

Methods for Creating Power-and Latency-Efficient 10T and 14T SRAM Cells

Mr M V MAHESWAR REDDY, Dr VASANTH SAMINATHAN , Mrs K APSANA

Abstract-

Since the introduction of smart phones in the previous decade, almost all signal processing components have merged into a single piece of hardware. Many functions that were formerly performed by dedicated hardware are now available on our smart phones, from calculators to sophisticated picture editors. The whole adder, the simplest computing unit, is taxed under the weight of all the arithmetic operations performed on signals (adding, subtracting, multiplying, dividing, numerically integrating, convolution, and filtering). Modern mobile apps' complicated algorithms need highly fast, low-power technology. To prevent read sensing failure and weakened cell stability from half-select writes, leaks must be sealed. An equalized bit line approach is proposed to reduce the likelihood of this happening by making leakage independent of data pattern. This has the added benefit of enhancing RBL sensing. In addition, a fast local write-back (WB) approach is used in this procedure to perform a half-select-free write operation. Using a hierarchical bit line design, it allows for a quick WB action after a local read to further protect the original data without slowing things down. The literature review provides a concise summary of the power and delay savings that may be achieved by using 10T and 14T SRAM with non-destructive column selection.

I. INTRODUCTION

The primary causes of power consumption in CMOS circuits are the currents flowing through them during a short circuit, leakage, and switching. Each is presented in its own section below. Both a pull-up (p-network) and pull-down (n-network) network are present in a static complementary metal oxide semiconductor (CMOS) circuit. The two networks' logic operations complement one another. In a normal configuration, where both the input and output states are stable, only one network is active, conducting the output to either the power supply node or the ground node, while the other network is passive, blocking the current. Whenever two networks are switched on while simultaneously active, a short circuit current occurs. For instance, an inverter's input signal may go from 0 to V_{dd} . The input voltage might be higher than V_{tn} for a brief period of time before falling below $V_{dd} - |V_{tp}|$.

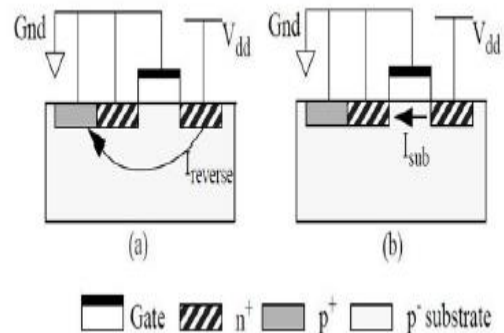


Figure 1: Two forms of leakage current: (a) current via a reverse biased diode, and (b) current below the diode's threshold.

Assistant Professor^{1,3} , Associate Professor²
 Department of ECE,

Viswam Engineering College (VISM) Madanapalle-517325 Chittoor District, Andhra Pradesh, India

Short-circuit current flows from the power supply line to the ground via both PMOS-transistor (p-network) and NMOS-transistor (n-network) transistors during this time period. The short-circuit current in a basic inverter [6] may be explored using simulation in SPICE, although the precise analysis is difficult. Short-circuit current is shown to be proportional to input signal slope, output load, and transistor size [5]. In a "well-designed" circuit, the short-circuit current will only account for around 10% of the total power used [7].

1. Power Loss:

Leakage currents are caused by two different sources: currents flowing through reverse biased diodes and currents flowing through non-conducting transistors. As the threshold voltage increases, the leakage currents increase exponentially with the leaking area. Unlike with certain logic types, designers are often unable to alter the leakage currents, which are a function of the underlying technology. With current technology, the leakage current is on the order of pico-Amperes, but it increases as the threshold voltage decreases. Leakage current is a major issue with some devices, such as very big RAMs. In most modern digital designs, leakage current is not a major issue. However, with 0.06 mm technology, leakage current may use just as much energy as switching current.

Leakage current in deep-submicron technology may be minimized by using several threshold voltages. Capacitances cause switching currents, which is why switching power is so important. Capacitances at the links between nodes.

