SWS

Design of Low Power Cam Memory Cell for the Next Generation Network Processors

K. Muralidharan[1], S. Uma Maheswari[2]

[1]Assistant Professor/ECE, Coimbatore Institute of Technology, Coimbatore, India
[2]Professor/ECE, Coimbatore Institute of Technology, Coimbatore, India

E-mail: [1]muralidharan.k@cit.edu.in

Abstract

In the modern world, high performance embedded applications in the field of multimedia, networking, and imaging are increasing day by day. These applications require high performance and more complex out-of-order superscalar processor. These complex dynamic instructions scheduling superscalar processors need higher levels of on-chip integration designs which are often associated with power dissipation. These out-of-order superscalar processors achieve higher performance compared to other processors by simultaneous fetching, decoding and execution for multiple instructions in out-of-order that are used in the next generation network processors. The main data path resources of the processor use CAM+RAM structure which is the major power consuming unit in the overall out-of-order processor design. The proposed new design of CAM+RAM with power-gating technique reduces the overall average power consumption compared to the conventional design without any significant impact on their performance.

Keywords: Static Power, dynamic power, CAM, SRAM

1. Introduction

Nowadays, our modern society moves towards the direction of high performance and energy efficient computing technology. Out-of-order superscalar processors are designed to attain high performance and are energy efficient. But this performance improvement will increase the overall chips power and their energy dissipation [1, 2]. Minimization of power consumption in the out-of-order superscalar processor has become important due to their need in network routers, next generation networks, all portable systems etc. Users demand longer battery life and this can only be obtained either by increasing the battery capacity or by

increasing the logic efficiency [3, 4]. But increasing the battery capacity also increases the weight of the system. Reducing power by using logic or minimizing the hardware is essential to meet the requirement in superscalar processor based systems [5].

As heat dissipation of components increases, it becomes more difficult to provide sufficient cooling by good packages, heat sinks or fans and ultimately the cost increases. Furthermore, higher temperatures increase the strain on the component and hence reduce its constancy [6, 7]. Other electrical issues need attention to provide a supply with proper capacity demands, like number of bond wires between the chip and the package, and a high amount of the signal routing area occupied by power distribution [8]. Power efficiency is made an important factor for the superscalar processor by presenting together the facts that, high current densities can lead to electro-migration at the system level, and highest power requirement demands huge and expensive power supplies [9-11]. Designing low power superscalar processor-based system requires different methodologies to be implemented at every level of abstraction. Power dissipation is either in static or dynamic manner. Static power dissipation is caused due to leakage and short circuit currents. The dynamic power dissipation is due to the occurrence of switching activities within the circuit. Dynamic power is the major contributor to the reduction of performance in the superscalar processor [12-14].

CAM+SRAM are used in out-of-order super scalar processors, to clear cache burst in the cache memory. In order to maintain the specified order of memory operations, all in-flight memory instructions are kept in a particular order [15]. Because of this, larger amount of CAM+RAM memory structures are required in the overall architecture. In addition to this, superscalar processors, wider their issue width in order to get higher performance.

2. Related Work and Background

The bypass hint mechanism for the reduction of register ports, by avoiding unnecessary reads and writes of register file where the values are bypassed, was proposed by Park et al. (2002) in the paper, "Reducing design complexity of the load/store queue" [16]. The Decoupled renaming technique was used to separate physical tagging and dependence of register operands. Write banking added with bypass mechanisms exhibit banking conflicts. Instructions commit their results in program order, was developed by Abella, Jaume et al. (2006) in the proposal, "SAMIE-LSQ: Set-associative multiple instruction entry load/store queue" for power reduction [17]. Sha et al. (2006) designed and developed a model without Store Queue (NoSQ) in out-of-order engine using Speculative Memory Bypassing (SMB)

methodology. Store load issue bypassing reorder predictor was used to identify whether the load bypasses or identifies the communicating store queue [18]. Tsai et al. (2008) analyzed the address allocation compression technique for all the load store queue in order to improve their scalability and power efficiency [19]. They presented their results in the article entitled "Power-efficient and scalable CAM+RAM design via address compression". Castro et al. (2005) implemented an efficient CAM+RAM filtering state mechanism for significant amount of energy reduction and presented it in their research article entitled "A Power Efficient and Scalable Load-Store Queue Design" [20]. The Load bypass and load forward were major technique utilized in the out-of-order processor [21]. Each entry of CAM+RAM gave either load effective or store effective address [22-24]. Every load instruction executes store queue to conclude whether load forward or load bypass is needed and every store instruction executes load queue for data forward in the same address location [25, 26].

