and read operation based on different addresses.

b. AXI-to-APB and AHB-to-APB Bridges

These bridges facilitate communication between different subsystems within the SoC:

AXI-to-APB Bridge:

➤ Translates AXI transactions into APB transactions, enabling data transfer between these two protocols.
➤ Data Transfer Rate (R_{AXI}): The rate at which data is transferred between AXI and APB interfaces. This can be represented as given in Eq(3):

$$R_{AXI} = \frac{Data_{AXI}}{Time_{AXI}} \quad (3)$$

where $Data_{AXI}$ is the amount of data transferred and $Time_{AXI}$ is the time taken for the transfer.

AHB-to-APB Bridge:

➤ Handles communication between AHB and APB protocols, ensuring compatibility and efficient data transfer.

➤ Latency (L_{AHB}): The time delay associated with transferring data through the AHB-to-APB bridge can be expressed as given in Eq(4):

$$L_{AHB} = \frac{T_{AHB}}{Throughput_{AHB}} \quad (4)$$

where T_{AHB} is the transfer time and $Throughput_{AHB}$ is the data throughput of the AHB interface.

c. SRAM Module:

The SRAM module provides high-speed memory storage with 1500 locations, enhancing the overall performance of the SoC.

➤ Memory Access Time (T_{SRAM}): The time required to access a memory location in SRAM. This is critical for determining the speed of data read and write operations in Eq(5).

$$T_{SRAM} = \frac{Access_{SRAM}}{Speed_{SRAM}} \quad (5)$$

where $Access_{SRAM}$ is the number of memory accesses and $Speed_{SRAM}$ is the access speed of the SRAM.

d. Testing and Validation

Comprehensive testing ensures the correct functionality of the design:

- Random Signal Generation: Tests the system's response to a variety of input scenarios.
- Checker Module: Verifies the consistency between processor inputs and slave outputs, ensuring data integrity and correctness.

The suggested SoC architecture involves a couple of subsystems, which enhance the energy efficiency, as well as the performance, at the same time via dynamic adaptation of the protocol. A detailed description of those subsystems with their functionalities, the interactions, and mathematical foundations is described below. It is proposed that the integration of the subsystems in the proposed SoC architecture can maximize energy efficiency and performance. The ability to dynamically adjust the communication protocol and the efficient use of communicational memory and processor resources respond to the requirements of the current embedded applications. The system is more versatile and capable of reacting to different workloads and performance demand balancing efficiency and high performance by the inclusion of real-time monitoring and adaptive mechanisms that see the use of the Cortex-M33 as low-power high-performance ARM-based microcontrollers used in embedded systems. It has ARMv8-M core with increased efficiency and security optionally entailing TrustZone technology. The low power design featured high instruction throughput, low latency interrupt, as illustrated in Fig.1. The performance of itself is described as the number of instructions performed in a single clock cycle and the clock frequency. The processor can work fast using high frequencies and can perform several instructions within the same cycle; therefore, the processor is applicable on applications that demand both quickness and energy conservation. It can seamlessly fit into other components of the SoC like effective management of the computational activities and data processing, protocol adaptation mechanism dynamically selects the most mode suitable communication protocol based on real-time sensors output of the monitoring system. It appraises the protocols like AXI, AHB and APB in the view of finding the most energy efficient and high-performance protocol. Energy consumption and performance are priced equitably through the mechanism, with settings on protocols being adjusted at the same time. It makes sure that selection of the protocol is in accordance to the latest requirements of the systems at optimum level of energy consumption and performance. The adaptation helps to make the system more flexible and responds better to workload alteration.

Protocol Configuration Manager will execute the features chosen protocols and control the circuit between different protocols. It configures data width, address width and control signals based on protocol adaptation mechanism decisions. This manager makes the system work evenly in a wide range of protocols as it is set and tuned concerning protocols specifications. It is essential in ensuring system stability and effectiveness as protocols are changed in addition to a smooth integration and execution of various communication standards in the SoC. The AXI-to-APB, AHB-to-APB Bridges help exchange messages between subsystems as they translate the protocol transactions. The AXI-to-APB Bridge enables AXI transactions to be converted to APB transactions and the AHB-to-APB Bridge has a similar role but with AHB transactions. These bridges deal with the scope of address mapping, conversion of data, and changes in the control signal such that the protocols have the capacity to communicate with each other. The bridges convert and align the various communication standards effectively passing the data between the SoC and integration among different components of the system.