II. POWER AND AREA REDUCTION TECHNIQUES AT CIRCUIT LEVEL

1. Transistor Reordering:

The functioning of a circuit including a series of Metal Oxide Silicon Field Effect Transistors (MOSFETs) is not affected by the order in which the transistors in the chain are linked. Flexibility in the relative arrangement of transistors is common in complex gates, where groups of transistors may be coupled in series. By taking advantage of this flexibility in transistor arrangement, lower power dissipation in CMOS devices is possible. Minimizing drain-source capacitance and using signal probability algorithms to decrease the number of transistors are two strategies to reduce power dissipation since transistor order in a CMOS gate substantially impacts the switching activity.

Second, Transistor Size:

Minimizing power consumption within a certain delay restriction may be achieved by appropriately sizing transistors in CMOS circuits. There are now two distinct classes of transistor size optimization algorithms: Timing constraints can be met and power dissipation can be minimized through linear programming and simulated annealing-based methods. The breadth-first method initially sizes each gate to maximize efficiency with respect to power consumption. If the delay restriction is met by the power-minimal arrangement, the procedure ends. Until the time goal is achieved, transistors on key paths are scaled down using power-delay optimum scaling. In [2], the authors offer such an algorithm.[3] This method is more complicated than its predecessors because it accounts for both the power dissipation caused by charging the circuit capacitance and the power dissipation caused by a short circuit. An further nugget of insight from [4] is that active area is not a good predictor of circuit power usage. It is shown that a CMOS circuit's power consumption is a convex function of its active area. The goal of a transistor size method to minimize power dissipation is distinct from that of decreasing the active area.

Reducing glitching activity in a circuit requires roughly balancing the delay of all true paths that converge at each gate, as shown in Figure 4. This is because path balancing results in nearly simultaneous switching on the various gate inputs, and any potential hazards at the output of the gate are removed. As a result, the circuit's average power dissipation decreases. Technology mapping is not required prior to doing path balancing. Typically, this is done by logic decomposition or selective collapse prior to technology mapping. Delay insertion and pin reordering, the results of a technology map, make this possible.

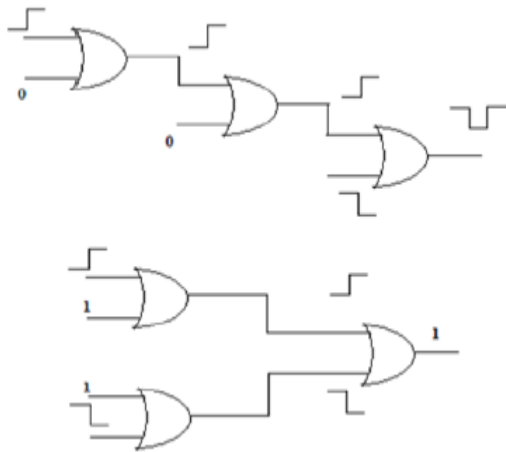


Figure 2: A real-world application of route balancing in action. Selective collapsing is based on the premise that a node's output arrival time may be adjusted by collapsing its fan-ins. Through logic decomposition, the level difference between the inputs of the nodes controlling high capacitive nodes may be minimized. The method of delay insertion aims to equalize the delays of all circuit pathways. The challenge of delay insertion is to minimize the amount of delay elements while maximizing their impact on reducing spurious switching. At last, rearranging pins may equalize delays along different paths. Because the delay characteristics of CMOS gates change depending on which input pin triggers a transition, this is feasible. Selective collapsing is based on the premise that a node's output may have its arrival time modified by collapsing its fan-ins. Through logic decomposition, the level difference between the inputs of the nodes controlling high capacitive nodes may be minimized. The method of delay insertion aims to equalize the delays of all circuit pathways. The challenge of delay insertion is to minimize the amount of delay elements while maximizing their impact on reducing spurious switching. At last, rearranging pins may equalize delays along different paths. Because the delay characteristics of CMOS gates change depending on which input pin triggers a transition, this is feasible. Multiple Power Sources and Variable Voltage Voltage:

