Fall 1-31-1998

Design, fabrication and characterization of the infrastructure for a two part non volatile RAM

Suma Madapur
New Jersey Institute of Technology

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

DESIGN, FABRICATION AND CHARACTERIZATION OF THE INFRASTRUCTURE FOR A TWO PART NON VOLATILE RAM

by
Suma Madapur

The present trend is to be able to develop cheaper and more convenient methods to design and fabricate non volatile RAM that will lead to an alteration of the computer architecture itself. This will help replace the DRAM, which the computers currently have in their memory. The proposed NVRAM is designed to overcome the common difficulties faced with the use of ferroelectric materials used in the fabrication of NVRAM. The new idea essentially consists of fabricating the circuitry component separately and the substrate component separately and binding the two with the help of a conducting material. Electrical properties such as contact resistance and elasticity of the conducting material are very important. The design, simulation and fabrication of the circuitry part were carried out which would help test the conducting material. The conducting materials could be conducting polymer or indium bump pads. In this work the indium bump pads were deposited on the circuitry component and patterned. A glass plate with a conducting tin oxide film was then flipped over the entire device to evaluate the effectiveness of the 128 X 128 indium bump pads. The results obtained were very promising for actual fabrication of NVRAM using this technique.

DESIGN, FABRICATION AND CHARACTERIZATION OF THE INFRASTRUCTURE FOR A TWO PART NON VOLATILE RAM

by
Suma Madapur

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Computer Engineering

Department of Electrical and Computer Engineering

January 1998

APROVAL PAGE

DESIGN, FABRICATION AND CHARACTERIZATION OF THE INFRASTRUCTURE FOR A TWO PART NON VOLATILE RAM

Suma Madapur

Dr. Durgamadhab Misra, Thesis Advisor Associate Professor, Department of Electrical and Computer Engineering, NJIT	Date
Dr. Kenneth Sohn, Committee Member Professor, Department of Electrical and Computer Engineering, NJIT	Date
Dr. Nuggehalli M. Ravindra, Committee Member Associate Professor, Department of Physics, NJIT	Date

BIOGRAPHICAL SKETCH

Author: Suma Madapur

Degree: Master of Science

Date: January 1998

Undergraduate and Graduate Education:

- Master of Science in Computer Engineering
  New Jersey Institute of Technology, Newark, New Jersey, 1998

- Bachelor of Science in Electronics and Communication,
  Sri Jayachamarajendra College of Engineering, Mysore University,
  Mysore, India 1995

Major: Computer Engineering

ACKNOWLEDGMENT

I would like to express my deep sense of gratitude to Dr. Durgamadhab Misra for his constant guidance, support and encouragement throughout the research.

My sincere thanks to Dr. Zhi Quan Chen from Princeton University without whom this research would not have been possible. It was his idea and interest which make this project a success.

I would like to take this opportunity to extend my sincere gratitude to Dr. K.S. Sohn and Dr. N.M. Ravindra for being part of my thesis evaluation committee. I would also like to thank Dr. Dentcho Ivanov and Ken O' Brien for assisting me throughout my work in the cleanroom.

I would like to thank Dr. Swain for his guidance and constructive suggestions throughout the thesis. Thanks also to Murat, Ganesh, Aldo and Wade for helping me throughout my thesis work.

Finally my sincere thanks and appreciation to my parents and family members for their constant moral support and help throughout my course of study.

This thesis is dedicated to my parents

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

1.1 Overview

The advantages realized through the use of NVRAM (Non Volatile RAM) are so tremendous that it has led to constant research and development of new and cheaper techniques to fabricate this device. One of the immediate advantages realized through the use of NVRAM is that there is no data loss on power failure and hence less need for constant disk backup. One prospective, true nonvolatile RAM utilizes ferroelectric technology. By replacing the conventional capacitor in a DRAM with a ferroelectric capacitor, a nonvolatile RAM is produced.

The conventional method for fabricating a ferroelectric NVRAM consists of fabrication of the MOS structure, formation of the ferroelectric capacitor, back-end processes such as deposition of inter-level dielectric layers, etch out contact holes, deposition of metal layer and etch out electrode patterns. High temperature annealing is usually performed after the formation of the ferroelectric capacitor to get high quality ferroelectric layer. There are however some inherent problems with this process. The optimum annealing temperature for the commonly used ferroelectric material (PZT) is above 600°C while the maximum temperature that the processed MOS structure can tolerate without degradation is 550°C. Also the back end process can be detrimental to the ferroelectric layer. To take care of these problems a more complicated procedure of annealing needs to be employed leading to very high costs of fabrication. However, if the

data recording medium which consists of the ferroelectric medium is fabricated as a separate substrate then both the problems generally encountered with conventional methods of fabrication can be avoided obtaining higher quality processing, higher yield and lower cost. The separation allows the ferroelectric material to be annealed at higher temperature and the backend process will be a simple one step annealing. In this way, a good, low cost, NVRAM can be fabricated.

1.1.1 Advantages of NVRAM

Currently computers have the DRAM or the Dynamic Random Access Memory for their main internal memory. The DRAM is volatile in nature and hence the recorded data disappears when the power is removed. So the data is transferred to magnetic disk or other nonvolatile media for storage purposes. The two-part non-volatile RAM can provide fast random accessibility, as well as nonvolatility. With the use of the NVRAM there is no need for system boot from the disk, since the computer works immediately on turn-on. Hence the hard disk can be eliminated. The elimination of hard disk has a number of advantages. It helps save space and energy. A notebook-computer with NVRAM replacing DRAM and disk drive can be very compact. Its battery will last much longer and it operates much faster. Since there is no mechanical motion, there are no moving parts. So every part can be securely fixed. Acceleration or vibration can do no harm and the computer can work reliably in violent motion as in air combat or missile launching.

The NVRAM can lead to an alteration of the computer architecture itself. Currently the computer's instruction set is stored in a Read Only Memory (ROM) section of the CPU chip. The Basic Input-Output System (BIOS) is also stored in external ROM

media. The stored ROM data cannot be altered. But the data stored in the NVRAM can be changed easily. Thus, with the NVRAM replacement of ROM, an instruction set or BIOS can be altered, according to need.

1.1.2 Existing NVRAMs

The conventional NVRAM functions as described below. The ROMS or the Read-Only Memories are used to store fixed information and have the advantage of being non-volatile. But often it is desirable to be able to modify ROM that has been programmed. The Erasable Programmable ROM is designed to meet this. The basic MOSFET cell shown in figure 1.1. Is replaced by a floating-gate transistor or stacked-gate cell. The basic structure of this cell is shown in figure 1.2.

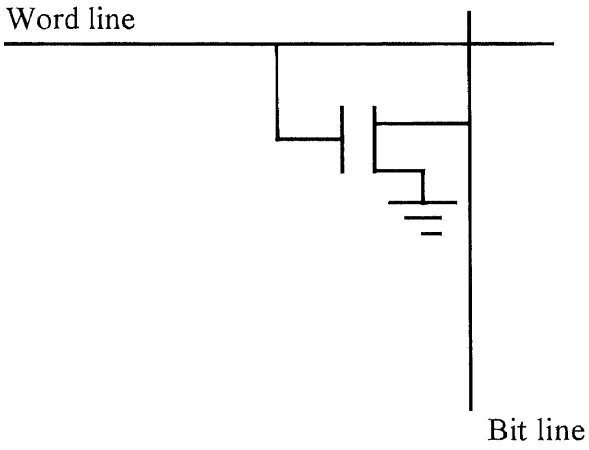

Figure 1.1 Basic MOSFET cell

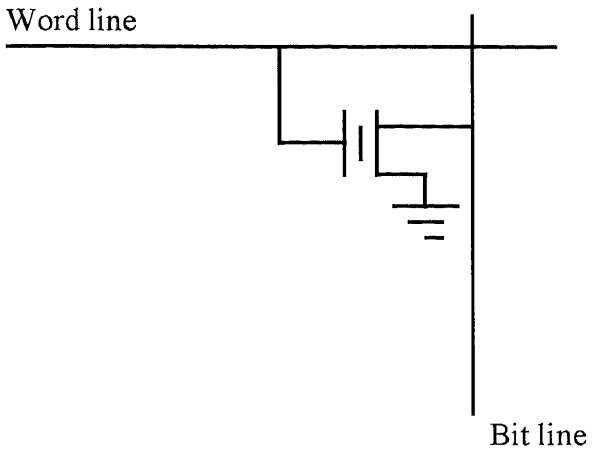

Figure 1.2 Floating gate transistor EPROM

The n-channel MOSFET has two polysilicon gates. The first, or the floating gate, is left unconnected. The second gate, or select gate, is operated to enhance the channel as in the usual MOSFET. Before being programmed, there is no charge on the floating gate and the device displays the typical enhancement MOSFET characteristics shown by curve (a) in Fig. 1.3. The threshold voltage is low in this case. A HIGH condition on the word line will turn ON the cell, and the column line will go LOW. This LOW on the column line is interpreted as a 1 at the column output selector.

To program the cell, a high voltage is applied to the drain and to the select gate through the address circuitry. A channel is established, and the MOSFET operates beyond the pinch-off. Electrons are accelerated in the high field between source and drain, and acquire enough energy to enter the conduction band of the gate oxide layer. There they are attracted by the positive potential on the select gate, and many of them lodge on the floating gate. This process is self-limiting because electrons on the floating gates begin to inhibit further electron injection.

When the programming voltage is removed, the electrons on the floating gate are trapped, and thus the floating gate constitutes a negative charge between the select gate and the substrate. This effect is that a much higher positive voltage is required on the select gate to enhance the channel than before, moving the device characteristic to (b) in Fig. 1.3. The new threshold voltage is greater than the voltage applied by the row decoder, and the transistor does not come ON when addressed. The cell is now programmed to 0.

Illuminating the cell with ultraviolet light will return it to the unprogrammed state.

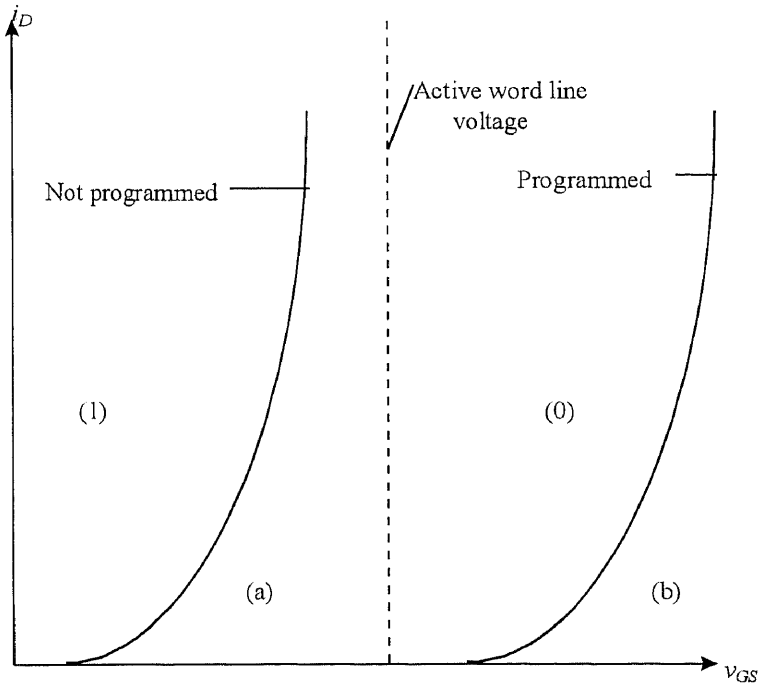

Figure 1.3 Transfer characteristic of a floating-gate transistor

1.1.2.1 Problems Associated with Existing NVRAMs

At present there are a number of nonvolatile memory devices that are being used. These are the EPROM, UV-EPROM, EEPROM, Flash memory, Battery backed SRAM, and Nonvolatile SRAM. But there are certain problems associated with most of these devices. The UV-EPROM can be erased by only removing and exposing the whole chip to ultraviolet light for 20 minutes. But UV-EPROM, EEPROM and Flash memories cannot be erased on a bit by bit basis. Also their erase speeds are very slow. Their construction is quite complicated and they are much more expensive than the DRAM. The Battery-backed SRAM requires a battery. Hence, its volatility is restricted. The size of this is very large and hence the memory cost also increases. The EEPROM and the Battery-backed SRAM are also complex in construction and hence high in cost, as well as low in bit density. The nonvolatile SRAM consists of a combination of EEPROM and SRAM and is even more complicated in construction and higher cost, as well as lower in bit density.

1.2 Objective

The perfect solution to the construction of the NVRAM would be to be able to design a ferroelectric NVRAM in which the general problems encountered during annealing and back end processing would be eliminated.

This thesis deals with the design, simulation, fabrication and testing of the infrastructure for this two part Non Volatile RAM.

This chapter deals with the description of a general idea about NVRAMs their advantages, the conventional NVRAMs as well as a detailed description of the proposed device and its various features. The second chapter deals with the design and circuit

simulation using Pspice. The third chapter describes in detail the various steps or processes used in the processing of any device. The fourth chapter deals with the actual processing done to fabricate the NVRAM and also illustrates the results obtained during process simulation using Supreme III. The fifth chapter deals with the results obtained from the test structures incorporated in the chip as well as the output characteristics obtained on testing the shift register circuit and the conducting material.

1.3 Principle of Operation

This particular NVRAM which is being fabricated utilizes the ferroelectric technology. This technique provides a radiation hardened device with the read/write speeds of fast static RAM, the density of dynamic RAM, and the nonvolatility of EEPROM. This will lead to a consolidation of the computer memory hierarchy, and produce machines, which will have increased throughput, in smaller, lighter packages.

This device involves replacing the conventional capacitor in a DRAM with a ferroelectric capacitor. The two polarization states in the ferroelectric capacitor represent the logical one or zero of the memory cell. The address selecting MOS array is the same as the DRAM. Nonvolatility is established through the remnant polarization of the ferroelectric material. It has some other advantages compared to the DRAM in addition to its nonvolatility property. It has faster access and write time than that of DRAM. Its power consumption is lower than that of DRAM, because no standing power is needed for the data retention. It is resistant to radiation, because it uses ferroelectric material. The NVRAM present which use the ferroelectric material use the following common steps during its fabrication

a) fabrication of the MOS structure

b) formation of the ferroelectric capacitor

c) back-end processes such as deposition of inter-level dielectric layers, etch out contact holes, deposition of metal layer and etch out electrode pattern.

