128 *128 SRAM in Cooperating Ultra Low Power Technique

[1]Mitali Agarwal , [2]Taru Tevatia, [3]Nitin Sehgal
[1]Research scholar, [2]Associate Professor, [3]Design Engineer
[1,2]Department of Electronics and Communication Engineering,
Al-Falah School of Engineering & Technology,Dhauj ,Faridabad
[3]Incise Infotech Pvt Ltd , Noida

***Abstract* - The primary technique used for power reduction is self-timed architecture. 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. This section explores the self-timed technique for the SRAM. The circuit diagram of self-timed IO block .The technique for achieving this uses a "dummy column" in the RAM to control the flow of signals through the core. A dummy column is an additional column of bit-cells. Bit-cells in the dummy column are forced to a known state by shorting one of the internal nodes to a given voltage.**

***Keywords* – Ultra low power SRAM, Row Decoder, Sense amplifier, Write driver circuit.**

I. INTRODUCTION

Static random access memory (SRAM) itself is a complicated circuit system. In this paper the basic working and characteristics of an SRAM and its major blocks, namely the memory cells, the row decoder, the read/write control block and the IO block have been described. Specifically, structure of SRAM, SRAM memory cell, word line related circuits; bit line related circuits and critical path analysis are described.

Fast low power SRAMs have become a critical component of many VLSI chips. This is especially true for microprocessors, where the on-chip cache sizes are growing with each generation to bridge the increasing divergence in the speeds of the processor and the main memory. Simultaneously, power dissipation has become an important consideration both due to the increased integration and operating speeds, as well as due to the explosive growth of battery operated appliances. This thesis 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 thesis investigates some circuit techniques which can be used in conjunction to scaling to achieve fast, low power operation. Conceptually, an SRAM consists of a matrix of 2m rows by 2n columns of memory cells. Each memory cell in an SRAM contains a pair of cross coupled inverters which form a bi-stable element. These inverters are connected to a pair of bitlines through NMOS pass transistors which provide differential read and write access. An SRAM also contains some column and row circuitry to access these cells. The m + n bit of address input, which identifies the cell which is to be accessed, is split into m row address bits and n column address bits. The row decoder activates one of the 2m word lines which connect the memory cells of that row to their respective bitlines. The column decoder sets a pair of column switches which connects one of 2n bitline columns to the peripheral circuits.

II. LOW POWER TECHNIQUES

Concept of Power Reduction Using Self-Timed Memory Design

The primary technique used for power reduction is self-timed architecture. 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. This section explores the self-timed technique for the SRAM. The circuit diagram of self-timed IO block is shown in figure 2.

The technique for achieving this uses a "dummy column" in the RAM to control the flow of signals through the core. A dummy column is an additional column of bit-cells. Bit-cells in the dummy column are forced to a known state by shorting one of the internal nodes to a given voltage.

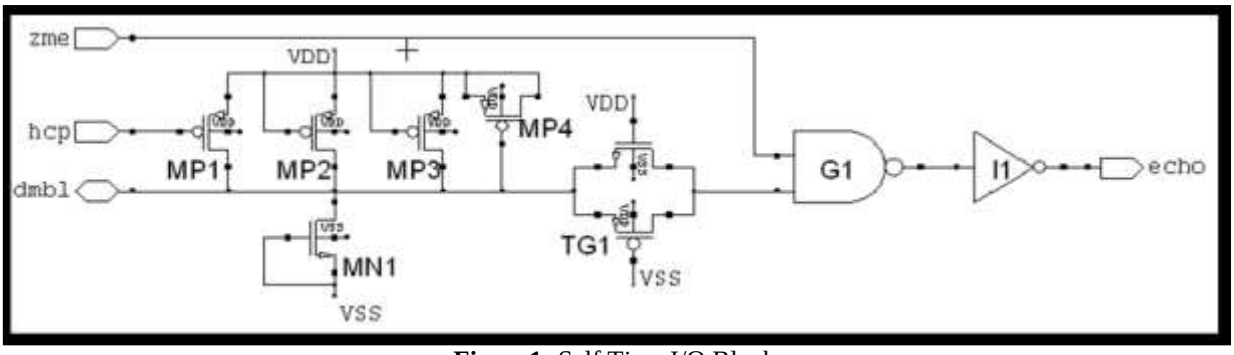

Figure1: Self Time I/O Block

When hcp is low the dummy bit line 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 it will have a high echo (reset) signal. If a rising edge of the hcp occurs, the dmbl will get discharged through the dummy row and we will have a low echo signal. This low echo signal resets the flip circuit in control block and kills the corresponding word-line shown in figure 1

Figure 2: Simulation of Self Time Operation

Sense Amplifier: Sense Amplifier 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. As with other ICs today, CMOS memories are required to have increased speed, improve capacity and maintain low power dissipation. These objectives are somewhat conflicting when it comes to sense amplifier design in memories. With increased memory capacity usually come increased bit line parasitic capacitances. This increased bit-line capacitance in turn slows down voltage sensing and makes bit-line capacitance swings energy expensive resulting in slower more energy hungry memories while decreased supply voltage lead to smaller noise margin which reduces sense amplifier reliability. Due to their great importance in memory performance sense amplifiers have become a very large class of circuits.