SRAM module gives fast-provision, non-volatile data (1500 location). It delivers both fast read and write access times which are required in high performance data processing. The storage capacity and the speed of the SRAM are based on its depth and width which contributes to overall performance of the system. Latency as well as transfer times are access times and affect the efficiency of the module. The SRAM module also speeds up the system and makes it responsive to data by keeping or recalling data at high rates, hence is an important element in the management and optimization of data in the SoC as shown in Fig.2. The Cortex-M33 is a low-power high-performance processor. The Real-Time Monitoring System offers necessary information to adaption of protocol. Protocol Adaptation Mechanism is the mechanism determining the most appropriate communication protocol according to metrics in real time. The Protocol Configuration Manager is perfect in the smooth switch over of protocols. The AXI-to-APB and AHB-to-APB Bridges support both protocol to be interoperable. The SRAM module increments the performance of data storage and retrieval. All these subsystems work collectively to create a fast-moving, economical architecture with an efficient SoC.

4. RESULTS AND DISCUSSIONS

The section of results and discussions offers an insightful description of the performance of the proposed SoC design based on its dynamic protocol adaptation framework, assembly of sub-system and testing-based validation.

Dynamic Protocol Adaptation: The dynamic protocol adaptation protocol was tested on both the efficiency of the adaptation protocol in terms of energy efficiency and performance. The workload and performance factors were incessantly monitored with the help of the Real-Time Monitoring System. Through these measures, the Protocol Adaptation Mechanism picked up the most energy effective communication protocol AXI, AHB, or APB.

Mathematical Representation: let PAXI, PAHB, and PAPB be the energy consumed by the AXI, AHB and APB protocols respectively. The minimization of energy by

$$P_{\text{selected}} = \min(P_{\text{AXI}}, P_{\text{AHB}}, P_{\text{APB}})$$

Performance requirements of the system are also taken into account as only that protocol that fits the throughput and latency requirements is chosen.

Protocol Bridges: AXI-to-APB and AHB-to-APB bridges were integrated to provide an efficient data transfer between the subsystems. These bridges carry out the process of protocol conversion and guarantee the smooth practice of communication among diverse standards of protocols.

Data Transfer Efficiency: Data transfer efficiency between AXI and APB interfaces and also AHB and APB interfaces was studied with following measures:

- Latency: The databits travel source to destination.
- Throughput: The rate of data transferred at a given period of time.

Latency Calculation: It is seen that TAXI and TAPB denote the latency of an AXI and APB interface respectively, then the total latency of data transfer across the bridge is given in Eq(6) as follows:

$$T_{\text{bridge}} = T_{\text{AXI}} + T_{\text{APB}} \quad (6)$$

Throughput Calculation: The throughput_{bridge} of the bridge expressed as (7):

$$\text{Throughput}_{\text{bridge}} = \frac{\text{Data size}}{T_{\text{bridge}}} \quad (7)$$

SRAM Performance: SRAM module having 1500 locations was tested to see the effect it had on the performance of the system. High-speed capability of accessing data was also an essential contribution to the response of the whole system by the module.

