DESIGN & IMPLEMENTATION OF SELF TIME DUMMY REPLICA TECHNIQUE IN 128X128 LOW VOLTAGE SRAM

[1]Mitali Agarwal, [2]Taru Tevatia
[1]Research Scholar, [2] Associate Professor
[1] Department of Electronics & Communication Engineering
[1]Al Falah University, Dhauj, Faridabad

Abstract—SRAM are useful building blocks in many applications such as cache memories, data storage embedded applications, microprocessors. The paper aims to propose the design for 128x128 SRAM with self time dummy replica technique for fast read & write access and low power consumption. The main functional blocks are 6T SRAM cell, row and column decoders, precharge circuit, read/write block and sense amplifier. Access time, speed & power consumption are the three key parameters for an SRAM memory design. Self-time technique has been implemented to optimize power and access speed of SRAM. Memory timing circuits need a delay element which tracks the bit-line delay but still provide a large swing signal which can be used by the subsequent stages of the control logic. The key to building such a delay stage is to use a delay element which is a replica of the memory cell connected to the bit-line, while still providing a full swing output. This technique uses a dummy column and dummy row in the RAM to control the flow of signals through the core. The 128x128 SRAM has been designed, implemented & analyzed in standard TSMC 180nm technology library using Cadence Virtuoso tool.

Keywords—SRAM, Decoder, Precharge Circuit, Sense Amplifier, Dummy column.

I. INTRODUCTION

With the rapid growth of communication and digital systems, handheld devices and consumer electronics are becoming increasingly popular. Large SRAM arrays that are widely used as cache memory in microprocessors occupy a significant portion of silicon area. In an attempt to optimize the performance of such chips, large arrays of fast SRAM help to boost the system performance. Low power SRAM array implementation is used to demonstrate the feasibility of low power memory design. This work explores the design of SRAMs, focusing on optimizing delay and power. While process and supply scaling remain the biggest drivers of fast low power designs, this report investigates some circuit techniques which can be used in conjunction to scaling to achieve fast, low power operation. To reduce the power consumption the first technique is to reduce the active duty cycle of the memory operation using self-timed architecture. An internal clock pulse with reduced on time is generated which controls all the memory operation. Second technique of power reduction is to use multi-stage row and column decoding which reduces the power consumption as well as it also improves the timing characteristics of memory.