3. Conventional Load and Store Queue

Existing CAM+RAM is considered for power analysis [27]. It is basically a CAM+RAM structure which includes the above features for high performance with low power. The structure of existing CAM+RAM is implemented and analysed for finding new technique to reduce the power further for the same performance as presented by Homayoun et al. (2011). Figure 1 shows the structure of the existing CAM+RAM implementation. The Upper segment taglines are inaccessible with the bottom segment taglines. Bottom segment is power gated during the normal period. Isolation between bottom bit line segment and corresponding upper bit line segment is achieved by tuning off transmission gate. During an associative search process operation, the tagline broadcasting in the processor is responsible for 46% of power consumption. In CAM+RAM, the tagline broadcasting must go to the entire width. Obviously, it has a large number of one-bit comparators which increases wire capacitance. The equivalent broadcast tagline capacitance is specified in the equation (1).

$$C_{tagline} = C_{gate}(CompareEN) * N_{rows} + C_{diff}(Compare\ Driver) + C_{metal} * T_{length} \qquad (1)$$

Where, N_{rows} indicates the number of CAM rows and T_{length} indicates the broadcast tagline length.

In CAM+RAM design, larger amount of power is consumed due its capacitance. The effective size of the Load/Store Queue is reduced dynamically in order to reduce the power dissipation. It detects whether the bottom bit line segment is empty or not in order to downsize

the Queue [29, 30]. Based on bit indication, it automatically reduces the size and the power consumption in the CAM+RAM memory cell as well as overall processor.
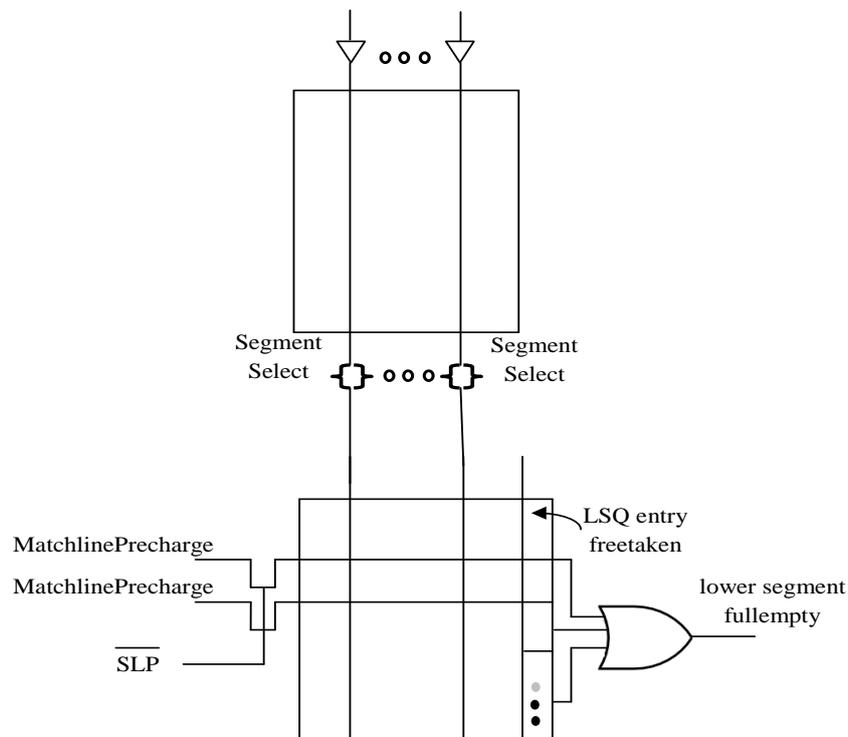

Figure 1. Structure of Existing CAM+RAM Implementation [28]

4. Proposed Instruction Queue

To reduce both the static and dynamic power dissipation in the CAM/RAM structure, gated-power technique is proposed and is shown in Fig.2. It enhances the robustness of the design against process variation. A priority encoder with CAM+RAM memory structure is used for getting input signals from the input request terminals. It provides an N-bit output code word to indicate the identification number for the highest-priority level from the active input demanding terminal.