Future digital systems will have to deal with the challenge of voltage scaling. The need to create complicated, high-performance systems on a chip is the primary motivation. To deal with strict power limits, it is expected that different ASICs and memory parts would switch to lower supply voltages. The lack of a comprehensive chip set for assembling systems with lower supply voltages is a major drawback of this strategy. Using a combination of supply voltages on the chip helps alleviate this issue

to some degree. Clocking-The clock may be a major source of wasted energy in synchronous systems. In most cases, this is the sole switching signal, and it must power a complex clock tree. In addition, there is sometimes a great deal of extra switching activity as a result of the clock changeover. As a result, modern circuits often have clocks that can be adjusted by the user. In other words, sub-clocks are generated from the master clock and may be adjusted to run slower than the master clock or even cease entirely. Therefore, the circuit is divided into sections, each of which operates on its own derived clock. The amount of energy saved this way depends on the specifics of the application. By turning off the electricity to the modules that aren't being used, we can save a lot of energy. Each node in the clock tree is assigned a binary string of 1s or 0s reflecting the active/idle state of the node in that time slot, which is then gated to turn ON/OFF sections of the clock tree during the active/idle mode.

III. LITERATURE SURVEY

Algorithms specific to digital signal processing and application-specific integrated circuits (ASICs), such as convolution, correlation, and digital filtering, depend on the efficient implementation of arithmetic circuits. These algorithms can only be executed on hardware with specialized ALU and MAC architectures. The primary components of these arithmetic units are adders, multipliers, and counters, and this chapter explores the many techniques and architectures suggested for their creation.[3]

In this study, by Aaina Nandal et al. (2018), we demonstrate ECRL using the sleepy keeper approach. The power consumption and power delay product (PDP) of the proposed adder are both lower. The circuits have been modeled in Mentor Graphics' 0.18μm CMOS simulation environment. The suggested Full adder has a latency of 135.97ns and a power dissipation of 129.819pW at 1.8V supply voltage. Multiple simulations have been run with supply voltages ranging from [1.0] V to [1.8] V. The suggested full adder circuit's power dissipation has been compared to those of previously published full adder designed circuits, and the results show that the proposed method yields superior results.

In this research, S. Lakshmi et al. (2018) offer a unique CMOS hybrid full adder circuit with energy efficiency of 1 bit, which can optimize performance factors such as power and latency. Both the traditional CMOS and the CMOS-TGL (complementary metal oxide semiconductor-transmission gate logic) models are used in the first design. To assure low power and fast speed, the suggested model is a combination of Complementary

metal oxide semiconductor (CMOS), pass transistor (PT), and modified gate diffusion input logics (MGDI). We examined and tabulated the performance metrics of 1 bit full adders, including power, delay, and area.

The cadence virtuoso tool was used to realize all the circuits in 90nm technology with a 1.2V supply. (2019) Gautam Nayan et al. When it comes to very large scale integration architectures like microprocessors and digital signal processors, addition is one of the most essential processes. Therefore, it is incumbent upon the adders to bring about a fast operation. A unique 8-bit adder design with a modified Gate Diffusion Input (m-GDI) implementation is proposed in this study. A 4-bit ripple carry adder (RCA) and a 4-bit carry look ahead adder (CLA) are two of the most fundamental building components of an adder. In comparison to conventional CMOS design, the suggested adder architecture uses 35% less power while using 70% less space and 71% less latency. Cadence Virtuoso Tool in 180nm technology is used to realize the suggested adder.

Reference: Somashekhar Malipatil et al. A novel method for designing low-power digital circuits, called Gate Diffusion Input (GDI), is suggested. With this method, digital circuits may have their space and energy needs drastically reduced. A three-transistor XOR gate and a two-transistor multiplexer (Mux) form the basis of the CMOS complete adder layout shown here. In this study, we develop a complete adder with just 8 transistors and implement voltage scaling by lowering the supply voltage. The entire area of this suggested complete adder is 144 m², and the power consumption is 4.604 mW. 2.