II. SRAM ARCHITECTURE

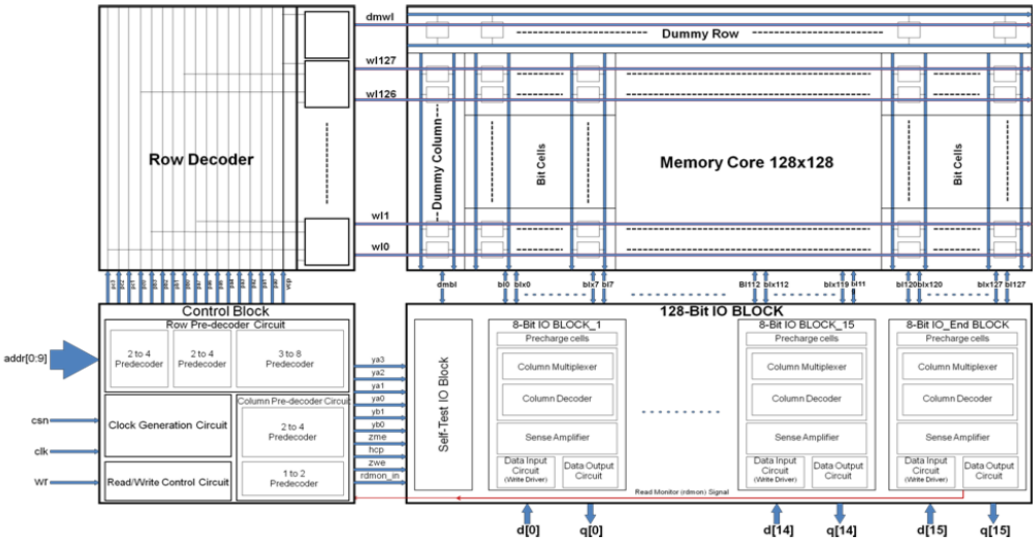

Figure 1: Block Diagram of a typical 1024x16CM8 SRAM Memory

The SRAM memory system design has been done based on an array-structured memory architecture at 180nm technology. The address signals (Addr [0] – Addr [9]) are divided into two groups. One group contains the row address bits (Addr [3] – Addr [9]). The other group contains the column address bits (Addr [0] – Addr [2]). Based on the seven row address bits, the row decoder produces 2^7 = 128 horizontal word lines. With three column address bits, the column decoder generates 8 select lines (true and complemented) for 8 bit multiplexer. This 8 bit multiplexer along with column decoder, precharge cells, sense amplifier and read/write circuitry forms 8 bit Input/output (IO8) block. There are sixteen 8 bit IO blocks connected horizontally to generate 128 vertical bit lines. The array produced by the intersections of the 128 horizontal word lines and 128 vertical bit lines is 1024x16x8 memory cell array.

III. DESIGN & IMPLEMENTATION

The main SRAM building blocks are as follows:

- SRAM Cell
- Precharge Circuit
- Write Driver
- Sense Amplifier
- Row Decoder /Column Decoder

SRAM Cell

A typical SRAM cell is made up of six MOSFETs. Each bit in an SRAM is stored on four transistors (M0, M1, M2 and M3) that form two cross-coupled inverters. This storage cell has two stable states which are used to denote 0 and 1. Access to the cell is enabled by the word line (WL in figure) which controls the two access transistors M_5 and M_6 which, in turn, control whether the cell should be connected to the bit lines: BL and BL'. They are used to transfer data for both read and write operations.

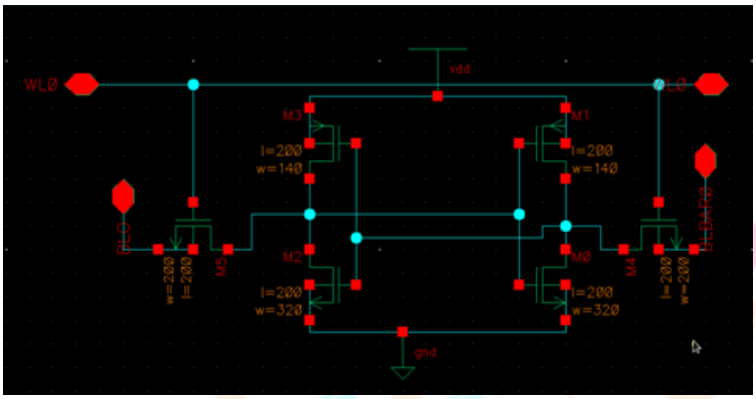

Figure 2: Six Transistor CMOS SRAM bit cell

Precharge Circuit

It is one of the essential components of SRAM. It forces each bit line to VDD and also equalizes their potentials. In both read & write operations, bitlines are initially pulled up to high voltage to VDD.

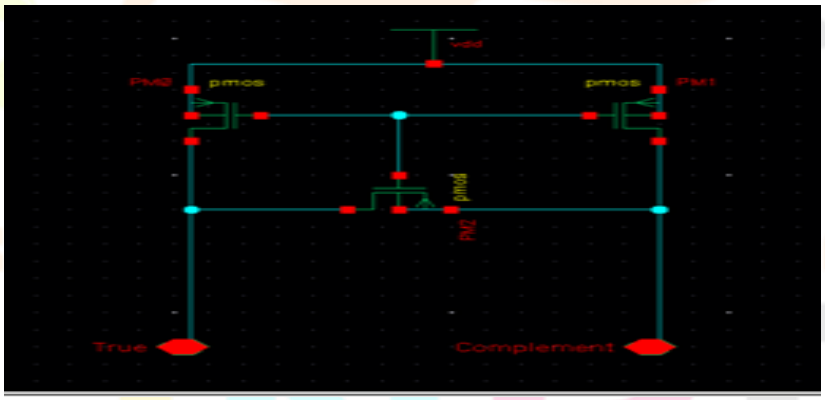

Figure 3: Schematic of Precharge Circuit

Write Driver

The circuit that writes the data into the cells is called a write driver. Generally the bitline is discharged all the way to ground during a write operation. This huge bitline swing can cause large power consumption during writes. During read operation, the bitline voltage swing is generally restricted to around 200mV & thus the writes can on an average consume about 1/8th more power than a read operation.