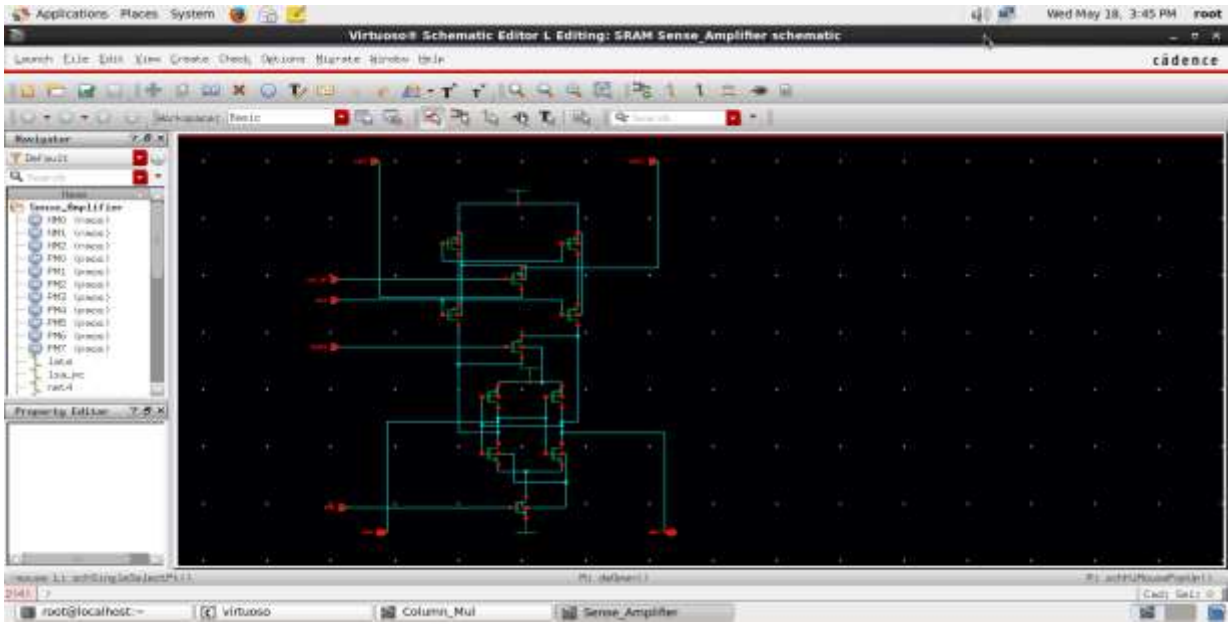

Figure : Schematic circuit diagram of Sense amplifier on cadence virtuoso

Write Driver circuit : In order to read data from an SRAM, the data must first be written into the memory. The circuit that writes the data into the cells is called a write driver or, from time reminiscent of only magnetic media, is called a write head. These terms are used interchangeably. In typical high-performance SRAMs, as has already been discussed, a pair of differential bit lines is attached to each cell and these bit lines are precharged into a high state. The cells have transfer devices that are NMOS. Thus the transfer devices drive a strong ―0‖ but do not drive a "1" very effectively. Figure shows the logic diagram of the data input or write driver circuit. The circuit consists of several NAND gates, inverters, and gated inverters. The inputs to the write driver circuit are horizontal clock pulse (*hcp*), memory enable signal (*zme: active high*), write enable signal (*zwe: writes when high*), and data input signal (*d[i]*)