In 2020, M. Keerthana et al. In digital and VLSI systems, adders are a crucial component. In digital systems, arithmetic operations play a crucial role. The primary focus of study in very large scale integration (VLSI) systems is the minimization of transistor size for the purpose of enforcing any other digital system. The suggested design is realized via the use of many logic systems, each of which fulfills a specific function inside the hybrid system. This architecture is a one-bit hybrid Full Adder cell. A 22-nm CMOS hybrid full adder is used to study the suggested approach. Simulation findings show that the suggested design is very efficient in terms of both power consumption and latency. Data route simulations revealed that entire adder circuits are used in today's high-speed central processor units. This hybrid Full Adder is often used in nanotechnology because of its decreased latency and improved efficiency. With a 0.8-V supply and 22-nm technology, we observed that the average power usage was under 1.1055 W, with a relatively low

latency of 7.0415 ps. When compared to earlier complete adder designs, this kind of adder allots far more efficiently in terms of power, high speed, and space.

E. Veera et al. (2018). Boopathy. In this study, we develop a low-cost Manchester carry chain (MCC) adder for 8 bits using multioutput domino CMOS logic. This adder generates parallel carries using two 4-bit carry chains. Short carry chains are a feature of the recommended 8-bit adder module. This feature allows it to be used in the realization of more complex adders, and it also leads to improvements in operation speed when compared to comparable adders based on the traditional 4-bit MCC adder module.

Citation: A. Sedhumadhavan et al. In this study, using SPICE, we investigate the speed and power performance of a reverse carry propagate full adder (RCPFA) based on static and dynamic circuits technology. A defining characteristic of RCPFA is that the carry signal travels from the most significant bit to the least significant bit. As a result, the input carry signal is given far more weight than the output carry signal in RCPFA circuits. With a 0.8 volt power supply and 22 nm CMOS strained silicon technology, the RCPFA circuits were developed. Last but not least, research into how temperature affects the speed and power performance of RCPFA adders has been conducted. The simulation findings reveal that the dynamic RCPFA circuits have superior speed and power than the static RCPFA circuits.

Tabassum, Zarin, et al., 2018. This study analyses and contrasts the transistor count, power dissipation, latency, and power delay product of four adders using four distinct logic techniques (Conventional, transmission gate, 14 transistors, and GDI based approach). Using the Cadence tool from the readily accessible GPDK - 90nm kit, the process is carried out on the virtuoso platform. NMOS and PMOS have fixed widths of 120nm and 240nm. The complete adder with a transmission gate has the benefit of fast speed but uses more energy.

IV. MOTIVATION FOR THE RESEARCH

Because of the dense integration of data channel components in portable application processing architectures, battery life is drastically shortened. The majority of the energy used by a digital CMOS device comes from dynamic power dissipation. Dynamic power dissipation may be significantly reduced by technological scaling, albeit this comes at the price of leaky power dissipation. Bringing down the supply voltage will lessen the dynamic power, but

it will cause a significant delay. Thus, the study is focused on lowering critical latency and limiting dynamic power dissipation at the circuit level. The study is focused on decreasing the gate count of CMOS arithmetic circuits, since the leakage power dissipation is related to the area of the CMOS circuits. Delay minimization is also taken into account with the reduction of power and area of data route components utilized in processing architectures in this thesis. Adders, multipliers, and counter circuits are taken into account. In this study, we suggest the following categories of circuits and examine how they could be used in various processing contexts. Adders-Decimal adder is an essential component of ALUs created for business and commercial applications, whereas Full Adder (FA) is the most common adder circuit used in processing systems. They are crucial because they are located on the critical route, which determines the system's overall efficiency. These adders need to be carefully designed to minimize processing architecture's latency and power consumption. There are restrictions imposed on FA designs that use the same logic style for both sum and carry. The threshold voltage drop issue plagues FA implemented with complementary pass logic, and the driving capacity of FA implemented with CMOS transistors and transmission gates is subpar. Newer hybrid FA with carry over Branch Based Logic (BBL).