When any one of the first group of input demanding terminal is identified as a logical TRUE value in the specified bit position of the N-bit output code word, it is active. Another input demanding terminal, should have a logical FALSE value for the specific bit position as disabled. The proposed new design of CAM+RAM with priority encoder in the conventional CAM+SRAM identifies memory order violations, consistency and inconsistent violations and misprediction. It reduces delay and power consumption, compared to the conventional CAM+RAM [27, 28]. The voltage on the Tagline is charged up to 0.7V which is very less than the corresponding VDD. Hence the new proposed architecture of CAM+RAM with priority

encoder provides both low-power and high-speed operation in the out-of-order superscalar processor.
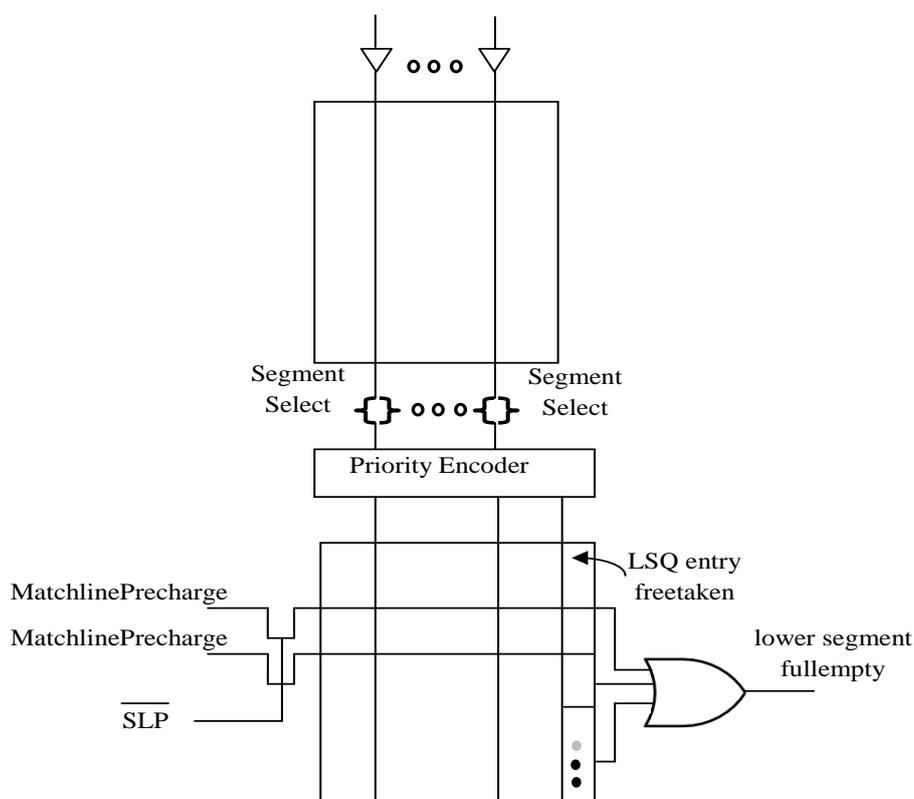

**Figure 2.** Circuit Implementation of proposed CAM+RAM

5.   Simulation and Results

Conventional design of CAM+RAM and the new proposed design CAM+RAM with priority encoder are designed using Tanner EDA tool. Precharge is kept as high, during the testing. Conventional designs of CAM+RAM and the proposed CAM+RAM with priority encoder output functional waveforms, are shown in Fig.3 and Fig.4 respectively. The performance of the new proposed design of CAM+RAM is compared with the conventional design based on 65nm, 45nm and 32nm technologies. The power reduction achieved by the proposed design compared with the conventional design in all the above technologies, indicates technology independency. Insertion of Priority Encoder in the design of CAM+RAM which reduces the associative search tag time for loading, memory order violations and broadcasting of data, also reduces the average power consumption.