In order to obtain high quality ferroelectric layer high temperature annealing is usually performed after the formation of the ferroelectric capacitor. This annealing is carried out in an oxidizing atmosphere to diminish the oxygen vacancies and other defects. The maximum temperature that the processed MOS structure can tolerate without serious degradation is 550°C. But the optimum annealing temperature of the commonly used ferroelectric material (PZT) is above 600°C. Also annealing in the normal fashion can be detrimental to the ferroelectric layer. Hence a more complicated annealing procedure is generally required. This in turn leads to a high cost in the fabrication of the NVRAM.

To tackle the above mentioned issues this method of constructing the NVRAM fabricates the data recording medium (which consists of the ferroelectric material) as a separate substrate. Hence we can anneal this ferroelectric material independently without harming the MOS structures and thus obtain the desired quality with higher yield and lower cost. This method eliminates any back end processing of the ferroelectric material and hence that problem is also completely taken care of.

Since we are able to anneal the ferroelectric material at a higher temperature we have a wider choice of materials. This also implies that the back end process will be a simple one step annealing. There is no need for a high cost, complicated annealing procedure. In this way, a good, low cost, NVRAM can be attained.

This also leads to higher density devices. The ferroelectric capacitor occupies far smaller area than the capacitors using ordinary dielectric materials. Since the capacitor is placed on a separate substrate, no extra area of the MOS circuitry is occupied. Techniques for fabricating high density DRAM can be adopted for fabricating the circuitry part, resulting in even higher density.

1.3.1 Design of NVRAM

The basic design is based on the construction of the DRAM, putting all the capacitors on a separate substrate and replacing the insulator of these capacitors with a ferroelectric film. The separate substrate forms the basis of the storage medium component while the original substrate forms the basis of the circuitry component. The storage medium component can be connected to the circuitry component using either the temporary connection method or the permanent connection method. In the permanent connection method, these two parts are connected permanently and packaged together after the fabrication. In the temporary connection method, the medium part will be aligned and pressed on the circuitry component during data writing or reading and can be separated from the circuitry component when not in use. In this case, the storage medium component is removable, like a floppy disk.

One of the most challenging aspects in this entire procedure is the bonding of these two components. To ensure good mechanical contact and electrical connection, the surface of the protruding conductive pads should lie in a plane. But there is bound to be some variation in height of the surface of both the circuitry component and the storage medium component of at least of a few tenths of a micrometer. If these pads were made of

a hard material then only a few of these pads would contact well with the storage medium component. To make sure that all the pads contact well, these pads should be made of soft material, so that the taller pads would yield a little at compression, allowing the shorter pads to contact well. Since in the case of the temporary connection repeated connections and disconnections are involved because the storage medium is removed and connected, this soft material should be elastic, and its elastic range should be larger than the variation of the height. Also the elastic pad should not cause a high resistance.

There are some conductive elastic materials that can meet the above requirements. One kind represents certain conjugate polymers. One kind represents certain conjugate polymers. Polyacetylene is one such example. Its resistivity can be as low as 10^{-5} Ω-cm. It is similar to an elastomer and it is fine grained. Another suitable method material is polyaniline. An elastic and conductive film can be obtained from this material or its blends, with other elastomers. Another kind of conductive elastic material is the metal particle filled elastomers. For example, silicone elastomer filled with fine-grained metal particles. The metal particle should be of submicron size, else it is difficult to obtain fine boundary for pads of micron-size. One method of obtaining such elastomer filled with fine grained metal particles is plasma polymerization and simultaneous metal vapor deposition.

After the formation of the conductive elastic material, photolithography and RIE (Reactive Ion Etching) techniques may be employed to pattern it into an array of pads.

The storage medium can be fabricated on a removable card, and can be pressed on the circuitry component only during writing/reading data. In this case, alignment is required to guarantee that a particular pad on the storage medium component presses on a same conductive pad of the circuitry component after removing the storage medium

component and reinserting. Otherwise, the address after removing and reinserting would be different from before. This is the problem of registration. Another problem that could arise would be due to any dirt particles. Cleanliness is a must and no dust particle can come between these two parts.

The elastic conductive pads can also be used for permanent connection. Some kind of clamping methods for bonding the two parts permanently by some means such as using adhesives is required.

Hence it is essential to establish the feasibility of electrically connecting two parts of the IC with myriad connection points. It is required to fabricate the elastic conductive pads, for ensuring good electrical connection. For this we have to have a method to measure the resistance of every connection point established. For this purpose, the address selecting function of the circuitry part will be used to advantage.

1.4 Objective of the Thesis

The aim of this thesis was to design and fabricate the circuitry component to test the conductivity and contact resistance of the conducting material which would be used in fabrication of the two part Non Volatile memory. For this it was necessary to generate a two dimensional array of access pads which was done using shift registers which were arranged and timed to so as to provide the desired output. The conducting material which was in the form of conducting polymer and indium bump pads would be deposited on the metal plates situated at the source of each of these access transistors. To test the conducting material a glass plate with a conducting material would be flipped over and

bonded to this circuitry component. This conducting material could be later substituted with the ferroelectric material. Mentor Graphics layout tool was used to design the circuit. Pspice was used for circuit simulation. The design was then translated into masks and the device was fabricated in the NJIT cleanroom. Supreme III was used for process simulation. The device was tested on completion of the circuitry part. HP4145B Semiconductor Parameter Analyzer was used to test the test structures to evaluate process parameters.

Since the objective of the thesis was to test the electrical characteristics of the conducting polymer as well as the indium bump pads the circuit was designed and fabricated using NMOS transistors only. Use of only NMOS transistors led to fewer mask levels being involved hence leading to lower costs and lesser time of fabrication.

CHAPTER 2

DESIGN

2.1 Introduction

The primary objective of the project was to be able to design a circuitry component in order to test the conductivity, elasticity and other properties of the conducting material used to connect the circuitry component and the substrate component of the NVRAM.

2.2 Design of Circuit

The device designed to test the elastic conductive material consisted of a memory circuit, which was made up of shift registers. The shift registers were used to generate a two-dimensional array of elastomer pads. The circuit was designed to help measure the resistance of every connection point. For this purpose the address selecting function of the circuit was such that at any point of time by choosing any word line and a bit line only one of the pads would be conducting.

Figure 2.1 shows the design of a single shift register unit that was used in this work. The output of this shift register unit was connected to the input of the next shift register unit. The shift register was composed of NMOS transistors whose W/L ratios were chosen based on the results of Pspice simulations. The entire layout was done using the IC tool in Mentor Graphics package. The IC verify tools i.e. the Design Rule Checking and IC extract were used to verify the layout and extract parasitics.

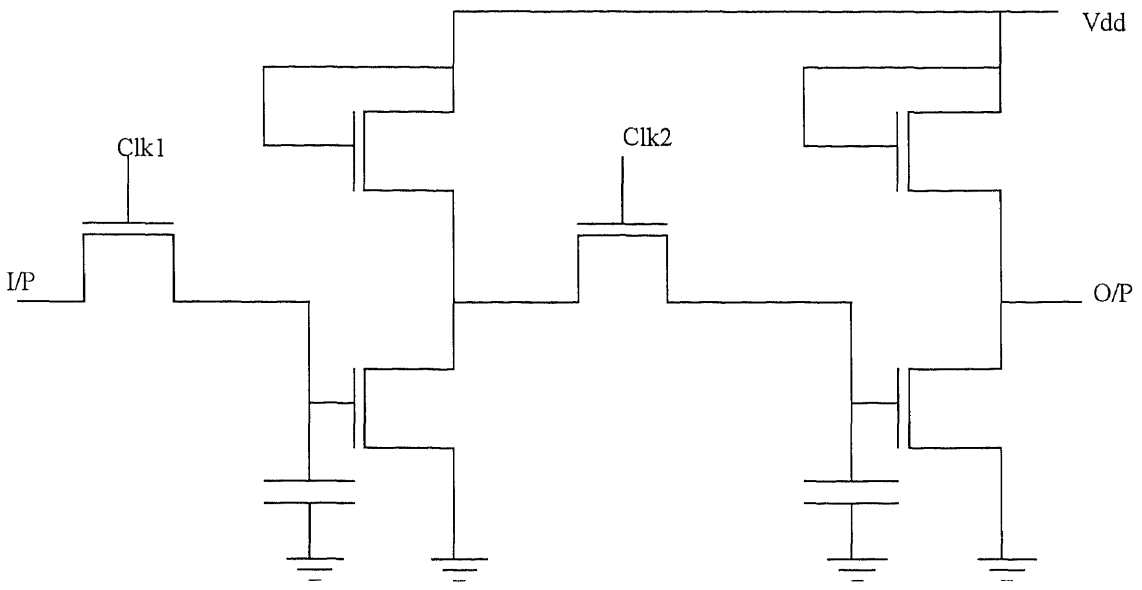

Figure 2.1 Schematic of the basic shift register unit

Figure 2.2 shows the layout representation of the shift register unit. The circuit fabrication required five mask levels, four of which were to be used to form the circuitry component of the two part nonvolatile RAM. The masks used to define the circuitry component were the active mask used to define the active area, the polysilicon mask used to define the gate and other interconnects, the contact mask used to contact metal and polysilicon to gate and active regions and the metal mask used to define all the metal patterning. The last mask was used to pattern the conducting polymer or the indium bump pads.

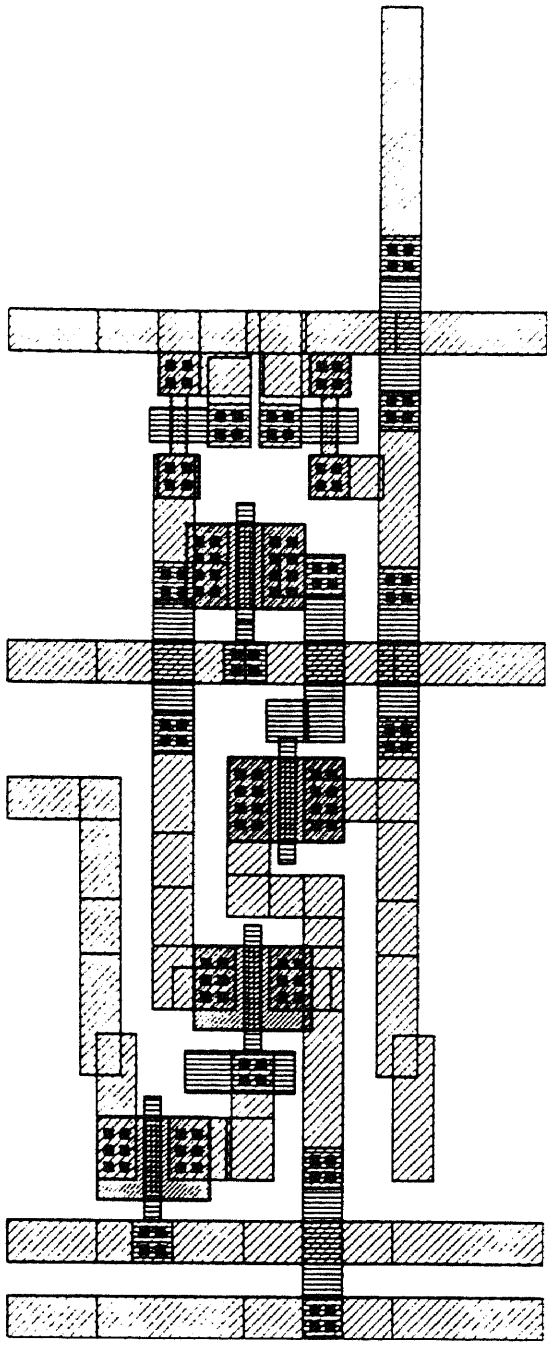

Figure 2.2 Layout of the basic shift register unit

There were two designs, which were done to test two different conducting materials, the conducting polymer and the indium bump pads.

2.2.1 Conducting Polymer

The design consisted of an array of 8 x 8 set of shift registers. This consisted of two lines of 8 shift registers each, which were placed at right angles to each other. The output of these shift registers helped form a grid matrix. At the junction of this grid at each point was present a transistor whose gate was connected to one output of the vertical set of shift register and whose drain was connected to one output of the horizontal set shift register. In the 8 x 8 set of the shift register the vertical set of the shift register was operating at 8 times the clock frequency of the horizontal set of shift register. So by the time the pulse traveled through all the shift registers in the horizontal direction the output in the vertical set of shift registers would advance by only one position upwards. This would give rise to only one transistor being conducting at any point of time. The order of conduction would be column by column moving from left to right, and after completing one row it would move upwards to the next row. The main inputs for any shift register were the ground, the Vin, the Vdd, and one clock input. The $\overline{\text{clock}}$ input was obtained by inverting this clock by passing it through an inverter. For the vertical set of shift registers identical design was used except for the clock timing. It had it's own synchronized appropriate clock period. The $\overline{\text{clock}}$ for the second clock was available through a different inveter.

There was also present a 32 x 32 set of shift register operating in the same manner as the 8 x 8 set of shift register. In the case of the 32 x 32 set of shift registers the clock of the vertical set was 32 times faster than the clock of the horizontal set.

The above described chip was basically intended to test the conducting polymer. The access transistors located at the intersection of each word line and bit line had a metal

pad at the source of the transistor. The conducting polymer was pattered and etched and a glass plate with some conducting coating flipped over this circuitry component and bonded to it to form a electric connection. The leads taken out from the bonding pads were connected to the respective inputs and the output measured to test the contact resistance of the conducting polymer. The dimensions of the NMOS transistors used was W/L = 20/6 micron and for the load transistor it was W/L = 6/12 micron. The pass transistor present at the intersection of each output line was of the dimension W/L = 20/6 micron. These values were chosen based on the result of Pspice simulation and taking into account the fabrication constraints.

The results of the Pspice simulations obtained from the Pspice simulation confirmed the working of the design although there was observed to be some voltage drop at the output as well as some delay at the output.

Figure 2.3 shows the design of the access transistors which were located at the intersection of the word line and the bit line. The word line is the output of the vertical stages of the shift register and the bit line are the outputs of the horizontal set of shift register. The conducting polymer was deposited on the metal plate located at the source of these access transistors.