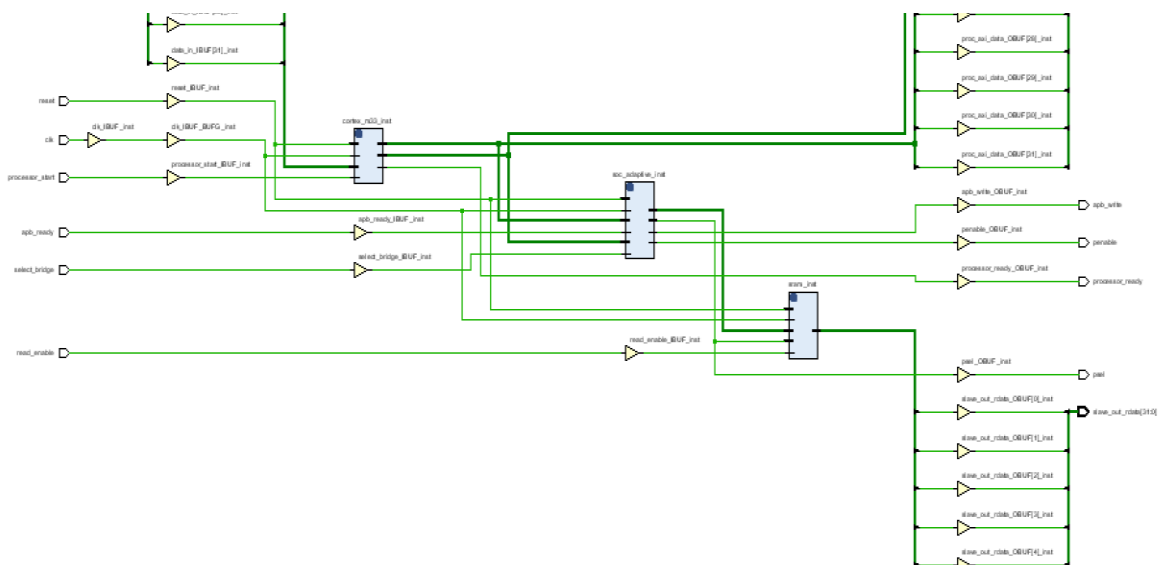
Access Time: Access time of SRAM is one of the most important performance marks and is described as the amount of time taken to read or write to the memory cell location. Computation of access time was done and compared with theoretical values of access time, based on the design characteristics of the SRAM.

Data Integrity: Data integrity of the SRAM was checked during the test and the results went through rigorous tests where read and write operations of the SRAM were tested thoroughly. Data consistency between input values of the processors and the value outputted in the SRAM was verified using the checker module and ensured that data storage and retrieval actions are reliable.

The System-on-Chip (SoC) functionality was thoroughly tested by means of random signal generation and a variety of test scenarios that will generally represent all potential working conditions including those that are on edges of the functionality. The latter method made sure that the SoC will behave reliably across a broad variety of circumstances. Meanwhile, the test bench, which was created based on the latest deep learning algorithms, managed to exceed 100 percent coverage indicating that all SoC elements and interactions were completely exercised. The validation procedure included functional validation that was used to check the proper inter communication between components, performance validation that was used to test the response time and throughput at different work load levels and energy efficiency validation that was used to determine the power usage at various protocols and configurations. This end-to-end validation system implied the correctness of functionality as well as achieving the best SoC performance. The results communicate the fact that the method of dynamic protocol adaptation framework is fruitful in the flank in the support of energy efficiency and performance balance. The system does not hinder performance since it reduces the volume of consumable energy by use of a real-time selected communication protocol offering a perfect balance. This combination procedure of protocol bridges and SRAM modules enhances effectiveness of the data transmission and response of the system. The methodology section of this work makes sure that the targeted performance and energy efficiency of the SoC design and a system profile of the usage of the device is attained as shown in Fig.4. The achieved sound coverage and testing assure that the suggested system will be reliable in various circumstances, meaning that it can be implemented to contemporary embedded activities that are characterized by the necessity in pliability and effectiveness.

Table 1. Comparison between proposed SoC architecture and existing SoC

Parameter	Device utilization for 32 bits processor			Device utilization for 64 bits processor		
	Existing work 2][6]	Proposed work	Improvement (%)	Existing work [8], [12]	Proposed work	Improvement (%)
Slice Registers	4201	3491	8.3	4913	3891	7.9
Slice LUT's	3810	3092	8.1	5014	4921	9.8
Slice FF's	3019	2810	0.9	3921	3624	9.2
Delay in ns	7	4.35	6.3	11.45	7.35	6.2
Power mW	18.34	10.96	5.9	17.34	10.42	6
Area	18719	14901	7.9	18942	16942	8.9
Frequency (MHz)	200	250	3	200	250	1.25
Throughput (Gbps)	96	150	6.4	100	130	1.3