The Power reduction is observed for the proposed design at 45nm fabrication technology when compared with conventional CAM+RAM.
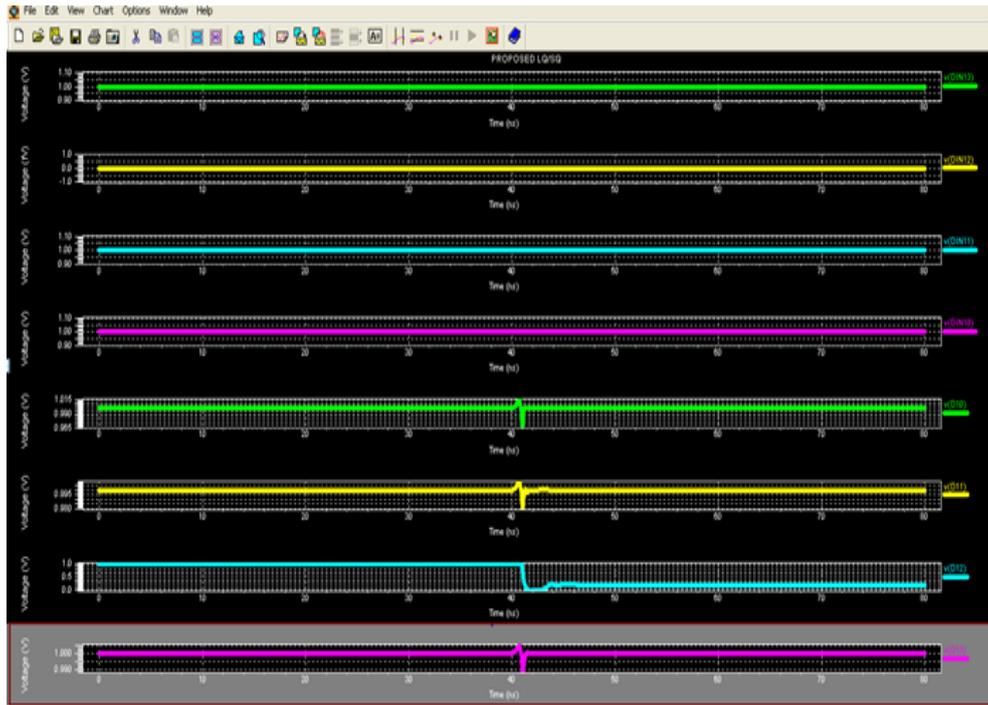
**Figure 3.** Output Waveform of the Conventional CAM+RAM
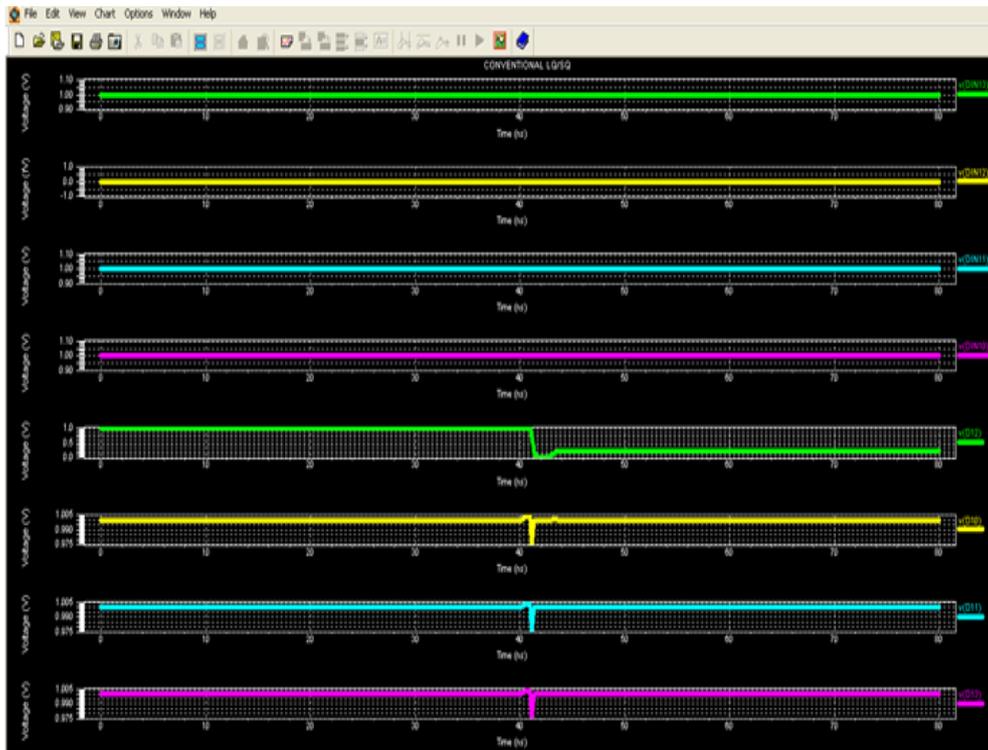
Figure 4. Output Waveform of the proposed CAM+RAM