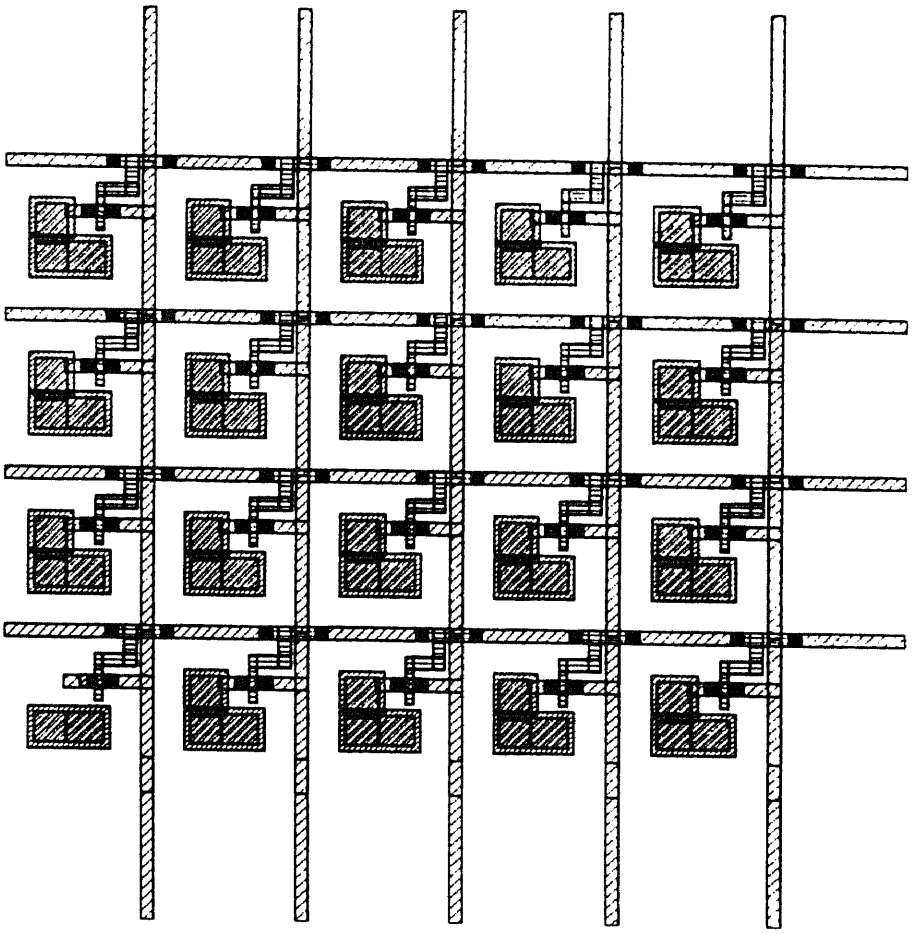

Figure 2.3 Two dimensional array of access transistors

2.2.2 Indium Bump Pads

The second design was a 128 x 128 set of shift register. The design was similar to the one described in section 2.2.1. However the dimensions of the driver transistors in this design is W/L = 20/4 and the load transistor is W/L ratio as 4/8 microns.

The indium bump pads were deposited by the lift off process on the metal pads present at the source of the access transistors. A glass plate with indium tin oxide was flipped over the circuitry component. Gold wires were bonded to the bonding pads to be accessed and these bonds were connected to gold patterns. Wire from the gold patterns were then connected to copper patterns which were used to apply the various voltages and clocks pulses. A measure of the contact resistance would determine the feasibility of the conducting material as well as prove if the bump pads made good contact with the glass plate. The results of the Pspice simulation done to determine the W/L ratios of the transistors used to form the shift register were confirmed with Pspice simulations.

Figure 2.4 shows the design of the access transistors which were located at the intersection of the word line and the bit line of the second design. The indium bump pads would be deposited on the metal plate located at the source of these access transistors.

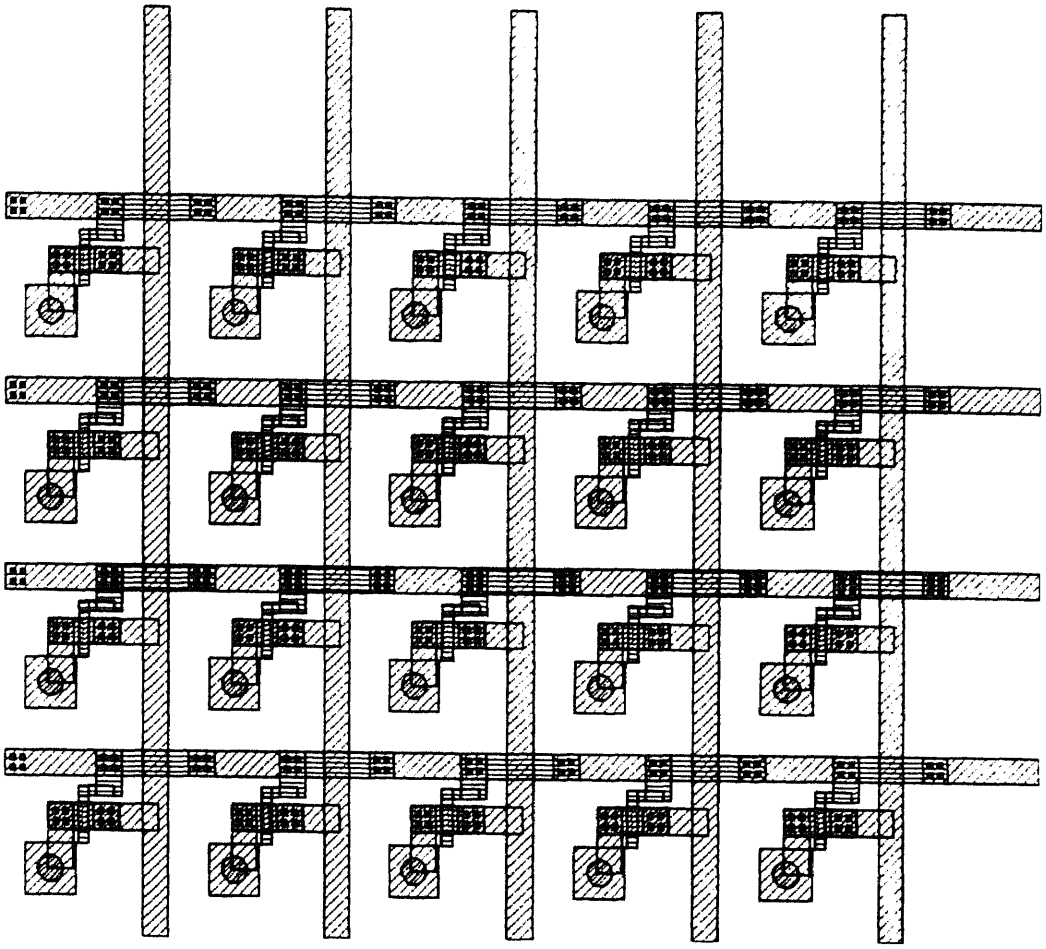

Figure 2.4 Two dimensional array of access transistors

Figure 2.5 shows the layout of the 8 x 8 set of shift register and depicts how the shift registers are arranged to generate the desired type of output.

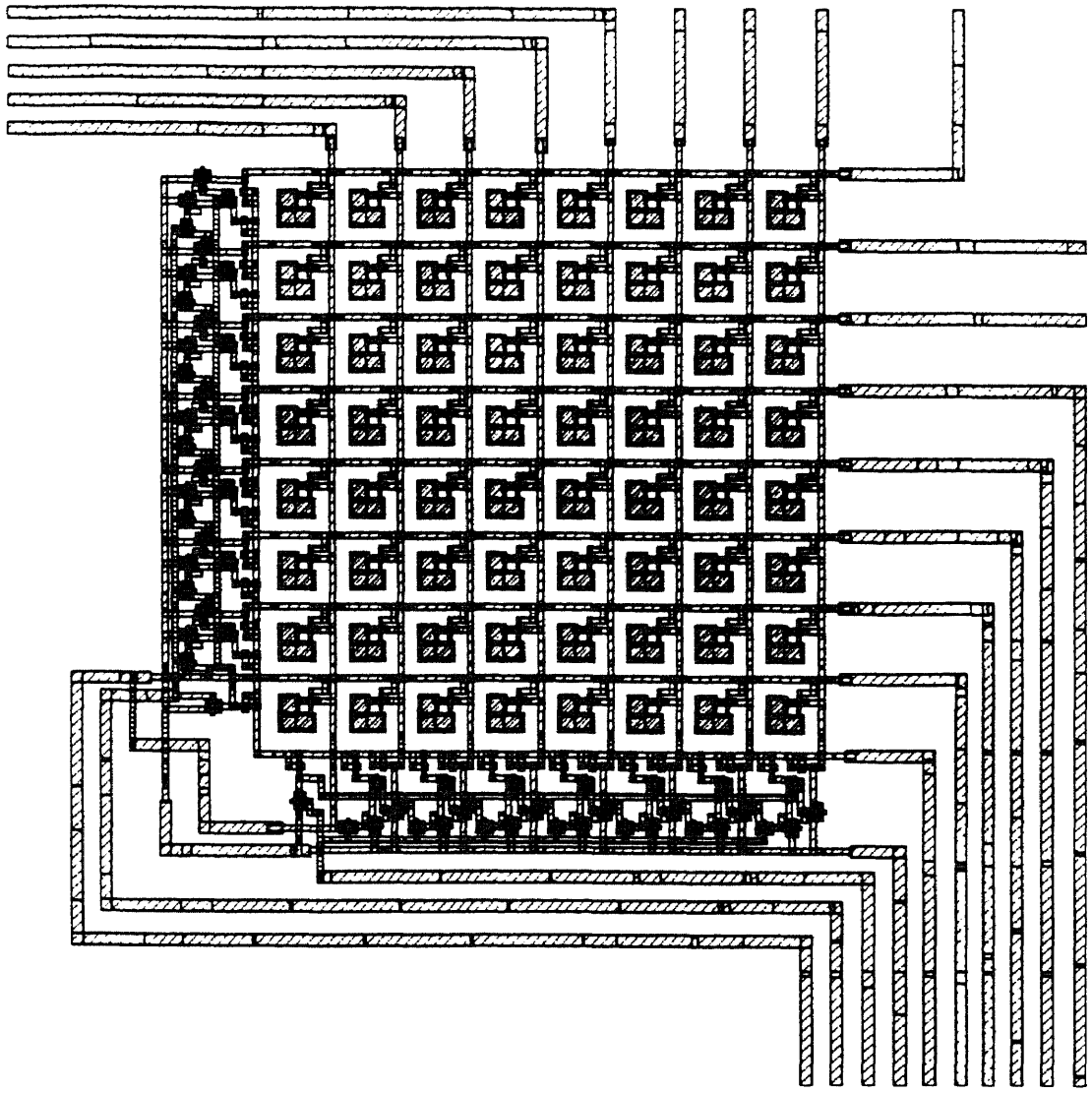

Figure 2.5 Layout of the 8 x 8 shift register

2.2.3 Test Structures

Apart from the address circuit design several test structures were also incorporated to measure various device as well as process parameters. Test structures here consisted of individual transistors of various W/L dimensions, structures to measure sheet and contact resistances, inverters and structures to measure the capacitances. The following shows the layout of some of these structures.

Test structures to measure the sheet resistance of metal and polysilicon were included.

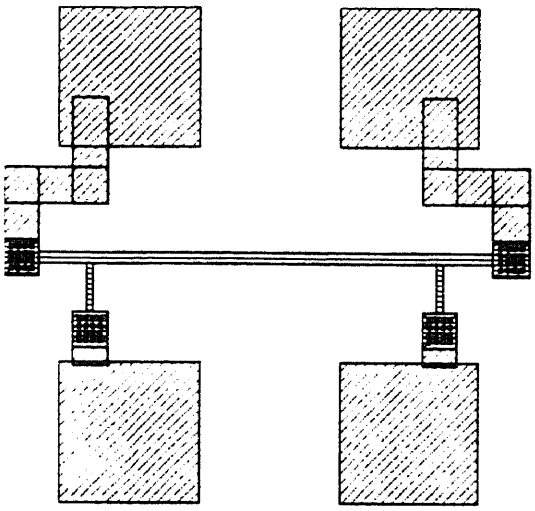

Figure 2.6 Layout of device to measure the sheet resistance

Test structures to measure the contact resistance offered by metal and N+ diffusion active region and metal and polysilicon were included.

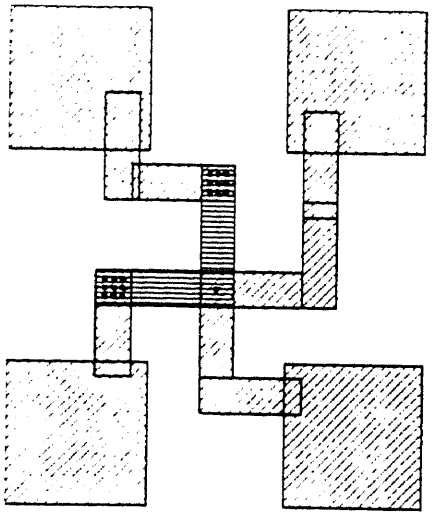

Figure 2.7 Layout of device to measure contact resistance

Test structures to measure capacitance formed between metal and polysilicon, metal and substrate, and polysilicon and substrate were included.

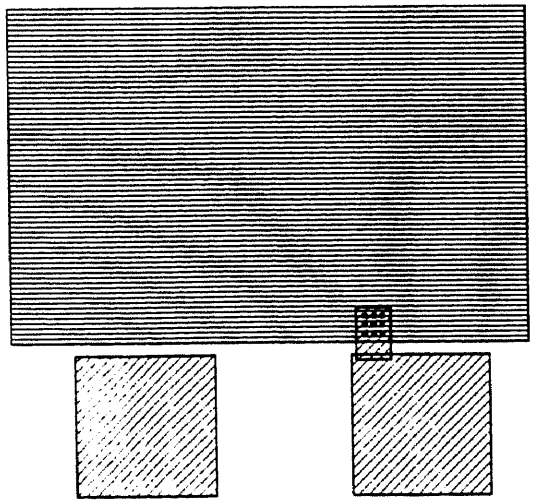

Figure 2.8 Layout of device to measure the capacitance

Transistors of various W/L ratios were included. HP4145B was used to measure their output characteristics as well as the threshold voltage.