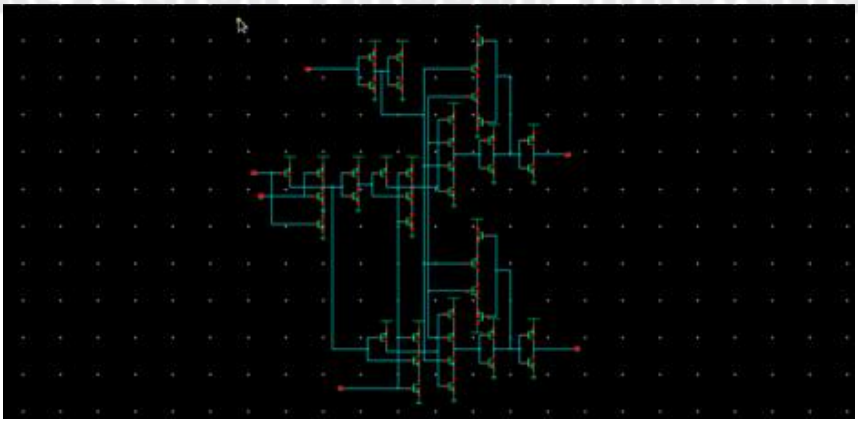

Figure 4: Schematic of Write Driver Circuit

Sense Amplifier

Sense Amplifier (SA) is the most critical circuits in the periphery of CMOS memory. The performance of sense amplifiers strongly affects both memory access time and overall memory power dissipation. The choice and design of a SA defines the robustness of bitline sensing, impacting the read speed and power. The primary function of SA in SRAM is to amplify a small analog differential voltage developed on bitlines by a read accessed cell to the full swing digital output signal thus greatly reducing the time required for a read operation.

Figure 5 shows the schematic of the full complementary positive feedback voltage sense amplifier. It is also known as latch type sense amplifier. They consist of a pair of cross coupled gain stages which are turned on with the aid of a sense enable sae, when an adequate input differential is set up. The positive feedback in the latch leads to a full amplification of the input signal to a full digital level. While this type consumes the least amount of power due to the absence of any biasing power, they could potentially be slower since some timing margin is needed in the generation of the sense enable signal.

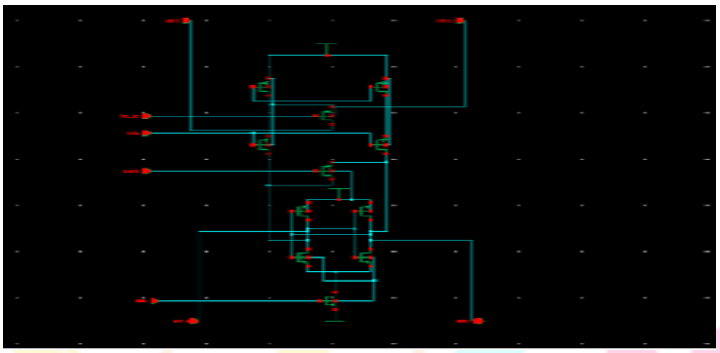

Figure 5: Schematic of Sense Amplifier

Row Decoder

An address arrives at an SRAM boundary to identify which address is being selected. Since the RA in SRAM stands for random access any location may be selected on any given cycle. The logical function of the decoder is equivalent to 2^n n-input AND gates, where the large fan-in AND operation is implemented in a hierarchical structure. The schematic of a two-level 7 to 128 decoder is shown in Figure 6. The first level is the predecoder where three groups of three (addr [3:5]), two (addr [6:7] and two addr [8:9]) address inputs are first decoded to produce three groups of predecoder outputs pa0-pa7, pb0-pb3 and pc0-pc3. The predecoder outputs are combined at the next level with vertical clock pulse (vcp) to activate the word line. The vertical clock pulse is generated by control circuit using global clock signal clk.