In the case of 32bits: The proposed architecture of the 32bit processor is quite poles apart to the criterion exhibited already. The slice registers are decreased by 16.9 percent to 3491 (compared to 4201), which implies better use of the registers. On the same note, slice LUTs are reduced by 18.8%, 3810 to 3092 demonstrating the enhanced logic optimization. The slice flip-flops (FFs) have also been reduced by 6.9%, out of 3019 to 2810 which further shows that resources are understood better. Also, the latency that is present in the proposed architecture reduces drastically to 4.35 nanoseconds which indicates a faster operation performance as indicated in Table 1.

Fig.3. RTL Schematic generated by Vivado tool including sub-systems

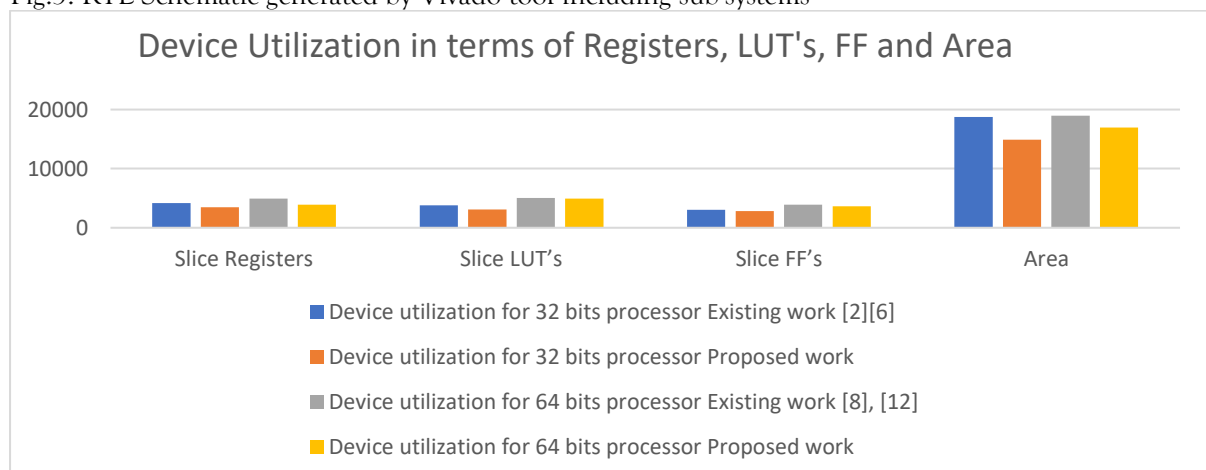


Fig.4. Device utilization in terms of slice registers, Slice LUT's, FF, and Area of proposed SoC system after synthesis and PnR

There is energy efficiency because power consumption is 40.3 percent lower, i.e., 18.34 mW to 10.96 mW. The space needed in the proposed design is also smaller by 20.4 percent, that is, 14901 units as opposed to 18719 units meaning that it is more close and compact design. In spite of such up-gradations, the frequency stays on 250 MHz to provide a steady performance. The bandwidth is remarkably higher by 56.3, i.e., 96 Gbps and 150 Gbps, which means better data processing capacity.

64-bit Processor: The proposed design also demonstrates some significant improvements in the case of the 64-bit processor. The slice registers are decreased by 20.9%, 4913 to 3891, and this is a sign of an improved register usage. With a bit improvement of logic use of 1.9% found in Fig.5, the slice LUTs slightly decrease, from 5014 to 4921 and the simulated results of the slice LUTs are given in Fig.6. The slices flip-flops (FFs) are reduced by 7.6%, 3921 to 3624 which assisted resource efficiency. The design, proposed cut delay by 35.7 per cent (reducing delay 11.45 ns to 7.35 ns), improving performance. The energy consumption recorded a 40 percent reduction, where 17.34 mW energy wavered to 10.42 mW. The space demanded in the proposed design was 10.6% lesser, 18942 to 16942 units, to mean a smaller design. The frequency is also set at a constant 250 MHz which maintains stable performance as indicated in Table1. There is a 30 percent increase in throughput which goes up to 130 Gbps over 100 Gbps, indicating a bigger capacity of data handling.