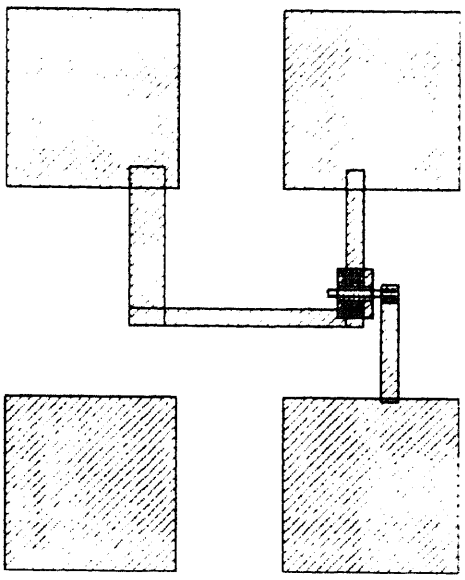

Figure 2.9 Layout of a single transistor

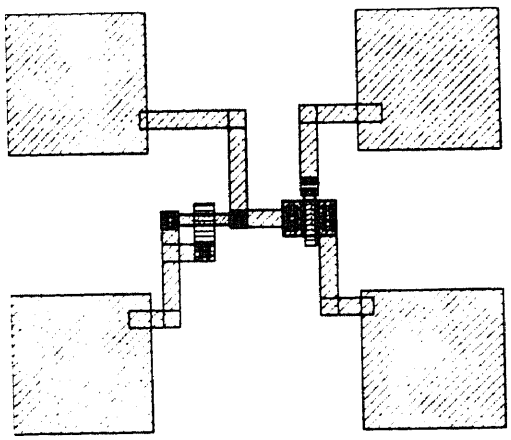

Figure 2.10 Layout to verify inverter characteristics

CHAPTER 3

FABRICATION

3.1 Introduction

In this chapter the detailed fabrication sequence is described. A brief process flow is described below. A cross sectional view of the device after every photolithography step is shown. All the process steps were done in the NJIT Microelectronics Research Center except the ion implantation steps. Ion implantations were performed at Ion Implant Services (Sunnyvale, CA). All implantation processes, such as the channel stop implantation, implantation for threshold voltage adjustment and source and drain implantation, were simulated using SUPREME III simulation package.

3.2 NJIT Clean Room

All the process work was done in the NJIT clean-room. NJIT clean room is a 1200-sq.-ft and class 10 fabrication line. It was equipped with all the necessary tools for processing wafers up to 150mm in diameter. These equipment include:

1) Wafer Inspection - microscope, Dektak profilemeter

2) Nanometrics optical line width

3) Wet chemical station Ultratek mask/wafer scrub

4) Semitool spin/rinse dryers

5) Karl Suss exposure system

6) Nanometrics FTM

7) Inspection microscope

8) MTI Coat and develop system

9) Drytek reactive ion etching system

10) Leitz MPV FTM

11) Varian sputtering system

12) BTU diffusion furnace

13) BTU LPCVD furnace

14) MDA toxic gas monitors

15) Tubewash station

16) MG Industries gas cabinets

3.3 Outline of the Process Steps for Fabrication

An outline of the processing steps that need to be followed is outlined. First the flow charts for the fabrication of the device is detailed and following this a schematic of the process is detailed.

3.3.1 Flow Chart for Fabrication of the Device

The following shows a flow diagram of the entire processing to fabricate the circuitry component of the two part NVRAM.

Silicon Wafer

↓

Deposition of denuding oxide

↓

Stripping of denuding oxide

↓

Deposition of pad oxide

↓

Deposition of LPCVD nitride

↓

Photolithography to pattern nitride to define active area

(Mask 1)

↓

Reactive Ion Etching (RIE) of nitride

↓

Channel Stop Implantation

↓

Steam oxidation to grow field oxide (LOCOS)

↓

Strip Silicon Nitride

↓

Deposition of Sacrificial gate oxide

↓

Implantation to adjust Threshold Voltage (V_t)

↓

Stripping of Sacrificial oxide

↓

Dry oxidation to grow gate oxide

↓

Deposition of polysilicon

↓

Polysilicon doping with $POCl_3$

↓

Photolithography to pattern polysilicon

(Mask 2)

↓

Reactive Ion Etching (RIE) of polysilicon

↓

Source and Drain implantation and Drive in

↓

Deposition of LTO

↓

Photolithography to pattern LTO

(Mask 3)

↓

Reactive Ion Etching (RIE) of oxide

↓

Sputtering of Metal

↓

Photolithography to pattern the Metal

(Mask 4)

↓

Wet Etch of the Metal

↓

Annealing of metal

3.3.2 Process Flow

A detailed description of the process is described below followed by the cross sectional view of the wafer after various steps of fabrication.

3.3.2.1 Starting Materials: Starting materials for the NJIT NVRAM process were p-typed (boron doped), <100> oriented silicon wafers, with a resistivity of 1.65-3.85 ohm-cm.

3.3.2.2 Denuding: Each wafer was scribed on the back for identification. Approximately 1000Å thick oxide was grown by steam oxidation at 950°C performed for 40 minutes that was stripped subsequently using 6:1 BOE. This step helped in removing the impurities and contaminants from the wafer surface.

Impurities can degrade the performance of circuits or even cause failure. Hence for high yield and device reliability it is essential to eliminate all source of contamination. Surface cleaning is important prior to high temperature processes because impurities react and diffuse at much higher rates at elevated temperatures. The most commonly used wet chemical cleaning technology is based on hot alkaline or acidic peroxide (H_2O_2) solutions. These are used to remove chemically bonded films from the wafer surface prior to critical steps. RCA cleans are based on a two step process: SC-1 followed by a SC-2. Both solutions incorporate the strong oxidizing capability of H_2O_2. SC-1 is an aqueous alkaline solution which removes organic films, while SC-2 is an acidic mixture used to remove alkali ions and cations, and metallic contaminants. SC-1 is typically a 5:1:1 solution of DI water, "unstabilized" H_2O_2 (30%), and ammonium hydroxide (27%). This solution is very effective in removing organic contaminants. SC-2 typically consists of 6 pats H_2O, 1 part H_2O_2 (30%), and 1 part hydrochloric acid (HCl, 37%) and is effective in removing heavy metals. The processing temperature is 80°C.

3.3.2.3 Definition of Active Area: A thin oxide (~250Å), called pad oxide was grown followed by LPCVD nitride of 1200Å thickness. The pad oxide helped to reduce stress between the silicon active layer and nitride. Photolithography was performed to define the active area using the active mask.

The photoresist (PR) is applied as a thin film to the substrate and subsequently exposed through a mask. The mask contains clear and opaque features that define the patterns to be created in the PR layer. The areas in the PR exposed to the light are made either soluble or insoluble in a specific solvent known as developer. Here resist was retained only on the active region and was removed from elsewhere. To start with the wafer was primed with a pre-resist coating of a material designed for better photoresist adhesion. The wafers were then ready to be coated with photoresist. Spin coating is the most widely used technique to apply a uniform and adherent film of desired thickness. This procedure was carried out by dispensing the resist solution on the wafer surface, and then rapidly spinning the resist until the resist was essentially dry. After the wafers were coated with resist, they were subjected to a temperature step, called soft-bake (or pre-bake). This step accomplished several important purposes, including driving off solvent from the spun-on resist and improving adhesion of the resist. After the wafer had been coated with resist and suitably soft-baked, it was ready to be exposed to some form of radiation in order to create a image on the resist. The degree of exposure was adjusted by controlling the energy impinging on the resist. Following exposure, the resist film was made to undergo development in order to leave behind the image, which would serve as the mask for etching. The wafer was then spun-dry and transported to the postbake module. Following development, an inspection was performed. The purpose was to insure that the steps of

the PR process up to this point have been performed correctly and to within the specified tolerance. Any inadequately processed wafers detected by this inspection could have their resist stripped and reworked. Post-baking was then performed just prior to etching. Its chief purposes include to remove residual solvents, to improve the adhesion, and to increase the etch resistance of the resist.

Nitride was etched using Phantom. Details of process conditions are:

40 sccm CF_4

Pressure = 250 mTorr

Power = 150 Watts

Temperature = 25 °C .

Visual inspection was made to make sure that the nitride was etched completely. Photoresist was stripped subsequently.

3.3.2.4 Channel Stop Implantation: Channel Stop Implantation was then done for isolations between the various NMOS transistors. Ion implantation is a process in which energetic, charged atoms or molecules are directly introduced into a substrate. The main objective of ion implantation is to introduce a desired atomic species into a target material. Details of the implantation were:

Species: $Boron^{11}$

Energy: 50 KeV

Dose: 8.0×10^{12}

The photoresist was then stripped using Mpyrol.

3.3.2.5 Field Oxidation: P-clean followed by a furnace pre-clean was performed before loading the wafers into the furnace. Furnace pre-clean was carried out in a 100:1 H_2O:HF for 5 minutes which ensures the removal of the pad oxide from all areas except the active area. 7500Å thick steam oxide was grown for 350 minutes at 1050 °C (LOCOS). Although there are several ways to produce SiO_2 directly on the Si surface, it is most often accomplished by thermal oxidation, in which the silicon is exposed to an oxidizing ambient (O_2, H_2O) at elevated temperatures. Thermal oxidation is capable of producing SiO_2 films with controlled thickness and Si/SiO_2 interface properties.

When Si is exposed to an oxidizing ambient at elevated temperatures, more rapid growth and thicker oxides are produced. The oxidation proceeds by the diffusion of an oxidant (molecular H_2O or O_2) through the existing oxide to the Si or SiO_2 interface, where the molecules react with Si to form SiO_2. The reactions occur at the Si /SiO_2 interface. Therefore as the oxide grows, silicon is consumed and the interface moves into the silicon. Based on the relative densities and molecular weights of Si and SiO_2 it is found that the amount of silicon consumed is 44% of the final oxide thickness. The reaction between oxygen and silicon occurs at the silicon-oxide interface and, to react with silicon, oxygen must diffuse through the growing oxide and reach the silicon surface.

Si_3N_4 was then stripped in a H_3PO_4 bath at 170 °C . A 100% overetch time was given to achieve complete removal of nitride. Subsequently the pad oxide was removed using 100:1 H_2O:HF for 5 minutes.

3.3.2.6 Threshold Voltage (V_t) Adjustment: V_t plays a crucial role in the performance of the devices. A sacrificial oxide (~900Å) was grown using steam oxidation. This oxide

helps in protecting the active area from the damages that would otherwise occur from direct ion bombardment. Details of the implantation are:

Species: BF_2

Energy: 60KeV

Dose: 2.0×10^{12}

After the implantation for V_t adjustment the sacrificial oxide was wet etched in a 6:1 BOE.

3.3.2.7 Gate Oxidation and Poly-I Deposition: Gate oxide plays a significant role in the performance of the devices. The objective here is to grow the purest possible oxide. For this the furnace tube was pre-purged at 1050 °C for one hour. A 400Å oxide was grown by dry oxidation at 950 °C. Immediately after that a 4000Å polysilicon was deposited by LPCVD at 600 °C. Polysilicon was doped by phosphorus diffusion/drive-in at 950 °C for two hours. Then the sheet resistivity was measured.

3.3.2.8 Gate Definition: Photoresist on top of Poly was patterned using the Poly mask using the process of photolithography as described above. Polysilicon was etched using Phantom etcher Details of process conditions are:

40 sccm SF_6

Pressure = 250 mTorr

Power = 150 Watts

Temperature = 25 °C .

3.3.2.9 Source and Drain Implantation: The wafers were shipped for N+ source and drain implantation. The details of the implantation are:

Ion species : As^{75}

Energy : 30 KeV

Dose : $2.0 \times 10^{16}/cm^2$

Subsequent to the implantations a drive-in was performed at 925 °C for 20 minutes.

3.3.2.10 LTO Deposition: Low temperature oxide (LTO) was deposited which acts as an insulator through which contacts to active and contacts to poly would be performed. The oxide thickness was 5500Å and was deposited at 425 °C .

3.3.2.11 Contact Hole Etching: Photolithography was performed using the contact mask. Oxide was etched and visually inspected so as to make sure that the etch was complete. Photoresist was subsequently stripped.

3.3.2.12 Al-Cu-Si Deposition and Patterning: Approximately one micron Al-Cu-Si metal was sputter deposited. Aluminum and its alloys are primarily used as material, which interconnects the device structures, formed in the silicon substrate. Sputtering is the term used to describe the mechanism in which atoms are dislodged from the surface of a material by collision with high energy particles. The sputtering process basically consists of four steps, ions are generated and directed at a target, the ion sputter target atoms, the ejected (sputtered) atoms are transported to the substrate, these atoms condense and form a thin film.

Photolithography was performed using the metal mask. Metal etch was performed using wet etch.

3.3.2.13 Annealing: The wafers were then annealed in forming gas at 400 °C for 30 min and the wafers were ready to be tested.

3.3.3 Supreme Simulations

The following graphs and table show the simulations done in SUPREME III to obtain the exact value of dose and energy for the various implantations.

Table 3.1 Ion Implantation's Involved

Process Steps	Species	Energy (KeV)	Dose (/cm2)
Channel Stop implant	B^{11}	50 KeV	8.0×10^{12}
Threshold adjustment (for N-MOS)	BF_2	60 KeV	2.0×10^{12}
N+ S/D Implant	As^{75}	30 KeV	2.0×10^{16}

Figure 3.1 shows the results of the SUPREME III simulations to determine the depth of channel stop implantation. The values used were dose $8e^{12}$ and energy level 50KeV of Boron which gave a junction depth of 1.75 micron.