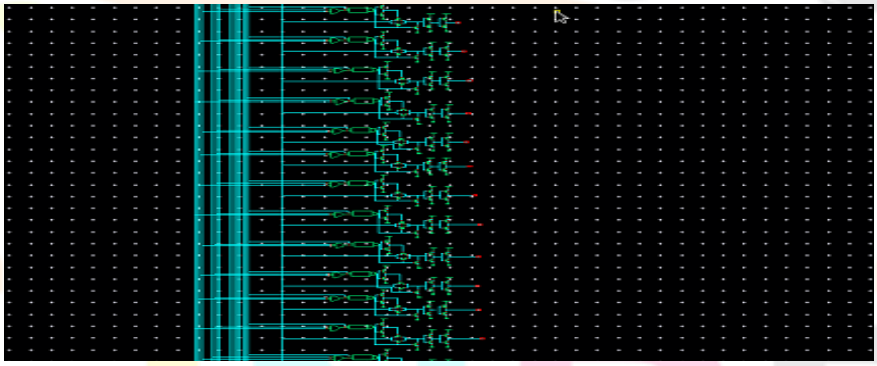

Figure 6: Schematic of Row Decoder

Column Decoder

Figure 7 shows the two level 3 to 8 column decoder. In the first level, the three column address lines addr [0:3] are decoded with two predecoder to generate six pre-decoded signals. The addr [0:1] are decoded with 2 to 4 predecoder to generate 4 pre-decoded signals ya0, ya1, ya2 and ya3. The addr [2] is decoded with 1 to 2 predecoder to generate 2 predecoder signals yb0 and yb1. In the next level, outputs of two pre-decoders are multiplied to produce eight final decoded outputs. These decoder output signals are used as column select signals. The column select signals select one of the eight columns in an 8-bit IO block.

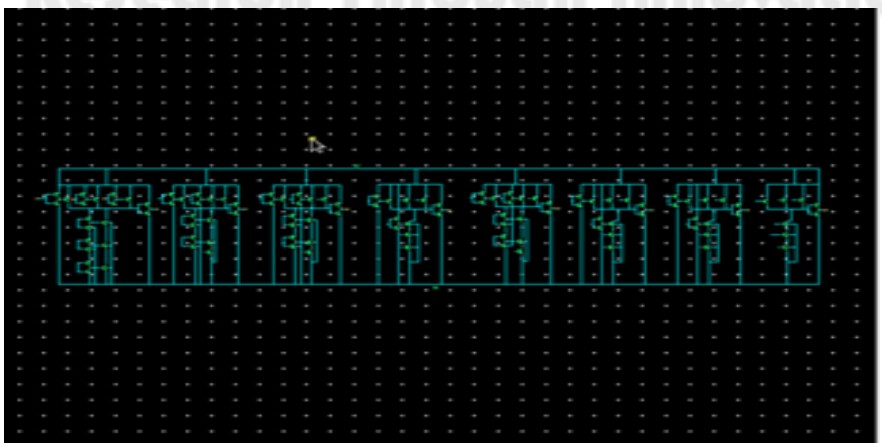

Figure 7: Schematic of Column Decoder

IV. SELF TIMING THE SRAM CORE

The technique for achieving this uses a dummy column in the RAM to time the flow of signals through the core. A dummy column is an additional column of bit cells and self timed IO block placed at the side farthest from the word drivers. Bit cells in the dummy column are forced to a known state by shorting one of the internal nodes to a given voltage. Figure 8 shows the schematic of the self time IO block.