V. SYSTEM LEVEL DESIGN

Both the hardware and software of a system may have an impact on its overall power consumption. gating, overlapping, and shunting are the three most common methods for charging and draining a node. Hardware/software separation, hardware platform choice (application-specific or general-purpose CPUs), resource sharing (scheduling) approach, and so on are all part of the system architecture. There is often the highest opportunity for power reduction when low power approaches are adopted during the system design phase.

Finding the optimal low-power solution in a vast design space is challenging at the system level since precise power analysis tools are few. However, software power optimization is possible provided, for instance, the instruction-level power models for a certain CPU are accessible [59]. It has been noticed that speedier code and more frequent cache utilization tend to result in lower power consumption. Power consumption, which is a function of the processor's internal switching, is affected by the instruction sequence as well. At the system level, two

of the most common low power approaches are power-down and clock gating. Power savings are achieved by turning off the inactive hardware components. As shown in Fig. 3, gating the clock driver may lower the switching activities, which account for 30–40% of the overall power usage.

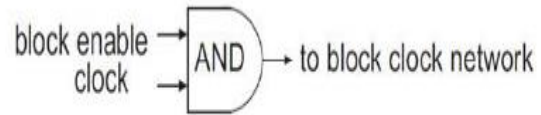
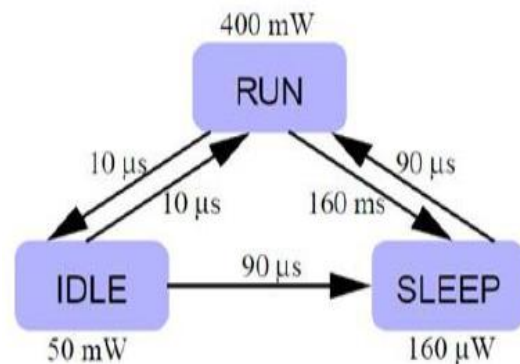


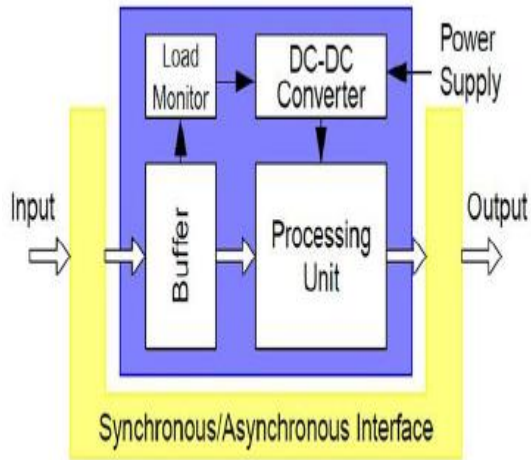
Fig 3. Clock gating.

Turning off the electricity to the whole infrastructure is an option. Sleep mode is often utilized in low power CPUs. There are three power states for the Strong ARM SA-1100 CPU, and the average power consumption changes throughout the modes [60]. ACPI, or advanced configuration and power management interface, allows applications to make use of these power states. In the last year, power management has been a major focus in OS development. Microsoft's desktop OS is one example of an OS that has APM (advanced power management). [4][5]



Power states diagram for a robust ARM SA-1100 microprocessor. The system was designed with high productivity in mind. The quantity of computing that is required, however, changes with time. Another low-power technique is to modify clocking frequency and/or dynamic voltage scaling to satisfy performance needs. After a certain amount of time has passed, the voltage of the power source may be decreased since the necessity for maximum performance has diminished. Either automatically (via load monitoring and voltage regulation) or at a set time, the voltage may be reduced. There is also a lack of research on asynchronous design techniques for low power design.