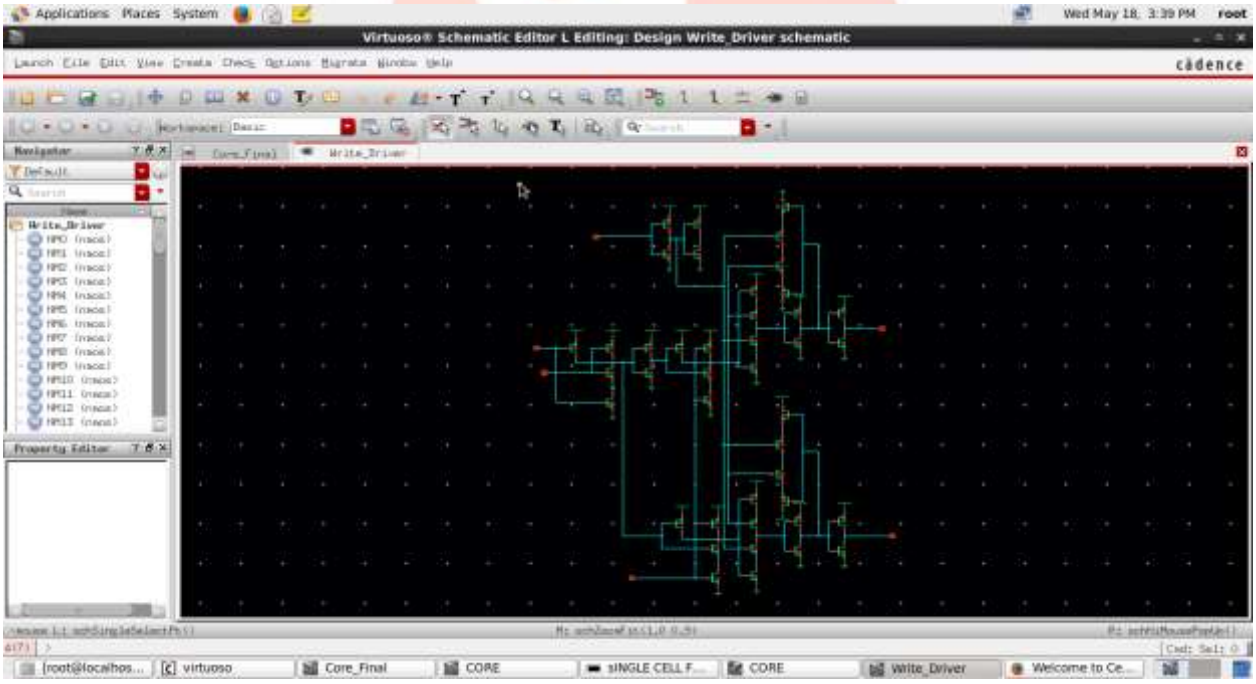

Figure : Schematic circuit diagram of write driver circuit on cadence virtuoso

III. CONCLUSION

The key to high speed in the SRAM data path is to reduce the signal swings in the high capacitance nodes like the bitlines and the data lines. Clocked voltage sense amplifiers are essential for obtaining low sensing power, and accurate generation of their sense clock is required for high speed operation. An energy efficient way to obtain low voltage swings in the bitlines is to limit the word line pulse width, by controlling the pulse width of the block select signal. The pulse widths are regulated by the aid of a replica delay element which consists of a replica memory cell and a replica bitline and hence tracks the delay of the memory cell over a wide range of process and operating conditions. Two variations of the technique is discussed .This is a very simple and robust technique with very low area overhead and is easily applicable to a wide variety of SRAM designs. This method can also be used to design a low swing data line operation for both the read and write accesses, by pulsing the control signal which gates the drivers of these lines. The pulse width is controlled by discharging a replica line which has the same capacitance as the worst case data line. This replica line is also used to trigger the amplifiers at the receiving end of these lines. We finally presented a technique to achieve low bitline write power, by using small voltage swings on the bitlines during the write operation. The memory cell structure is modified such that the cells can be used as latch type sense amplifier to amplify and store the small swing write data presented on the bitlines..

IV. REFERENCES

[1] Incise infotech pvt .ltd . (All the confidential specification provided by the company)
[2] RF Silicon Technology Pvt. Ltd
[3] A. P. Chandrakasan, et. al., ―Low-Power CMOS Digital Design‖, *IEEE Journal* .
[4] Janm Rabaey , Anantha Chandrakasan , Borivoje Nikolic . ―Digital Integrated Circuits‖,
[5] Tegze P.Haraszti, ―CMOS memory circuits", Kluwer Academic Publishers‖, 2000
[6] Vishwani D. Agrawal. CMOS SRAM Circuitry Design and Parametric Test in Nano-Scaled .
[7] Yuh-Kuang, Tseng *Industrial Research and Technology Institute*, Chapter 49.
[8] Meixner and J. Banik. Weak write test model: An SRAM cell stability design .
[9] James B Kuo, Jea-Hong Lou, ―Low-Voltage CMOS VLSI Circuits‖, pg 235-343
[10] K. Itoh, *VLSI Memory Chip Design*. Springer-Verlag, 2001
[11] A. Agarwal, B. Paul, S. Mukhopadhyay, and K. Roy, \Process variation in embedded.
[12] C.-T. Chuang, S. Mukhopadhyay, J.-J. Kim, K. Kim, R. Rao, "High-performance SRAM in nanoscale CMOS: Design challenges and techniques," in IEEE Int. Workshop., Memory.