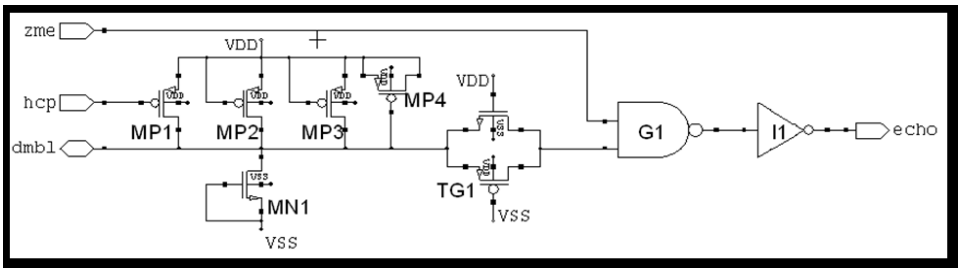

Figure 8: Self Time IO Block

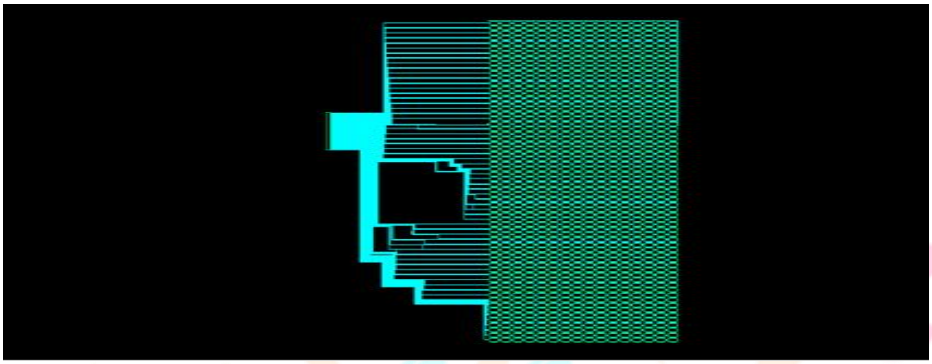

Figure 9: 128x128 SRAM Memory Core

V. SIMULATION RESULTS

Figure 10 shows the simulation result, which illustrates the working of self time IO block.

When hcp is low the dummy bitline dmbl is connected to the power supply (VDD). The dummy bitline dmbl is connected to one input of the NAND gate G1 followed by an inverter I1. The other input of the G1 is connected to memory enable signal which is high when the chip is selected. Hence we will have a high echo (reset) signal. If a rising edge of the hcp occurs, the dmbl will get discharged through the capacitors form of MN1 and we will have a low echo signal. This low echo signal resets the flip-flop circuit in control block and kills the corresponding wordline.

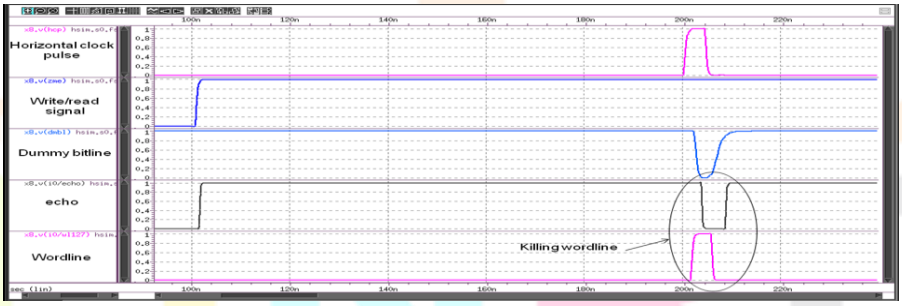

Figure 10: Simulation of Self Time Operation

Figure 11 shows the simulation result of the sense amplifier. Once a wordline and a column select signals asserted, one of the bitline start discharging. When the sufficient difference is made between the two bitlines the saefb signal separate input and output of the sense amplifier and then the sense enable signal sae is asserted. After the assertion sae the latch action will take place and we will have full complementary outputs at the sense amplifier outputs sao and saox. The minimum sensing voltage measured is 180mv and the sensing delay is around 2.5nS.