When compared to their synchronous equivalents, asynchronous systems have a number of benefits. These include lower peak current, no spurious transitions, automatic power down, and non-global clocking. The dynamic voltage scaling method is only one example of a low-power strategy that may be used with the asynchronous design approach to further reduce power consumption [6]. This concept is shown in Fig. 5.



Voltage scaling in an asynchronous design, as seen in Figure 5.

Algorithm Complexity: The amount of energy used is significantly affected by the algorithm selected. By using a fast Fourier transform instead of a direct computation of the DFT, the number of operations for a 1024-point Fourier transform, for example, may be reduced by a factor of 102.4, and the power consumption is also projected to be reduced by an equivalent amount. It is the responsibility of algorithm designers to find a solution that minimizes energy use while yet satisfying all criteria. The computational and communication/storage components of algorithm costs are taken into account. Both the number of operations and the related communication and storage expenses contribute to an algorithm's overall complexity. The number of required operations, the cost per operation, and the volume of data sent across long distances are the primary concerns of algorithm selection. Regularity and locality in algorithms are powerful techniques for making them more efficient [7]. If an algorithm's complexity can be reduced, it could use less energy since fewer operations will need to be performed. If an algorithm's concurrency can be raised, it may be possible to apply alternate approaches, such as voltage scaling, to reduce power

usage. An algorithm's regularity and locality affect its ability to govern and communicate with hardware.

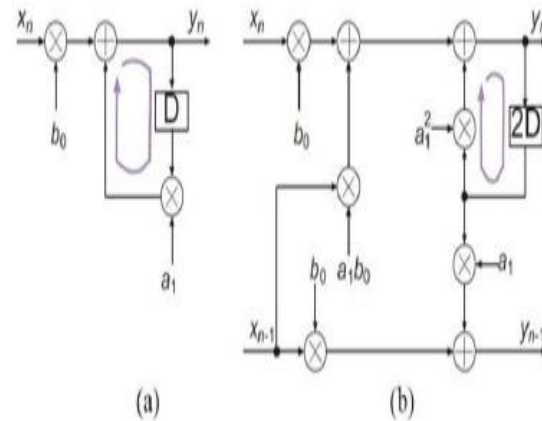


Figure 6(a): Initial data flow diagram. Unrolled version of the signal flow diagram (b). In certain cases, such with wave digital filters, faster algorithms may be used with voltage-scaling for energy-efficient applications. Saving energy at the circuit level is usually less feasible than at higher, more conceptual levels. But this cannot be ignored. The potential for energy savings is substantial because of the widespread use of these basic cells. An improvement of only a few percentage points in the efficiency of D flip-flops might have a significant effect on power consumption in highly pipelined systems. The switching in CMOS devices is the root cause of their variable power consumption. In conventional combinational logic, spurious transitions account for 10%-40% of the total switching power utilized. In certain cases, such when multiplying an array, spurious transitions might be common. To reduce the occurrence of spurious transitions, it is suggested that the delays of signals from registers that converge at a gate be close to identical. Two techniques [12] include adding buffers and making sure devices are the right size.

The total load capacitance increases after adding a buffer, yet the erroneous switching is still reduced. This strategy is known as path balancing. Many logic gates have interchangeable inputs, meaning that changing one input does not change the logic operation of the gate. Different gates exist, such as NAND, NOR, XOR, and so on. However, power consumption might shift with the order of inputs. In a two-input NAND gate, for example, the A-input, which is nearer the output, consumes less power than the B-input, which is nearer the ground, for the same switching activity factor. It is recommended to prioritize the usage of input pins with reduced power consumption for switching. Using less energy in this

way doesn't cost anything. The statistics of switching activity factors for different pins must be known in advance for pin ordering to be beneficial [13][14][15].