The average power consumption observed from the CAM+RAM design for various transistor channel length sizing 65nm, 45nm and 32 are shown in Fig.5, Fig.6 and Fig.7 respectively. These results show that CAM+RAM with priority encoder is adaptable for all the
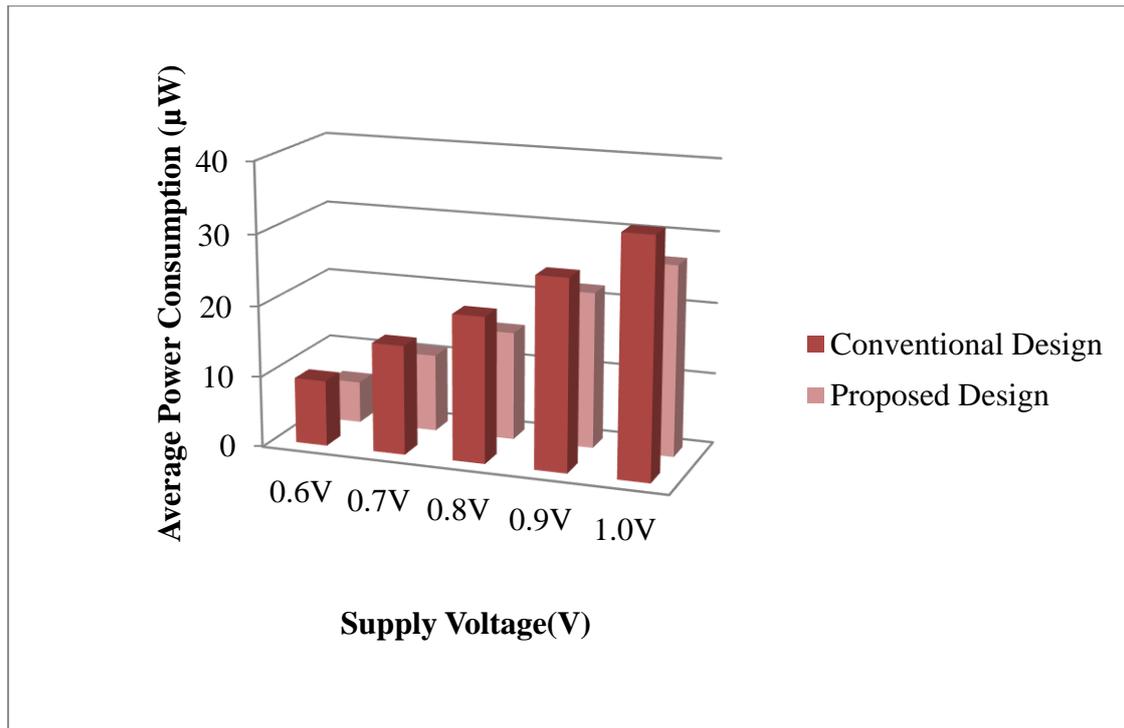