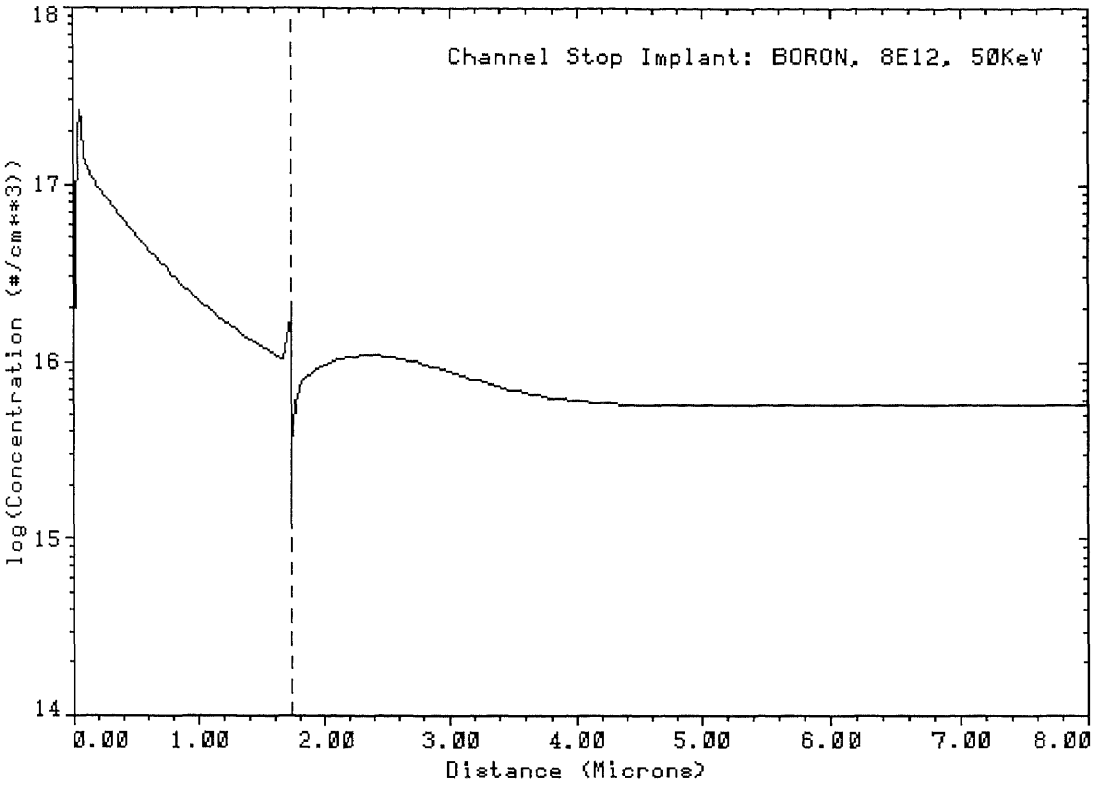

Figure 3.1 Channel stop simulation

Figure 3.2 shows the results of the simulation done to control the threshold voltage (Vt). BF_2 was used at a dose of $2e^{12}$ and energy value of 60 KeV.

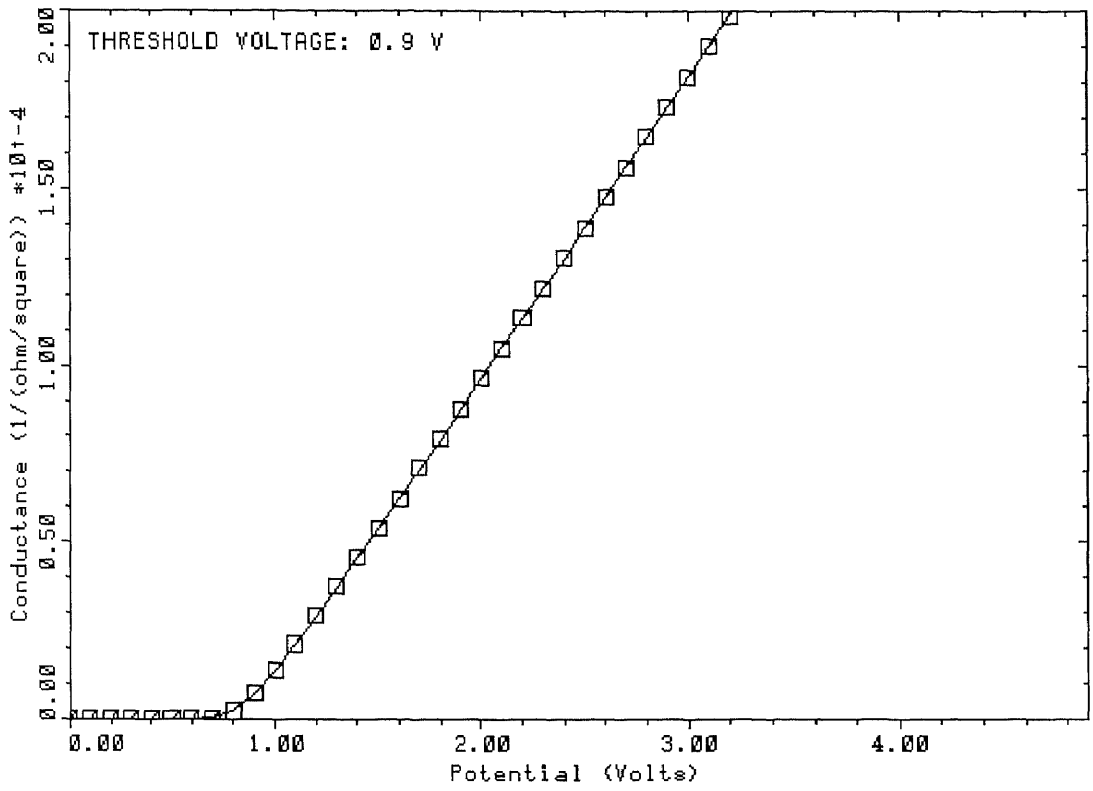

Figure 3.2 Threshold voltage simulation

Figure 3.3 shows the simulations done to determine the dose and energy values for source and drain implant. As^{75} was used with a dose value of $2e^{16}$ and energy value of 30KeV. The junction depth after drive in was observed to be 0.40 microns.

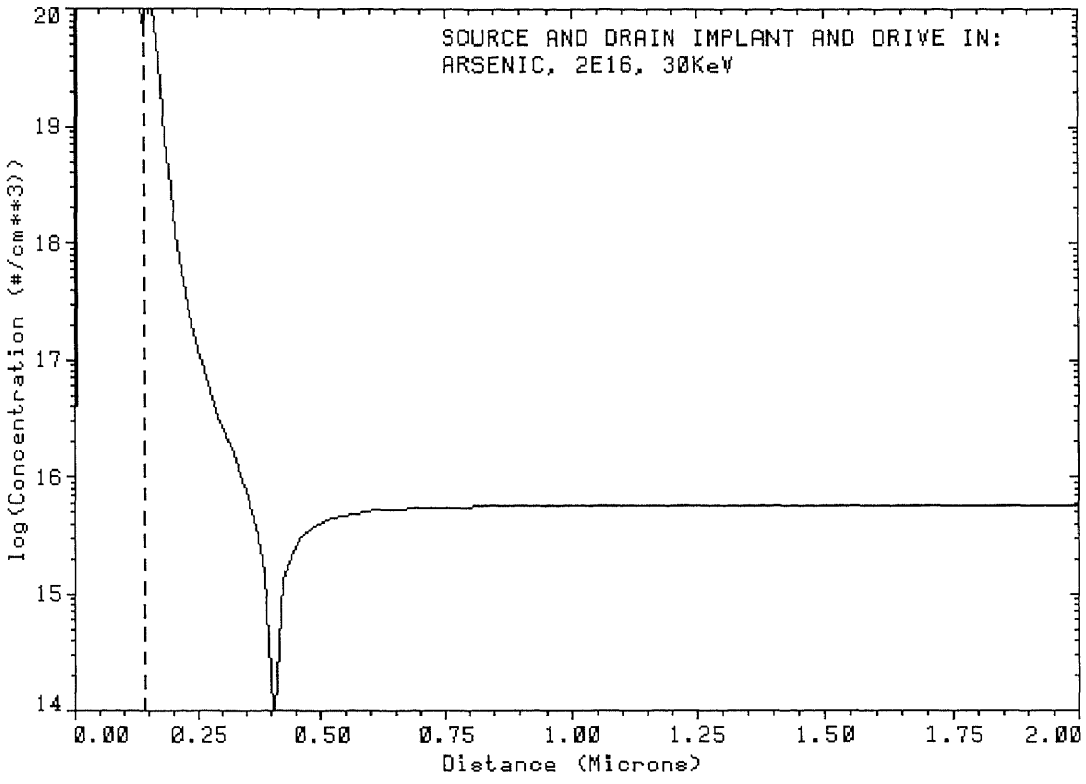

Figure 3.3 Source and drain implantation

3.3.4 Mask Definitions

The following table depicts the alignment sequence of the masks.

Table 3.2 Mask alignment sequence

Mask Name	Type	Field	Aligning Sequence
Active	Chrome	Clear	
Poly	Chrome	Clear	Active
Contact	Chrome	Dark	Poly
Metal	Chrome	Clear	Contact
Polymer	Chrome	Clear	Metal

3.3.5 Cross Sections

The following shows the step by step processing of the wafer and it's cross sectional view after each step.

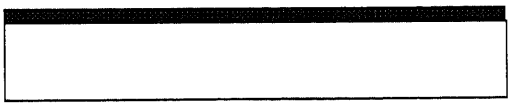

Figure 3.4a The original Silicon wafer was p-type. 1000 Å steam oxide was grown on the wafer which was the denuding oxide which was subsequently stripped.

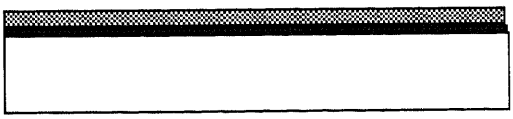

Figure 3.4b The oxide was then stripped and a fresh pad oxide was grown on the wafer. On this oxide a 1200 Å thick silicon nitride was deposited.

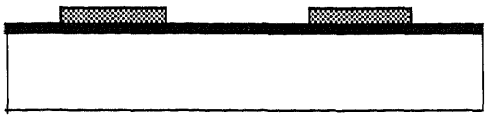

Figure 3.4c The nitride was etched using RIE. (Active mask)

Figure 3.4d 0.75 micron of thick field oxide was grown (LOCOS).

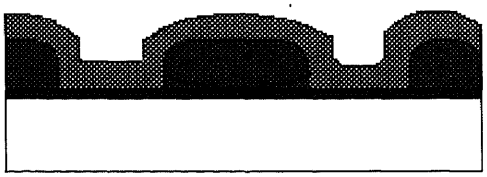

Figure 3.4e The nitride was stripped and the pad oxide removed. A thin gate oxide (dry oxidation) of 400 Å was grown followed by deposition of 4000 Å polysilicon. The polysilicon was doped with Phosphorus (POCl3).

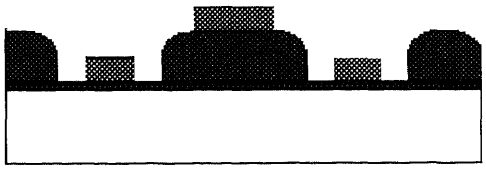

Figure 3.4f The polysilicon was patterned and RIE etched as shown. This polysilicon over the active area acts as a self-aligning mask during the formation of the source and drain.

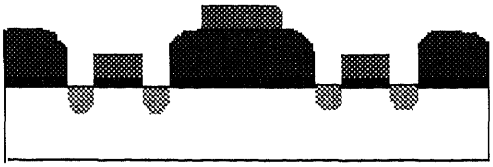

Figure 3.4g Source and Drain implantation was then done.

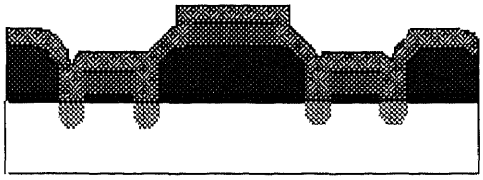

Figure 3.4h 6000 Å thick low temperature oxide was deposited

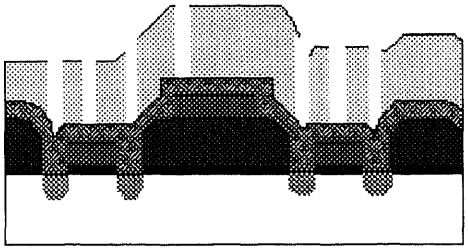

Figure 3.4i Photoresist was then deposited and was patterned as shown using contact masks.

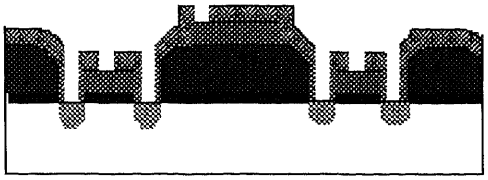

Figure 3.4j The oxide was then etched using RIE so as to open up holes for metal contact.

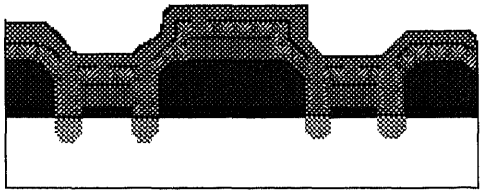

Figure 3.4k Metal of thickness 1.0 micron was then deposited as shown.

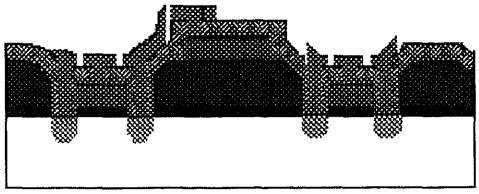

Figure 3.4l Metal was patterned using metal mask as shown above.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Overview

The method used to test the shift register circuit as well as the test structures is detailed below. A discussion of the results obtained is also included.

To test the test structures the HP 4145B Semiconductor Parameter Analyzer was used. The HP 4145B has four SMUs (Source Monitor Unit), two voltage monitors (Vm) and two voltage sources (Vs). Connections are defined according to the device being measured. However, there are certain basic testing programs already defined like the FET, diode, BJT that can be modified as desired.

4.2 Test Structures

Among the test structures tested to verify the process adopted for fabrication as well as those to test the transistor parameters were the sheet resistance measuring structures, contact resistance measuring structures, capacitance structures, inverters and transistors of different W/L ratios. The sheet and contact resistance measuring structures were there for different layers, such as, metal and polysilicon. Similarly there were capacitance test structures with capacitance being formed between different layers such as metal and polysilicon, polysilicon and silicon, and metal and silicon with oxide being the dielectric.