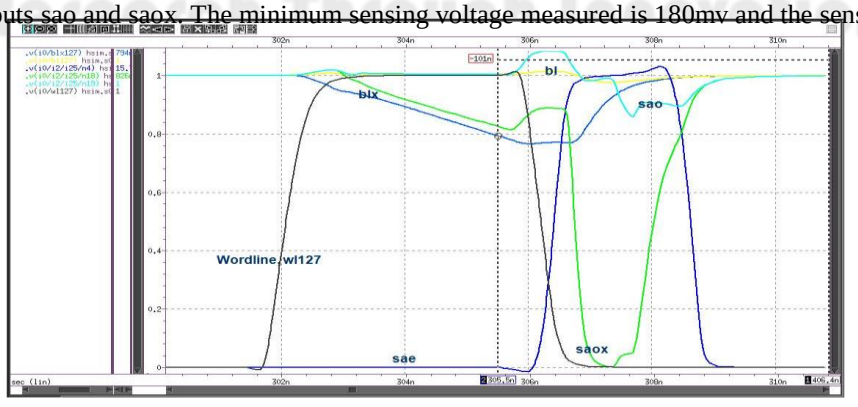

Figure 11: Simulation of Sense Amplifier

Figure 12 shows the simulation result which illustrates the complete write/read operation. The IO block is tested with complete memory

chip.

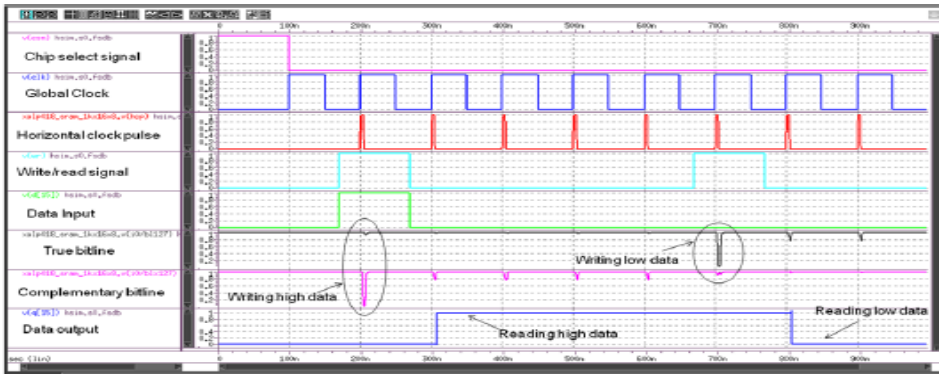

Figure 12: Complete Write / Read operation

The memory access time is the time between 0.5VDD of global clock to the 0.5VDD data output q[i]. The access time measured for different loads and clock slope. The load capacitance is varied from 0 to 1pF for a particular global clock slope.