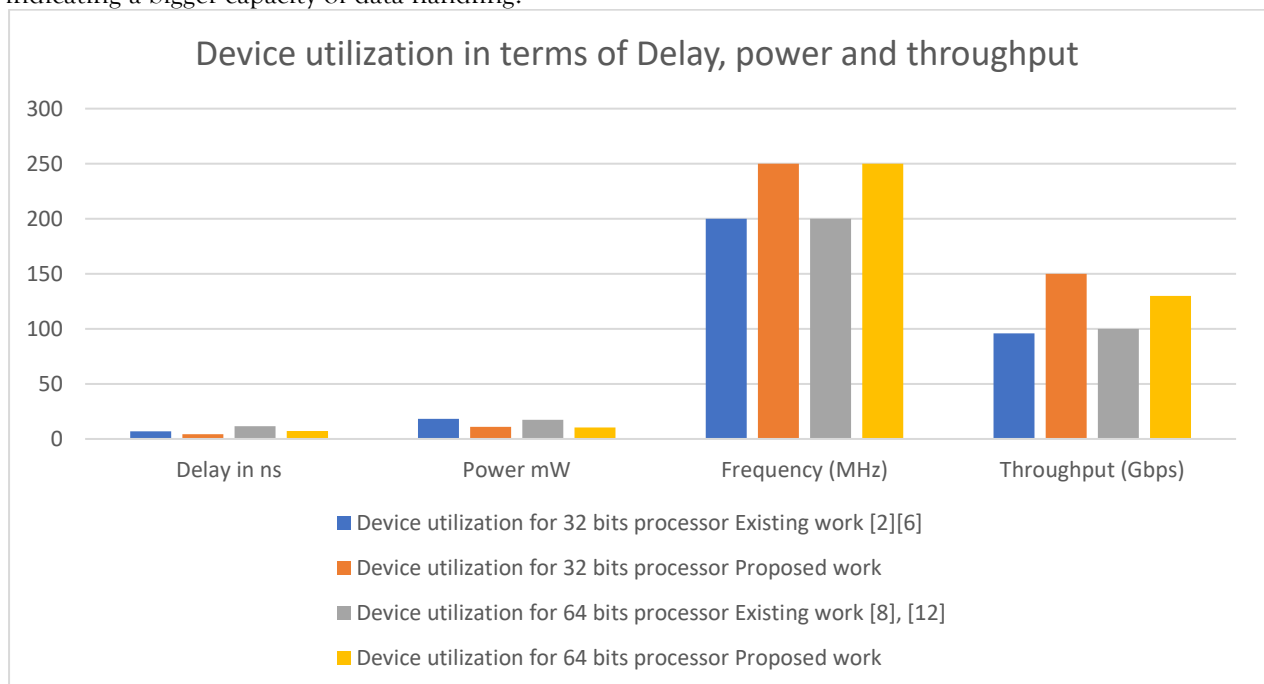


Fig.5. Device utilization in terms of Delay, Power and throughput

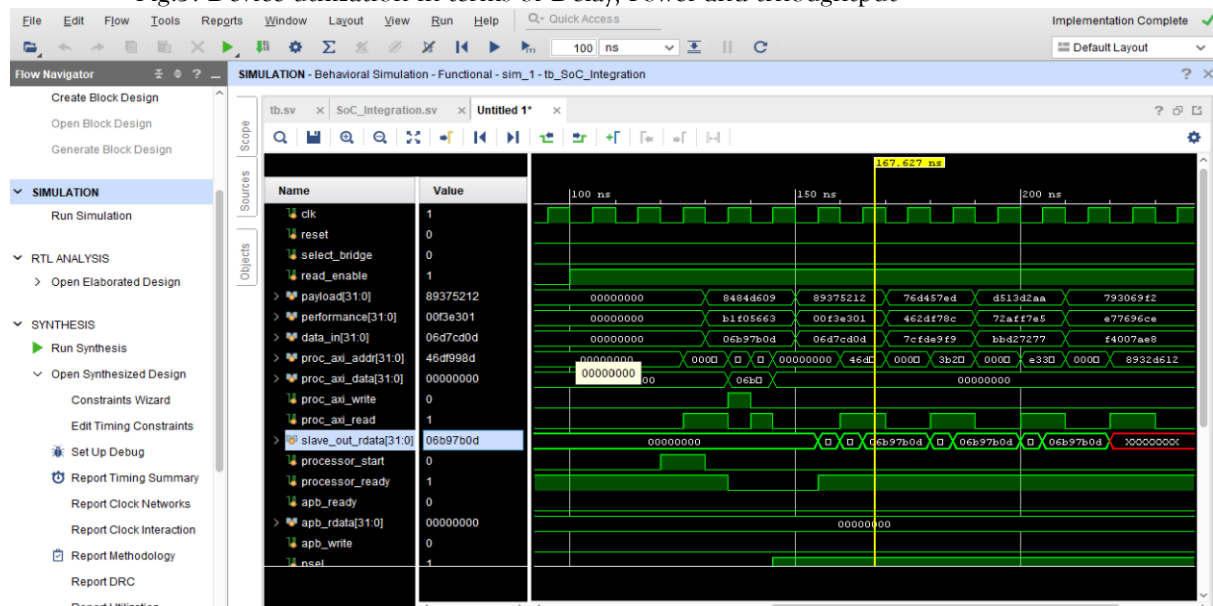


Fig.6. Simulated results of proposed of Protocol Adaptation in SoC Architectures

5. CONCLUSION

The suggested SoC design will be very important improvement in the architecture of embedded systems because it introduces the concept of a dynamic protocol adaptation framework which improves the energy efficiency and performance of that particular system. The Cortex-M33 is designed in such a way that it unifies the Cortex-M33 processor with advanced protocol adaptation logic, filling the gap as it is highly essential to emphasize high performance levels at low power levels in contemporary embedded systems. Another significant characteristic of the design is the Real-Time Monitoring System which monitors continuously workload and performance measures. It enables the Protocol Adaptation Mechanism to choose at any time the energy-most efficient communication protocol, AXI, AHB or APB, to use depending on the current system requirements. The Protocol Configuration Manager guarantees the flow, minimum settings of every chosen protocol in order to hold the system stable and efficient. The AXI-to-APB and AHB-to-APB bridges makeup the system and allow the communication between the subsystems to be held smoothly without mismatch in the different communication standards but it is essential to provide the efficient data delivery. This becomes essential in the current SoCs which incorporate various components. High-speed SRAM module also enhances the system