4.2.1 Sheet Resistance

To measure the sheet resistance the Vander-Pauw structures were used. The Vander-Pauw structure consists of 4 leads. Two of the outer leads are used to drive a current and the voltage is tapped across the other two leads. The ratio of the measured voltage to the drive current per number of squares between the two output (voltage) leads gives the value of sheet resistance. Table 4.1 of device shows the measured values of sheet resistance of the n+ diffusion in active area, metal and the polysilicon layers. The sheet resistance is determined from the van der pauw formula

$$Rs = \prod R(+/- I) \ / \ \ln(2)$$

where R(+/- I) is the incremental resistance determined from the current and voltage measurements for both directuions of the current and both contact orientations

Table 4.1 Sheet Resistance

Parameter	Test Device	Range (Ω/□)		Av. Value (Ω/□)
Sheet Resistance	N+ diffusion in active area	Max	60	50
		Min	40	
	Poly-1 on gate oxide in N+ active	Max	35	30
		Min	25	
	Metal	Max	0.08	0.07
		Min	0.05	

4.2.2 Contact Resistance

To measure the contact resistance the current was driven between the two bonding pads in the transverse direction (as opposed to longitudinal ones done in sheet resistance measurement). The developed voltage was measured across the other two bonding pads. Table 5.2 shows the measured values of contact resistance of the metal and n+ diffusion in active area, metal and the polysilicon. The contact resistance Rc is given in ohms and is defined as

Rc = Vf / I i.e. is the ratio of the voltage drop across a conductor-to-semiconductor interface, to the current being forced through the interface. The Transmission Line Model (TLM) is the model often used to describe the contact region. It defines

$$Vf / I = Ve/I \cosh \sqrt{(Rs/\rho_c)}\, d$$

where ρ_c = specific contact resistance, Rs = sheet resistance of diffusion level, d = length of contact window, Vf = voltage at the front of the TLM, I = current entering the contact, Ve = voltage at the end of the TLM. Here the Kelvin resistor structures are used to extract the value of the contact resistance.

Table 4.2 Contact Resistance

Parameter	Device	Range	(Ohms)	Av. Value (Ohms)
Contact Resistance	Metal-1/N+ diff-usion of active	Max	145	140
		Min	130	
	Metal-1/Poly1	Max	10	9
		Min	8	

4.2.3 Capacitance

The various capacitances measured using the test structures were the capacitance between the metal and polysilicon, capacitance between the polysilicon and substrate. The capacitance-voltage characteristics of an MOS structure depend on the state of the semiconductor surface. Depending on the gate voltage, the surface may be in accumulation, depletion or inversion region. In the p-substrate an accumulation layer is formed when gate voltage is less than 0. The negative charge on the gate attracts holes toward the silicon surface. When an accumulation region is present the MOS structure behaves like a parallel plate capacitor. The gate conductor forms one plate of the capacitor, the holes in p-substrate the other plate of the capacitor. The

$$\text{capacitance } Co = (\varepsilon_{SiO2}\,\varepsilon_o / t_{ox})\,A$$

where A = area of gate

and ε_{SiO2} = dielectric constant

In the depletion region the capacitance value is given by

$$C_{dep} = (\varepsilon_o \varepsilon_{Si} / d)\,A$$

where d = depletion layer depth

and ε_{Si} = dielectric constant of silicon.

The capacitance between the metal and polysilicon had the LTO as the dielectric layer. The capacitance was measured using the 4145B in conjunction with a Boonton 7200B capacitance meter, which was operated at a frequency of 1MHz. The voltage applied between the probed points in the test structure was actually applied through the capacitance meter. The capacitance meter has the capability of converting the measured

capacitance value between the parallel plates of the capacitor to the corresponding value of current, which it feeds back to the 4145B. Given below are the values of the various capacitances per unit area.

Capacitance:

Poly/p substrate	34 aF/μm^2
Metal/Poly	38 aF/μm^2
Gate oxide thickness:	367Å

4.2.4 Inverter

Inverter structures were included in the design. The W/L ratios for the NMOS transistor as well as the load resistance transistor were obtained by Pspice simulation. The W/L ratio of the NMOS transistor was 20/4 and that of the load resistor was 4/8. The load transistor was connected in the enhancement mode. Figure 4.1 shows the schematic of the inverter circuit.

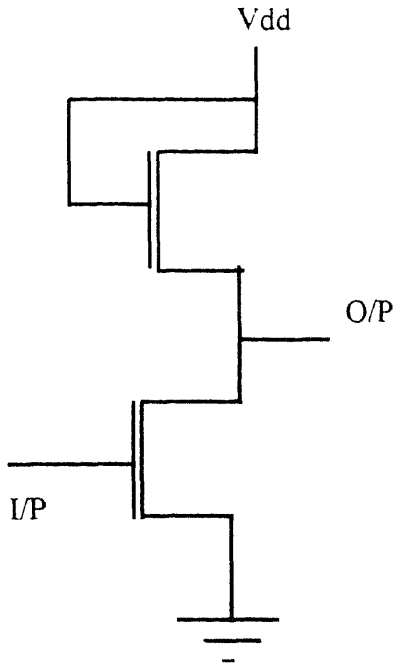

Figure 4.1 Schematic of the inverter circuit

The Vdd is connected to 5V, the i/p is a pulse of amplitude 7V, gnd to 0V and the output value is measured. The output is expected to be inverted and follow the Vdd. As we can see in figure 4.2 the amplitude of the output is very low and this can be attributed to the leaky nature of the transistor.

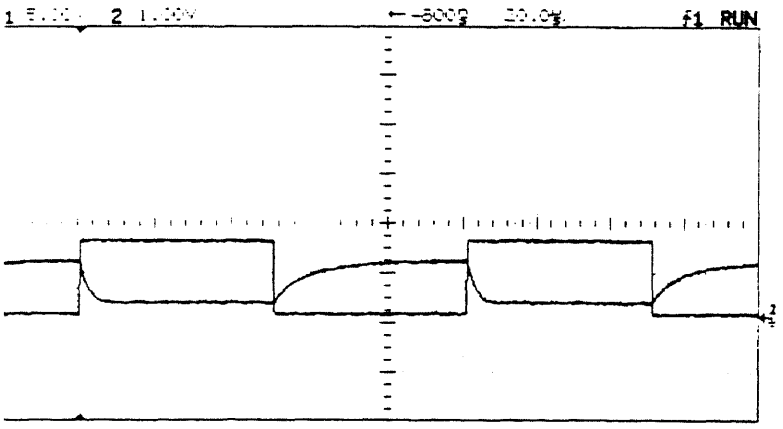

Figure 4.2 Inverter characteristics obtained from NMOS

4.2.5 NMOS transistor

NMOS transistors of various W/L ratios were included in the design. Figure 4.3 shows the typical drain current Vs. drain voltage of an NMOS transistor. The threshold voltage is varied from 0V to 5V.

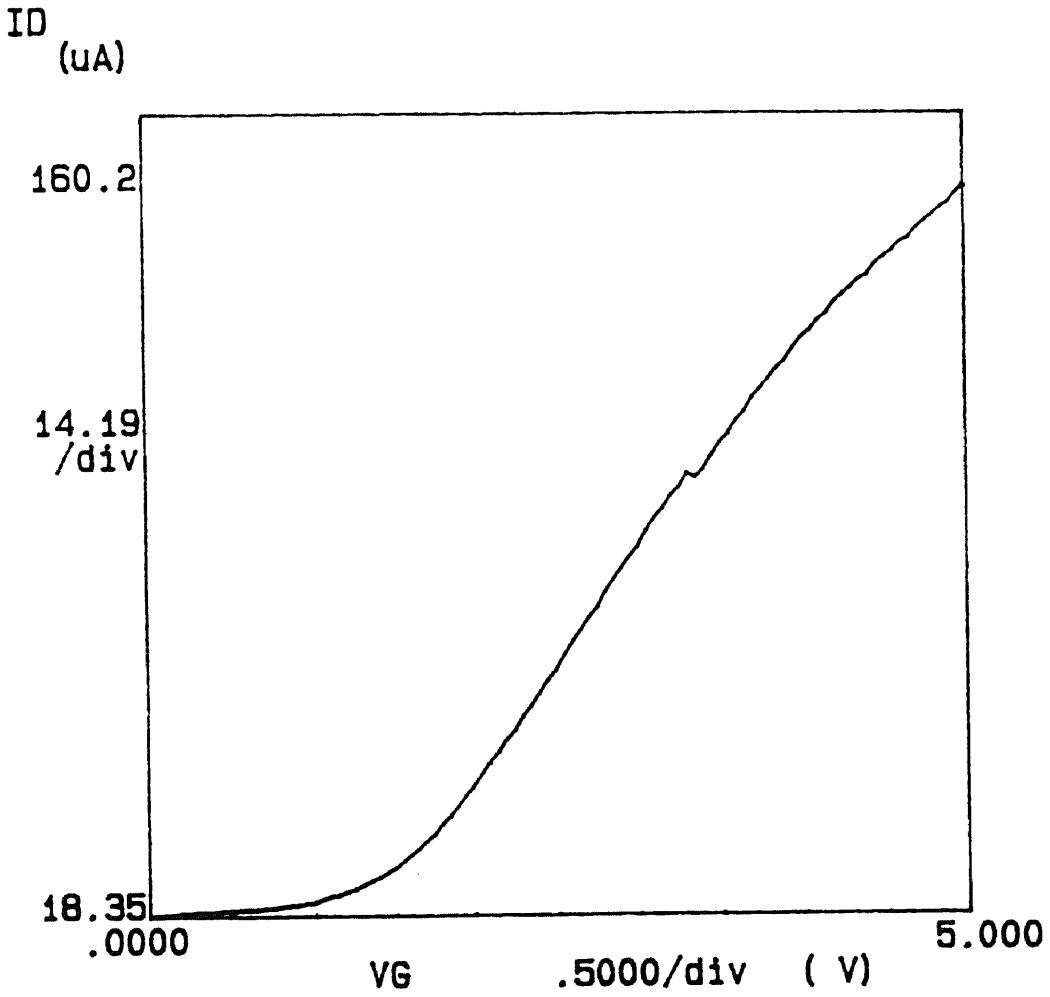

Figure 4.3 Drain current Vs. drain voltage characteristic of NMOS

The MOS transistors have three regions of operation, Cut-off region, linear region, and saturation region.

The drain-to-source current Ids is described as

Ids = 0 in the cut-off region i.e. when Vgs - Vt <= 0

Ids = β[(Vgs - Vt)Vds - V^2ds / 2] in the linear region i.e. when 0 < Vds < Vgs - Vt

Ids = (β / 2) (Vgs - Vt) ^{2}

where Vgs = gate-to-source voltage, Vt = device threshold, β = MOS transistor gain factor which is dependent on both the process parametrs as well as the device geometry.

Figure 4.4 shows the drain current Vs. threshold voltage at constant Vd of 5V. The threshold voltage is 0.9 volts and is close to the simulated value in SUPREME III.

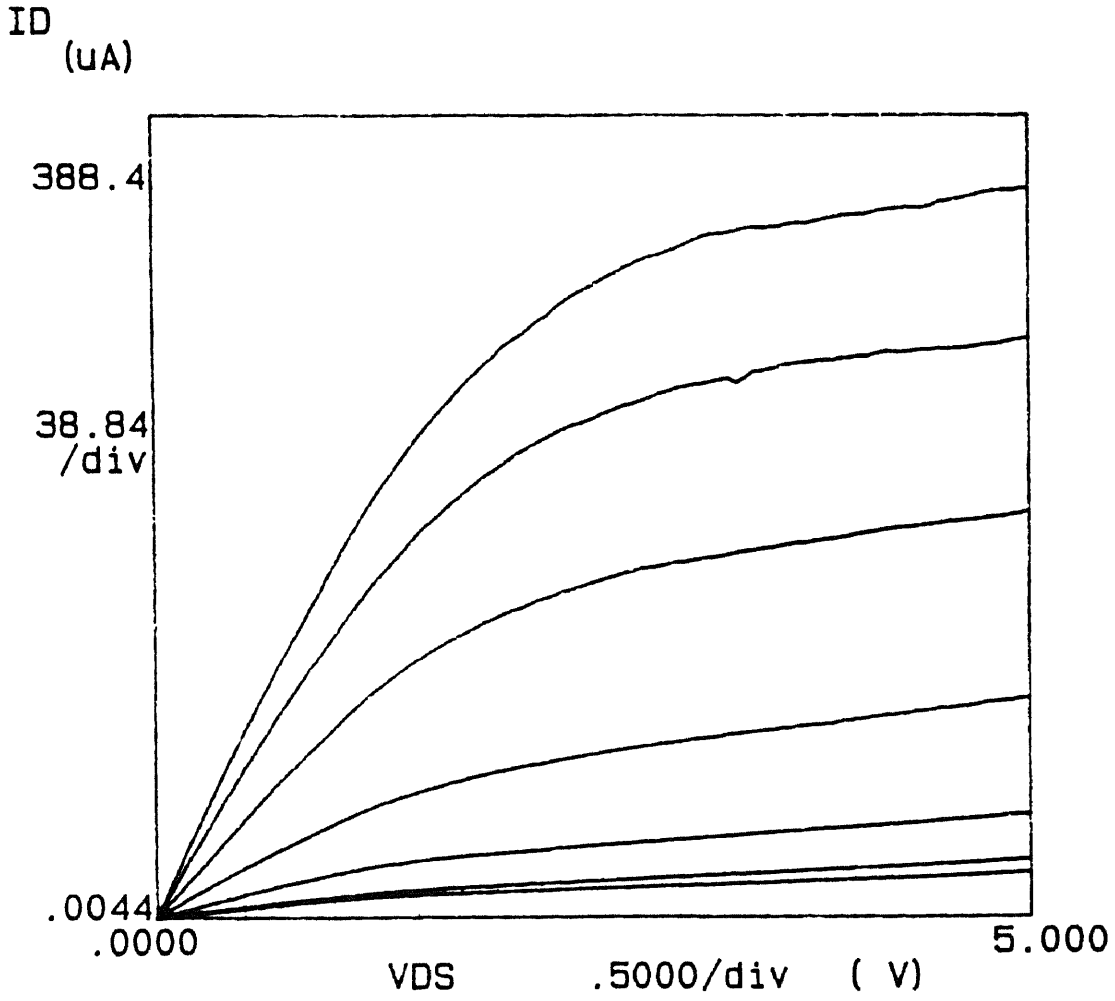

Figure 4.4 Drain current Vs. gate voltage characteristic for NMOS

4.3 Shift Register Circuit

The shift register circuit was tested. For the purpose of testing the shift register circuit an input pulse was applied using the pulse generator and the clock pulse was applied using the function generator. The shift register operates on the principle that simple ratioed inverters are connected by pass transistors clocked with a two phase clock. Charges stored on the gates of the driver transistors holds the inverter output state between the clock periods.

To measure the output characteristics of the shift register circuit a constant voltage source was applied for the Vdd. The output was measured using an oscilloscope. The details of one such measurement is shown below in figure 4.5. The shift register circuit essentially consists of an input pulse, two clock inputs, the Vdd, ground terminal and the output terminal. The two clock inputs which were applied were the clock and the clock inputs. The clock frequency was at 500Hz. The input applied was at 50MHz and the output was observed to be following the frequency of the input clock at 500Hz as seen from the graph. The amplitude of the output pulse was observed to be equal to 3.8 volts while the amplitude of the input pulse was observed to be 6.2 volts.