above technologies and especially more suitable to implement in 45nm technology for higher power reduction while compared to other technologies.
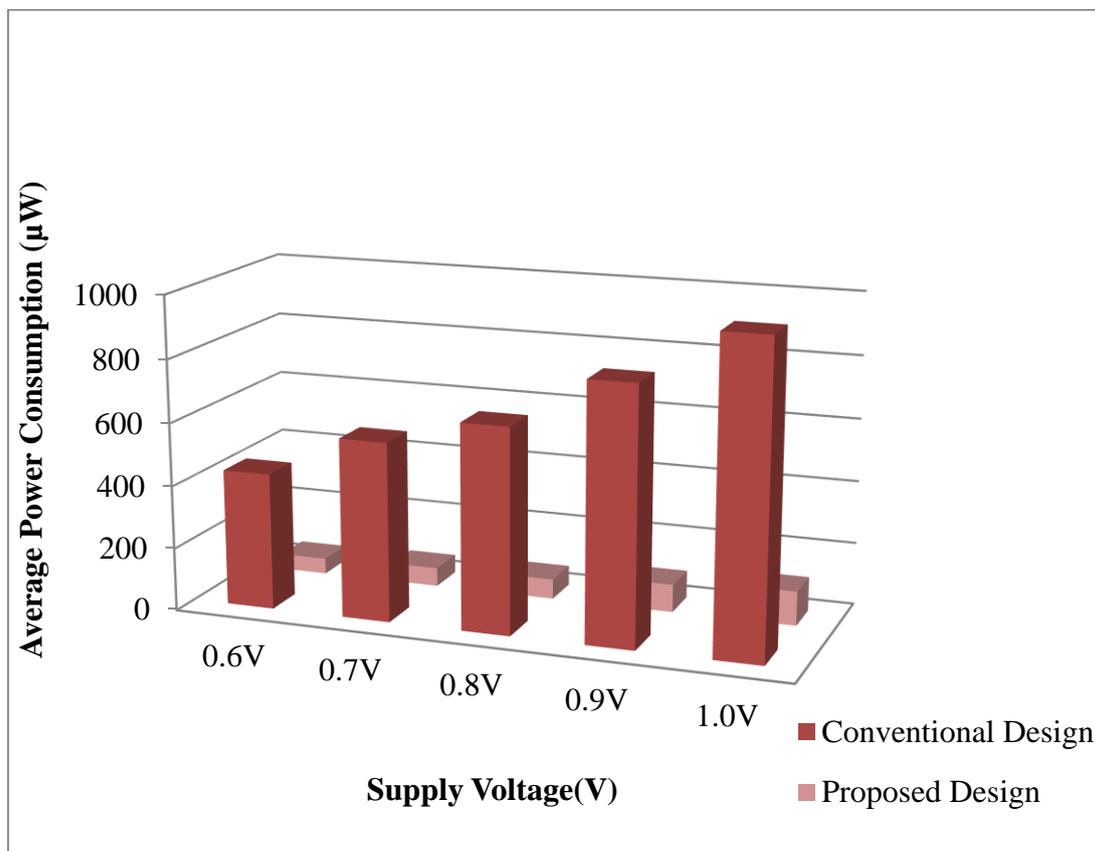
Figure 5. Power Analysis of CAM+RAM (65nm)
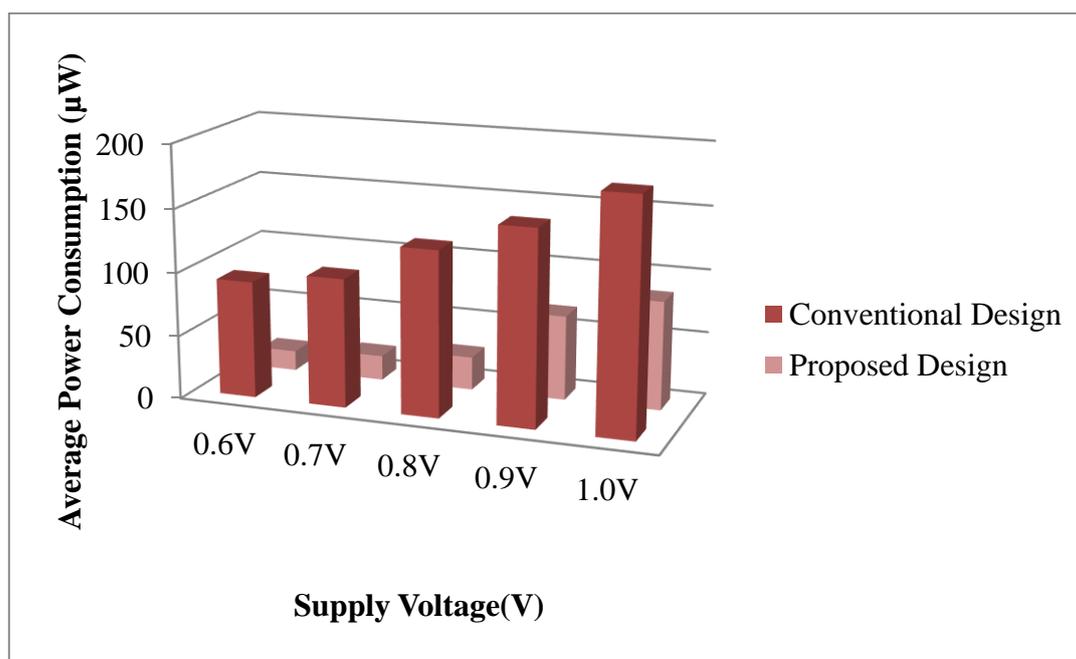
Figure 6. Power Analysis of CAM+RAM (45nm)