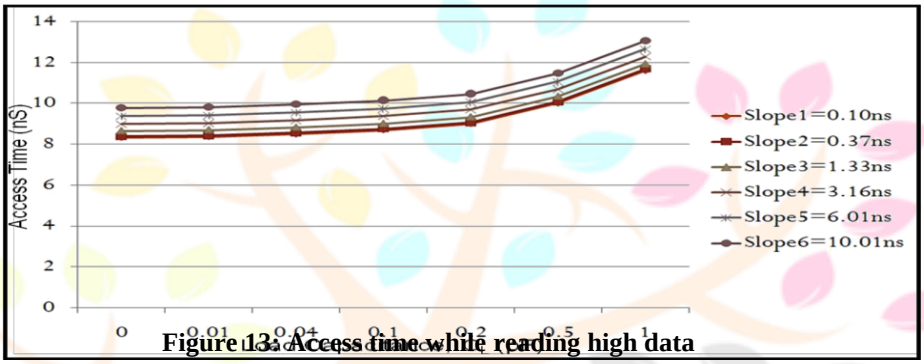

Figure 13: Access time while reading high data

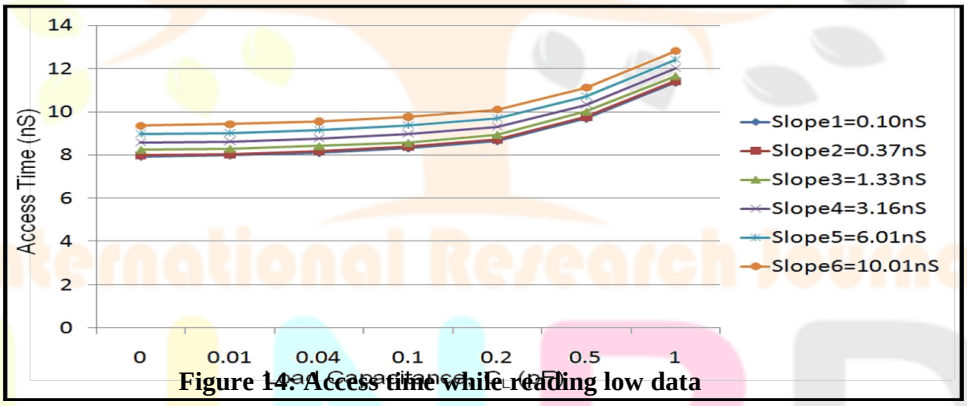

Figure 14: Access time while reading low data

VI. CONCLUSION

The schematic of 128x128 SRAM memory using self time architecture is designed and implemented in TSMC 180nm technology using Cadence Virtuoso tool. The SRAM access path is split into two portions: the row decoders and the read data path. With the predecoder the total path effort becomes independent of the exact partitioning of the decode tree, which will allow the SRAM designer to choose the best memory organization, based on other considerations. The minimum sensing voltage of sense amplifier measured is 180mv and the sensing delay is around 2.5nS. A graph plotted between load capacitance and access time for different clock slope analyzes that the access time is almost independent load capacitance. The access time is slightly increasing as the clock slope is increasing.

REFERENCES

[1] P. Barnes, "A 500MHz 64b RISC CPU with 1.5Mb On-Chip Cache", 1999 IEEE International Solid State Circuits Conference, Digest of Technical Papers, pp.86-87.

[2] S. Hesley, "A 7th-Generation x86 Microprocessor", 1999 IEEE International Solid State Circuits Conference, Digest of Technical Papers, pp.92-93.

[3] Martin Margala, "Low Power SRAM Circuit Design", Department of Electrical and Computer Engineering, University of Alberta, IEEE, 1999.

[4] Ao Cnandrakasan, R. Broderson. "Minimizing Power Consumption in Digital CMOS circuits," Proceedings of the IEEE, p-498-523, April 1999.

[5] Bharadwaj S. Amrutur," Design and Analysis of Fast Low Power SRAMs", Department of Electrical Engineering, Stanford University, IEEE, August 1999.

[6] Bharadwaj S. Amrutur and Mark A. Horowitz, "Fast Low-Power Decoders for RAM", IEEE Journal of Solid-State Circuits, Vol. 36, Pp. 1506-1515, No. 10, October 2001.

[7] K. Sasaki, "A 7-Ns 140-Mw 1-Mb CMOS SRAM with Current Sense Amplifier", IEEE Journal Of Solid State Circuits, Vol. 27,

Pp. 1511-1517, No. 11, November 1992.

[8] K.W. Mai, "Low-Power SRAM Design Using Half-Swing Pulse-Mode Techniques", IEEE Journal Of Solid State Circuits, Vol. 33, Pp. 1659-1671, No. 11, November 1998.

[9] Sreerama Reddy G M and Dr. P. Chandrasekara Reddy, "Negative Word Line Scheme Based Low Power 8kb SRAM for Stand Alone Devices", European Journal Of Scientific Research, Vol. 26, No.2, Feb 2009.

[10] Sreerama Reddy G M and Dr. P. Chandrasekara Reddy, "Design And VLSI Implementation Of 8mb Low Power SRAM in 90nm", European Journal Of Scientific Research, Vol. 26, No.2, Feb 2009

[11] Neeraj Kr. Shukla, R.K.Singh, and Manisha Pattanaik (2011) "Design and analysis of a Novel Low power SRAM bit-cell

[12] structure at Deep-Sub-Micron CMOS technology for mobile multimedia applications" in International Journal of Advanced Computer Science and Applications, pg 43-46 Vol.2, No.5,2011.