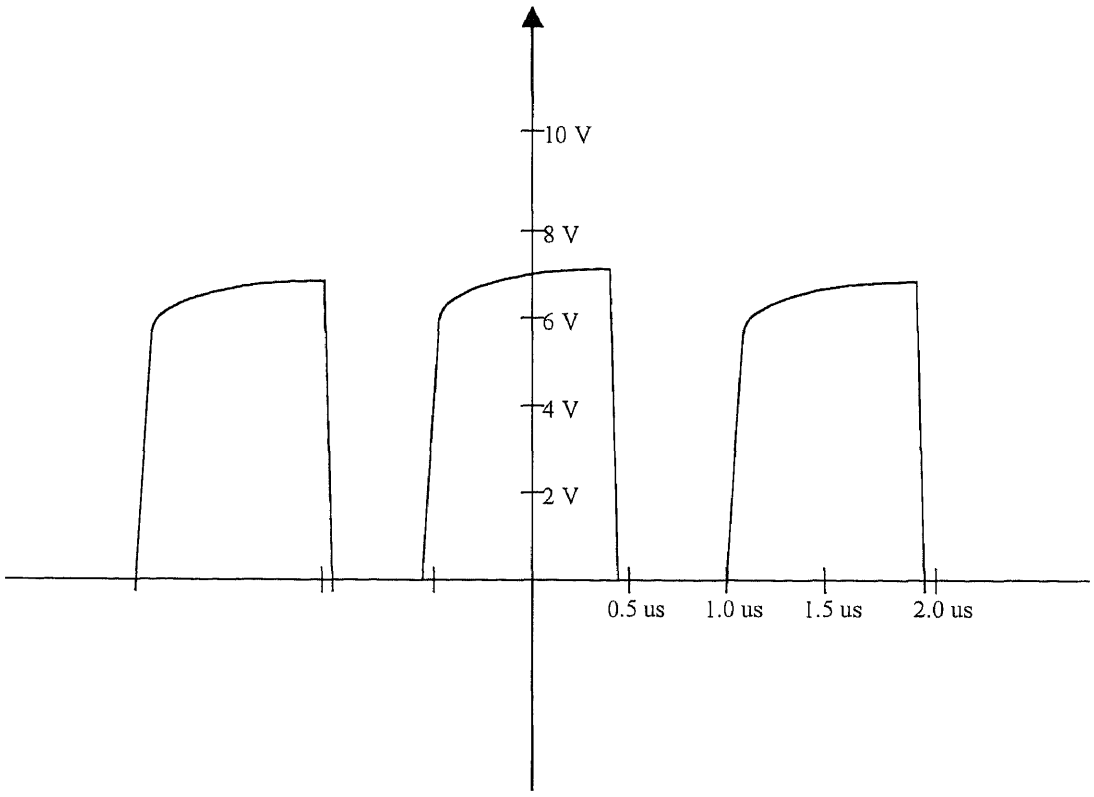

Figure 4.5 a. Shift register input

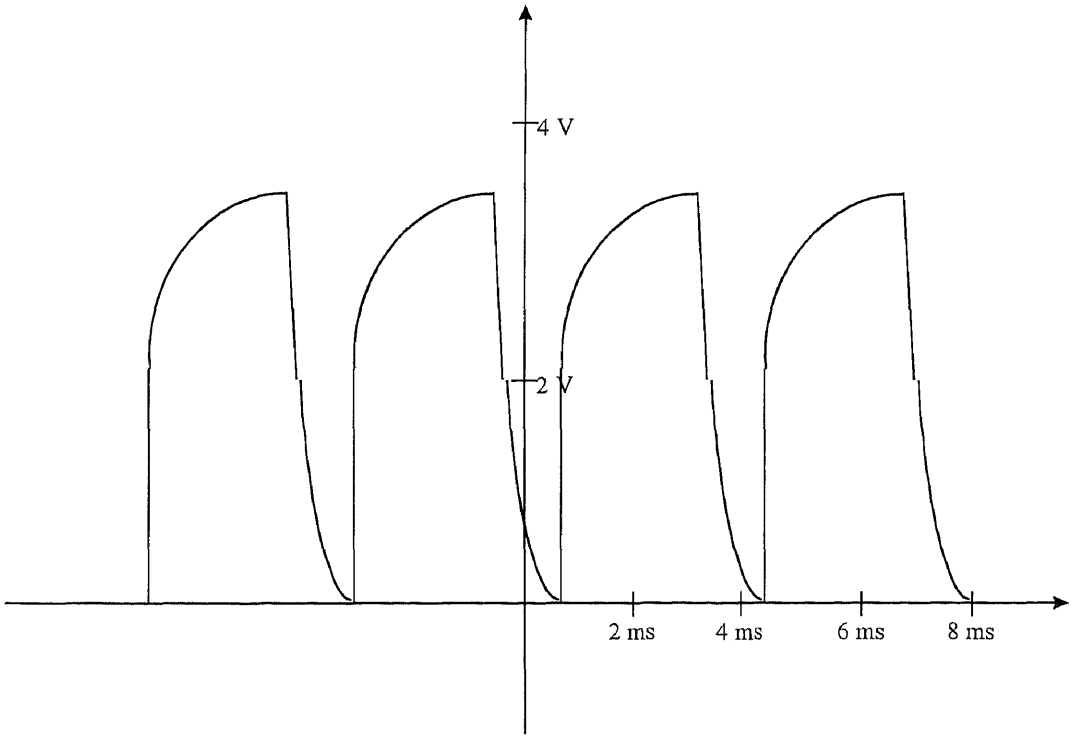

Figure 4.5b Shift register output

4.4 Indium Bump Pads

The indium bump pads were deposited on the circuitry part using the lift-off technique. The dies were then diced. To test the conductivity of the indium bump pads as well to estimate the contact resistance uniformity throughout the circuit, a glass plate with indium tin oxide (ITO) was flipped over this circuitry component. This ITO would be later replaced by the actual ferroelectric material. Gold wiring was done to connect the bonding pads of the circuit to copper pads situated on the surface on which the device was mounted. Hence, the copper pads could be probed for testing purposes. The output was taken from the glass plate with the conducting material. The glass plate with the conducting material was further connected through a resistor to the ground. The results obtained are shown in figure 5.6. For the purpose of testing two separate clocks were applied to the two sets of shift registers. In order to be able to have only one transistor conducting at any point of time the output of the shift register placed horizontally was connected to clock1 which was operating at 2.8 microseconds and the output of the shift registers placed vertically had the clock termed as clock2 operating at

$$128 \times 2.8 \text{ microseconds} = 372 \text{ microseconds.}$$

In this way by the time all the columns (bit line) were made high once the vertical output (word line) would advance upward by one position. The columns would go high one by one again but this time the transistors present in one row above would be turned on. The input pulse applied through a pulse generator was of time period 2.5 microseconds. The Vdd was at 7 volts. Hence it is expected that the output as seen connected to the conducting material should follow the clock1 i.e. one transistor at a time should be turned

on at the same rate as the clock1. As seen from figure 4.6 the output and the clock1 are in phase with each other.

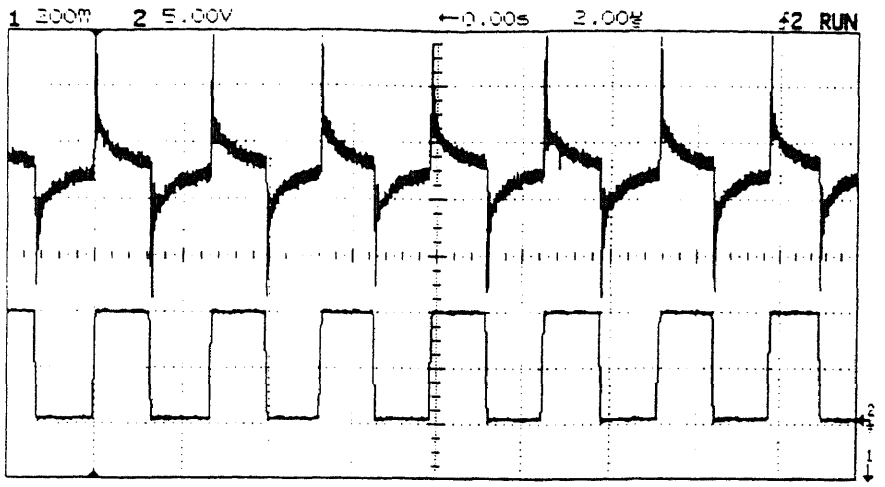

Figure 4.6 Results on testing the indium bump pads

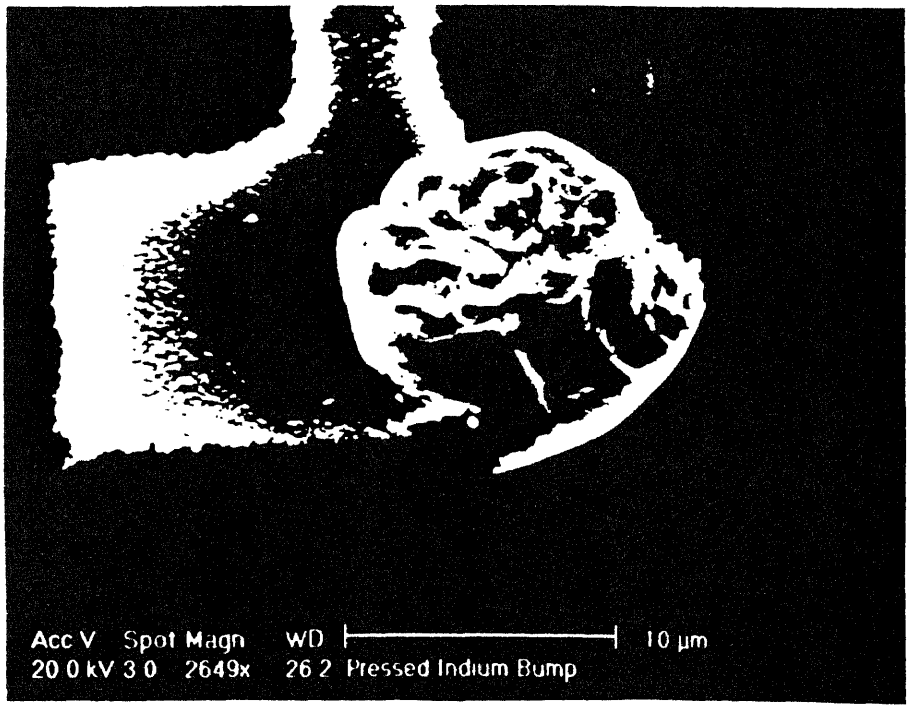

Figure 4.7 Indium Bumps deposited and pressed by the glass plate

4.5 Discussions

The parametric analysis of the test structures such as contact resistance and the sheet resistance shown in table 4.1 and table 4.2 indicate that the process was well under control. The measured value of the threshold voltage (1.0V) shown in figure 4.4 was also very close to the simulated value (0.9V).

Since the results of the process were very encouraging, the shift register circuit was tested. The results obtained on testing the shift register circuit were also as simulated.

Since the objective of the design and fabrication of the circuitry component was to be able to test the conductivity, contact resistance and other properties of the conducting materials, the indium bump pads were deposited and patterned on the circuitry component. The glass plate with indium tin oxide was then flipped over and the completed device was then tested. Successful operation of the preliminary design, fabrication and testing of the two part NVRAM chip illustrates the potential capability of this idea to be implemented in ULSI applications.

CHAPTER 5

CONCLUSIONS

5.1 Overview

Successful implementation of the idea of fabricating a two part NVRAM making use of a conducting material capable of transforming information from the circuitry part to the substrate part (made of ferroelectric material), with 100% faithful transform efficiency is a milestone achievement in the development of the non volatile memory technology. However successful operation of the preliminary design, fabrication and testing of the two-part NVRAM was demonstrated. The test results obtained were very good and matched well with the simulation results during various steps of processing. The testing of the shift register gave good results. Conducting material i.e. the indium bumps were deposited, patterned, and a glass plate with indium tin oxide (ITO) was flipped over. The results of testing this final device gave promising results about the possibility of using such a conducting material for the fabrication of a NVRAM.

5.2 Future Scope

Preliminary simulations, design and fabrication of the chip employed to test various characteristics of the conducting material operating with a substantial yield as reported in the thesis hints towards a potential utilization of this concept in a matured manufacturing environment. As mentioned earlier a NMOS technology with a fairly large design rule was used in the present investigation due to time as well as the various cost considerations.

However, a CMOS technology with a tighter design rule can be implemented to improve the speed, reliability and yield of such a working device.

APPENDIX A

LPCVD LTO THEORY AND DEPOSITION CONDITIONS

A.1 LPCVD Reactor

The deposition reactor is schematically shown in the figure A.1. This reactor was manufactured by Advanced Semiconductor Materials America Inc.(ASM America, Inc.) as a poly silicon micro-pressure CVD system. The horizontal reaction chamber consists of a 13.5 cm diameter fused quartz tube and a 144 cm long encapsulated with a three-zone, 10 kwatt, Thermco MB-80 heating furnace. The flow of oxygen into the reaction chamber is controlled by a MKS mass flow controller.