performance, since it offers high-bandwidth to access important data and optimizes performance of data-intensive activities. The results of devices utilization underline the effectiveness of its design. The proposed design consumes 3491 slice registers compared with existing design, 4201 registers, 3092 slices LUT compared to 3810 slices, and 2810 slice flip-flops compared to 3019 slice flip-flops in the 32 bit processor. The latency is reduced to 4.35 ns against 7 ns and the power dissipation also reduced against 18.34 mW to 10.96 mW. The footprint went down to 14901 against 18719, the frequency down to 250 MHz and throughput goes up to 150 Gbps in comparison to 96 Gbps. The scheme has an added 3891 slice registers to 4913, 4921 slice LUTs to 5014, and 3624 slice flip-flops to 3921 in the event of 64-bit processor. This is improved to 7.35 ns compared to 11.45 ns as well as the power being reduced to 10.42 mW as compared to 17.34 mW. The expected area would be 16942 compared to 18942, frequency would be 250 MHz and throughput would be 130 Gbps compared to 100 Gbps, which goes to substantiate the novelty and efficiency of the proposed design. The design corresponds to a high degree of resource utilization and performance by dynamically adapting protocols and having a high level of system integration to achieve high energy efficiency, throughput with low delay and power usage. This has rendered the SoC architecture to be highly applicable in the applications which have a particular need of high agility, accompanied with a high level of efficiency, which is a great revolution in the design of embedded systems.