Figure Error! No text of specified style in document.**.** Power Analysis of CAM+RAM
(32nm)

6. Conclusion

The implementation of proposed CAM+RAM for improving the performance of the superscalar processor is explained. The priority encoder incorporated with the logic structure of load store queue for achieving further power reduction shows improvement in average power consumption compared to the existing CAM+RAM. The power analysis of the proposed structure in 65nm, 45nm and 32nm technology shows the power reduction and technology independency. This power gating technique with priority encoder can be used with any other orthogonal techniques to achieve power reduction of overall processor design which is especially designed for next generation networks.

References

[1]   Bansal, Malti, Harmandeep Singh, and Gaurav Sharma. "A Taxonomical Review of Multiplexer Designs for Electronic Circuits & Devices." Journal of Electronics 3, no. 02 (2021): 77-88.

[2]   Vaithiyanathan, D., Vikrant Gupta, Santosh Kumar, Alok Kumar Mishra, and J. Britto Pari. "Performance Analysis of Implicit Pulsed and Low-Glitch Power-Efficient Double-Edge-Triggered Flip-Flops Using C-Elements." In International Conference on

Communication, Computing and Electronics Systems, pp. 261-275. Springer, Singapore, 2020.

[3]   Chen, Joy Iong Zong, and Lu-Tsou Yeh. "Graphene based Web Framework for Energy Efficient IoT Applications." Journal of Information Technology 3, no. 01 (2021): 18-28.

[4]   Kumar, N. Mohan. "Energy and power efficient system on chip with nanosheet fet." Journal of Electronics 1, no. 01 (2019): 52-59.

[5]   Shahul, Nooriya, and Siddharth Shelly. "Bidirectional Battery Charger for Electric Vehicle." In Proceedings of International Conference on Sustainable Expert Systems: ICSES 2020, vol. 176, p. 185. Springer Nature, 2021.

[6]   Mugunthan, S. R. "Wireless rechargeable sensor network fault modeling and stability analysis." Journal of Soft Computing Paradigm (JSCP) 3, no. 01 (2021): 47-54.

[7]   Shrestha, Sujan, and Subarna Shakya. "Technical analysis of zigbee wireless communication." Journal of trends in Computer Science and Smart technology (TCSST) 2, no. 04 (2020): 197-203.

[8]   Kunaraj, A., J. Joy Mathavan, and K. G. D. R. Jayasekara. "Hardware Implementation of Automatic Power Source Controller Cum Power Optimizer." (2021).

[9]   Smys, S., Abul Basar, and Haoxiang Wang. "Artificial neural network based power management for smart street lighting systems." Journal of Artificial Intelligence 2, no. 01 (2020): 42-52.

[10]   Kumar, N. Mohan. "Energy and power efficient system on chip with nanosheet fet." Journal of Electronics 1, no. 01 (2019): 52-59.

[11]   Joy, Vijo M., and S. Krishnakumar. "Efficient Load Scheduling Algorithm Using Artificial Neural Network in an Isolated Power System." In International Conference on Inventive Computation Technologies, pp. 615-621. Springer, Cham, 2019.

[12]   Suma, V., and Shavige Malleshwara Hills. "Resource Intensification for Mobile Devices Using the Approximate Computing Entities." Journal of trends in Computer Science and Smart technology (TCSST) 2, no. 1 (2020): 26-36.

[13]   Kavinkartik, E., S. Suseendiran, and N. Mohankumar. "Design and Randomness Evaluation of Mixed-Signal TRNG for Low Power Applications." In International Conference on Intelligent Computing, Information and Control Systems, pp. 105-113. Springer, Cham, 2019.

[14]   Chen, Joy Iong Zong, and Kong-Long Lai. "Machine learning based energy management at Internet of Things network nodes." Journal: Journal of Trends in Computer Science and Smart Technology September 2020, no. 3 (2020): 127-133.