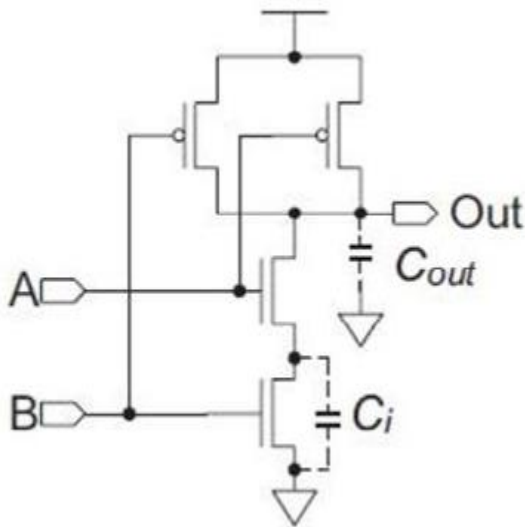


Fig 7. NAND gate.

VI. CONCLUSION

This article summarizes recent studies of the numerous possible layouts for one-bit full adder cells. The power consumption, latency, operating frequency, and transistor density of several full adder cells are compared. According to the study, the new 14T and 10T are much superior to all low-supply-voltage existing designs in terms of signal strength, power consumption, and data transmission rate. This circuit is well suited for usage in arithmetic circuits and other VLSI applications because to its low power consumption and high performance. In this research, we use just one bit of information to analyze typical full adder circuits. The factors of power, time, their product, distance, and loss at a threshold level were analyzed. Due to its importance as a building component of many types of computer hardware, the transistor-level design of the full adder circuit has advanced throughout the years. The most widely cited designs from the last two decades were chosen for this comparison. This article may serve as a useful resource for VLSI designers and researchers when deciding which circuit to utilize in a certain computing block or where to hunt for optimization possibilities.

REFERENCE

- [1] Aaina Nandal; Manoj Kumar Design and Implementation of CMOS Full Adder Circuit with ECRL and Sleepy Keeper Technique 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) Year: 2018 DOI: 10.1109/ IEEE Greater Noida (UP), India.
- [2] S Lakshmi; C Meenu Raj; Deepti Krishnadas Optimization of Hybrid CMOS Designs Using a New Energy Efficient 1 Bit Hybrid Full Adder 2018 3rd International Conference on Communication and Electronics Systems (ICCES) Year: 2018 DOI: 10.1109/ IEEE Coimbatore, India.
- [3] Gautam Nayan A Comparative Analysis of 8-bit Novel Adder Architecture Design using Traditional CMOS and m-GDI technique 2019 International Conference on Communication and Electronics Systems (ICCES) Year: 2019 DOI: 10.1109/ IEEE Coimbatore, India.
- [4] Somashekhar Malipatil; Vikas Maheshwari; Marepally Bhanu Chandra Area Optimization of CMOS Full Adder Design Using 3T XOR 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET) Year: 2020 DOI: 10.1109/ IEEE Chennai, India.
- [5] M. Keerthana; T. Ravichandran Implementation of Low Power 1-bit Hybrid Full Adder using 22 nm CMOS Technology 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) Year: 2020 DOI: 10.1109/ IEEE Coimbatore, India.
- [6] Boopathy. E Veera; Priva. M Swadhi; S. Sujitha; R. Susmitha; S. V. Sonia Realization of High Speed Low Power MCC Adder using Dynamic CMOS Transistors 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) Year: 2018 DOI: 10.1109/ IEEE Coimbatore, India.
- [7] A. Sedhumadhavan; S. Sabariesh; V. Shanmathi; K. Ramya; R. Venukumar; J. Ajayan Study of Performance Comparison of Static and Dynamic Approximate Reverse Carry Propagate Adder Using 22 nm CMOS Technolog 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) Year: 2020 DOI: 10.1109 IEEE Coimbatore, India.
- [8] Zarin Tabassum; Meem Shahrin; Anika Ibnat; Tawfiq Amin Comparative Analysis and Simulation of Different CMOS Full Adders Using Cadence in 90nm Technology 2018 3rd International Conference for Convergence in Technology (I2CT) Year: 2018 DOI: 10.1109/ IEEE Pune, India.