[15] Karuppusamy, P. "Design and analysis of low-power. high-speed baugh wooley multiplier." Journal of Electronics 1, no. 02 (2019): 60-70.

[16] Park, I, Powell, MD & Vijaykumar, TN 2002b, 'Reducing register ports for higher speed and lower energy', Proceedings of the Annual International Symposium on Microarchitecture, MICRO, vol. 2002-January, pp. 171-182

[17] Abella, Jaume, and Antonio González. "SAMIE-LSQ: set-associative multiple-instruction entry load/store queue." In Proceedings 20th IEEE International Parallel & Distributed Processing Symposium, pp. 10-pp. IEEE, 2006.

[18] Sha, Tingting, Milo MK Martin, and Amir Roth. "Nosq: Store-load communication without a store queue." In 2006 39th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'06), pp. 285-296. IEEE, 2006.

[19] Tsai, Yi-Ying, Chia-Jung Hsu, and Chung-Ho Chen. "Power-efficient and scalable load/store queue design via address compression." In Proceedings of the 2008 ACM symposium on Applied computing, pp. 1523-1527. 2008.

[20] Castro, Fernando, Daniel Chaver, Luis Pinuel, Manuel Prieto, Michael C. Huang, and Francisco Tirado. "A power-efficient and scalable load-store queue design." In International Workshop on Power and Timing Modeling, Optimization and Simulation, pp. 1-9. Springer, Berlin, Heidelberg, 2005.

[21] Mesa-Martinez, FJ, Nayfach-Battilana, J & Renau, J 2007, 'Power model validation through thermal measurements', ACM SIGARCH Computer Architecture News, vol. 35, no. 2, pp. 302-302.

[22] Michaud, P & Seznec, A 2001, 'Data-flow prescheduling for large instruction windows in out-of-order processors', in The Seventh International Symposium on In High-Performance Computer Architecture, 2001. HPCA. , pp. 27-36.

[23] Moreshet, T & Bahar, RI 2002, 'Complexity-effective issue queue design under load-hit speculation', in In Proceedings of the Workshop Complexity-Effective Design.

[24] Moreshet, T & Bahar, RI 2003, 'Power-Aware Issue Queue Design for Speculative Instructions', pp. 634-637.

[25] Moreshet, T & Bahar, RI 2004, 'Effects of speculation on performance and issue queue design', IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 12, no. 10, pp. 1123-1126.

[26] Moshovos, a, Breach, SE, Vijaykumar, TN & Sohi, GS 1997, 'Dynamic Speculation And Synchronization Of Data Dependence', Conference Proceedings. The 24th Annual International Symposium on Computer Architecture.

[27] Houman Homayoun , "Reducing power in all major CAM and SRAM based processor units via centralized, dynamic resource size management"IEEE transactions on Very large scale integration system,vol.19,no.11,nov 2011.

[28] Dhanalakshmi, G., M. Sundarambal, and K. Muralidharan. "Diminution of power in load/store queue for CAM and SRAM-based out-of-order processors." International Journal of Advanced Intelligence Paradigms 15, no. 1 (2020): 89-97.

[29] P John, T & G.Dhanalakshmi, M 2014, 'Reducing Power in Processor Unit via Centralized Dynamic Resource Size Management', IOSR journal of VLSI and Signal Processing, vol. 4, no. 1, pp. 38-43.

[30] Pagiamtzis, K & Sheikholeslami, A 2006, 'Content-addressable memory (CAM) circuits and architectures: A tutorial and survey', IEEE Journal of Solid-State Circuits, vol. 41, no. 3, pp. 712-727.

Author's biography

K. Muralidharan works in the Department of Electronics and Communication Engineering at Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India. His area of research includes Low power VLSI design, Wireless Networks, Internet of Things, Computer Networks, Image processing, Mobile Communication and Cryptography.

S. Uma Maheswari works in the Department of Electronics and Communication Engineering Coimbatore Institute of Technology, Coimbatore. Tamil Nadu, India. Her area of research includes VLSI Design, Machine learning, Artificial Intelligence, Speech Recognition, Digital Signal Processing, Signals and Systems, Digital System Design, IoT, Microprocessors and Microcontrollers.