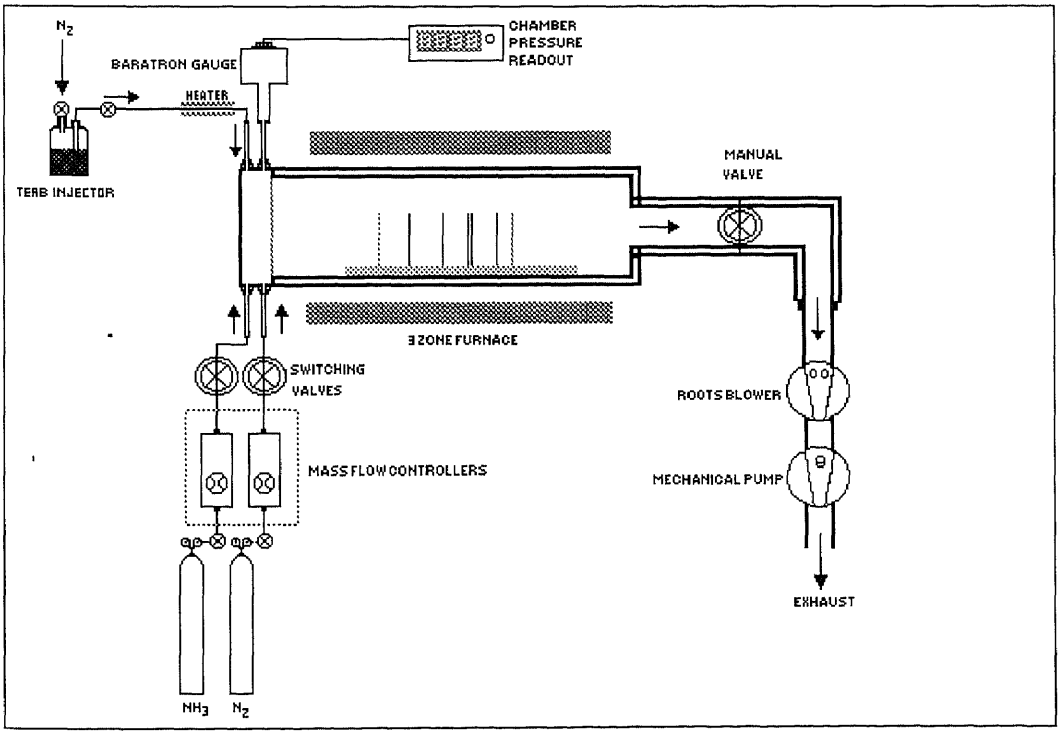

Figure A.1 Schematic representation of the LPCVD reactor

The other end of the reaction chamber is connected to a vacuum station comprised of a Leybold-Heraeus Trivac dual stage rotary vane pump backed by a Leybold-Heraeus roots pump to create the necessary vacuum in the system. An oil filter system is used to filter unnecessary particles from oil and thereby increasing the lifetime of the pump.

A.2 Flow Chart Showing the Oxide Deposition

P-clean

5:1 H_2SO_4: H_2O_2

110^{O}C, 10minutes

↓

Rinse Hot DI Water

10 minutes

↓

Rinse Cold DI Water

5 minutes

↓

Spin Dry

↓

Furnace Pre-clean

100:1 H_2O:HF

1 minute

↓

Rinse Cold DI Water

10 minutes

↓

Spin Dry

↓

LPCVD LTO

Table A.1 Oxide deposition parameters

SiH_4	300sccm
O_2	7.5sccm
Temp	425°C
Pressure	500m Torr
Time	7hours 45 min

APPENDIX B

POLYSILICON DEPOSITION AND DOPING METHODS

B.1 Methods of Polysilicon Deposition

Polysilicon can be deposited in a variety of ways and using many types of precursors.

APCVD : Atmospheric pressure chemical vapor deposition

LPCVD : Low pressure chemical vapor deposition

PECVD : Plasma enhanced chemical vapor deposition

Polysilicon occurs on occasion when the growing of single-crystal epitaxial silicon is attempted, and it will generally have large grain size. These occurrences are sporadic and undesirable. When polysilicon films are to be used in there own right, a uniform and generally small grain size is required, which usually implies a considerably lower deposition temperature than that used for epitaxy. Further, the applications themselves normally require relatively low-temperature depositions in order to minimize diffusion and /or wafer damage. Typically, low-pressure tube depositions with silane as the silicon source are now used. An LPCVD reactor provides for a very economical deposition and silane decomposition proceeds at temperatures as low as 400°C. When much thicker polysilicon layers are required-for example, in dielectrically isolated wafers speed of deposition, as well as an inexpensive feedstock is desirable, and the process used is much like that of high-temperature epitaxy.

The deposition rate of LPCVD polysilicon depends on temperature, silane partial pressure, and whether or not doping gas such as phosophine, arsine or diborane is present.

B.2 Flow Chart for Polysilicon Deposition

P-clean

5:1 H_2SO_4:H_2O_2

110°C

10 minutes

↓

Rinse Hot DI

10 minutes

↓

Rinse Cold DI

5 minutes

↓

Spin Dry

↓

Deposit Polysilicon

Target thickness = 3200 Å

Table B.1 Polysilicon deposition conditions

$Si H_4$	300sccm
Pressure	400m Torr
Temp	600°C
Time	2hours 45 min

B.3 Flow Chart for Polysilicon Doping

P-clean

5:1 H_2SO_4:H_2O_2

110°C

10 minutes

↓

Rinse Hot DI

10 minutes

↓

Rinse Cold DI

5 minutes

↓

Spin Dry

↓

Phosphorus Diffusion

APPENDIX C

SPUTTERING

C.1 Sputtering

While evaporation is widely used to deposit aluminum and its alloys, sputtering of these materials has become more practical because of the increased deposition rate and more uniform step coverage and contact hole filling. Also, when refractory metals are used, either sputtering or CVD is required to achieve realistic throughput in manufacturing.

Sputter is similar to a billiard-ball event. Ions are accelerated in an electric field toward a target of material to be deposited, where they 'knock-off' (sputter) target atoms. The sputtered ions then deposit onto wafers, which are conveniently placed facing the target. Argon is typically used for sputtering because it is inert and readily available in pure form. It is ionized by colliding with high electrons in the chamber, and then accelerated in an electric field toward the negatively biased target. The momentum of ions incident on the target is then transferred to surface atoms of the target material, causing their ejection. Therefore, during sputter deposition, material is removed from the target and deposited onto the wafers.

The incident energy must be large enough to dislodge target atoms, but not too large to cause penetration into the target material (ion implantation). Typical sputtering ion energies range from 500-5000 eV. The number of atoms sputtered from the wafer surface per incident ion is defined as the sputter yield. This number varies from 0.5-1.5, depending on the momentum of ions and their angle of incidence.

There are many types of sputtering. They are :

a) Ion beam sputtering

b) Magnetron sputtering

c) Reactive sputter deposition

d) Bias sputter deposition

C.2 Flow Chart Showing the Sputter Deposition Process

P-clean

5:1 H_2SO_4:H_2O_2

110°C

10 minutes

↓

Rinse Hot DI

10 minutes

↓

Rinse Cold DI

5 minutes

↓

Spin Dry

↓

Deposit Aluminum

↓

Measure sheet resistivity

Table C.1 Sputter deposition parameters

Parameters	Values
Base pressure	$8 * 10^{-7}$ torr
Temperature	300°C
Deposition rate	10 Å/min
Thickness	10000 Å

APPENDIX D

PHOTOLITHOGRAPHY

D.1 Introduction

The basic steps of the lithographic process are shown in the figure. The photoresist (PR) is applied as a thin film to the substrate(e.g. SiO_2 on Si), and subsequently exposed through a mask. The mask contains clear and opaque features that define pattern to be created in the PR layer. The areas in the PR exposed to light are made either soluble or insoluble in a specific solvent known as developer. In the case when irradiated (exposed) regions are soluble, a positive image of the mask is produced in the resist. Such material is therefore termed a positive photoresist. On the other hand, if the non irradiated regions are dissolved by the developer, a negative image results. Hence the resist is termed a negative resist. Following development, the regions of SiO_2 no longer covered by resist, are removed by etching, thereby replicating the mask pattern in that oxide layer.

The resist is seen to perform two roles in this process. First, it must respond to exposing radiation in such a way that mask image can be replicated in the resist. Second, the remaining areas of resist must protect the underlying substrate during subsequent processing. In fact the name resist evolved from the ability to resist etchants.

Although both negative and positive resists are used to manufacture semiconductor components, the higher resolution capabilities of positive resists have virtually made them exclusive choice for VLSI applications. Conventional positive optical lithographic processes and resists are capable of producing images on VLSI substrates

with dimensions as small as 0.8-1.5μm. For submicron features, however, diffraction effects during exposure may ultimately cause other higher resolution techniques to replace optical lithography

D.2 Description of Lithographic Process

The first step in this process is the application of photoresist. The photoresist is applied by spin coating technique. This procedure involves three stages : a)dispensing the resist solution onto the wafer; b) accelerating the wafer to the final rotational speed; and c) spinning at a constant speed to establish the desired thickness (and to dry the film).

The dispensing stage can either be accomplished by flooding the entire wafer with resist solution before the beginning the spinning, or by dispensing a smaller volume of resist solution at the center of the wafer and spinning at lower speeds to produce a uniform liquid layer across the wafer.

In the next stage the wafers are normally accelerated as quickly as is practical to the final spin speed and finally spinning at the constant speed to obtain desired thickness.

In this work the rotational speed was maintained at 1500 rpm. The wafers were spun at this speed for a period of 20 seconds. This gave a photoresist of 2μm thickness. Prior to exposure the wafers are baked for one minute at 115°C or for a period of 20 minutes in an air oven maintained at the same temperature in order to remove the moisture present. The baking is done to remove the moisture from the wafer. Moisture can reduce the adhesion.

The exposure was carried out using a SUSS MA6 mask aligner. The lamphouse is equipped with a 350 W mercury high pressure lamp and a SUSS diffraction reducing

optics. The usable wavelength falls between 350-450 nm. The lamphouse has an ellipsoidal mirror, and a 45° cold light mirror. The type of exposure lamp depends on the optical range selected. The cold light mirror reflects the desired short wavelength UV light through a fly's lens and transmits the longer wavelengths to a heat sink located in the bottom of the lamphouse. The lamphouse also contains a condenser lens, diffraction reducing lens plates, a 45° turning mirror and a collimation lens. A holder is provided in the mirror house for a filter. SUSS diffraction reducing exposure system provides a high resolution over the entire exposure area, resulting in steep resist edges and small diffraction effects.

The wafers were exposed for 20 seconds in the SUSS MA6 mask aligner. After the exposure the wafer must undergo "development" in order to leave behind the image which will serve as a mask for etching. The developer is poured on the wafer and allowed to develop for 30 second before it is spun at a high rpm(~2000-3000). This procedure is repeated for another 10 seconds of development. The wafer is then washed with DI water and spun again to remove all the water. The wafers are baked as before to rid of the moisture that may been absorbed by the substrate. The windows are then inspected under the microscope for their integrity.

APPENDIX E

ETCHING

E.1 Introduction

In general etching process is not completely attainable. That is etching process are not capable of transferring the pattern established by protective mask into the underlying material. Degree to which the process fail to satisfy the ideal is specified by two parameters: bias and tolerance. *Bias* is the difference between the etched image and mask image. *Tolerance* is a measure of statistical distribution of bias values that characterizes the uniformity of etching.

The rate at which material is removed from the film by etching is known as etch rate. The units of *etch rate* are Å/min, μm/min, etc. Generally high etch rates are desirable as they allow higher production throughputs, but in some cases high etch rates make the control of lateral etching a problem. That is since material removal can occur in both horizontal and vertical directions, the horizontal etch rate as well as vertical etch rate may need to be established in order to characterize an etching process. The lateral etch ratio, L_R, is defined as the ratio of the etch rate in a horizontal direction to that in a vertical direction. Thus:

$$L_R = \frac{\text{Horizontal etch rate of material}}{\text{Vertical etch rate of material}}$$

In the case of an ideal etch process the mask pattern would be transferred to the underlying layer with a zero bias. This would then create a vertical edge of the mask.

Therefore the lateral etch rate would also have to have been zero. For nonzero L_R, the film material is etched to some degree under the mask and this effect is called undercut. When the etching proceed in all directions at the same rate, it is said to be *isotropic*.

By definition, however, any etching that is not isotropic is anisotropic. If etching proceeds exclusively in one directions (e.g. only vertically), the etching process is said to be completely *anisotropic*. A typical anisotropic etch profile is shown in the figure.

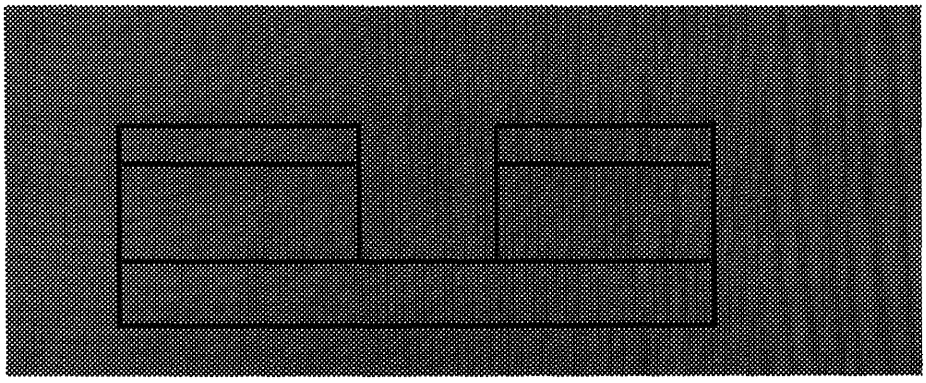

Figure E.1 Anisotropic etch profile

So far it has been assumed that the mask is not attacked by the etchant and did not consider that the layers under the etched film can also be attacked by the etchant. In fact both the mask and the underlying layer materials are generally etchable, and these effects may play a significant role in specifying etch processes. The underlying material subject to attack may either be the silicon wafer itself, or a film grown or deposited during a previous fabrication step. The ratio of the etch rates of different materials is known as *selectivity of an etched process*. Thus both the selectivity with respect to the mask material ands the selectivity with respect to the substrate materials are important characteristics of an etch process.

E.2 Wet Etching Technology

Wet etching processes are generally isotropic. They are inadequate for defining features less than 3μm. Nevertheless for those processes that involve patterning of linewidths greater than 3μm, wet etching continues to b a viable technology.

The reason wet etching has found widespread acceptance in microelectronic fabrication is that it is a low cost, reliable, high throughput process with excellent selectivity for most etch process with respect to both mask and substrate materials.

In general wet etch processes can be broken down into three steps:

- diffusion of the reactant to the reacting surface
- reaction
- diffusion of reaction products from the surface.

The second step can obviously be further differentiated into adsorption prior to, and desorption subsequent to, the actual reaction step. The slowest of the steps will be rate controlling. That is, the rate of that step will be the rate of the overall reaction.

Chemical etching can occur by several processes. The simplest involves dissolution of the material in a liquid solvent without any change in the chemical nature of the dissolved species. Most etching process, however, involve one or more chemical reactions. Various types of reactions my take place, although one commonly encountered in semiconductor fabrication is oxidation-reduction (redox). That is, a layer of oxide is formed, then the oxide is dissolved and the next layer of oxide is formed, etc.(e.g. wet etching of Si and Al)

In semiconductor applications, wet etching is used to produce patterns on the silicon substrate or in thin films. A mask is typically used to protect desired surface regions

from the etchant and this mask is stripped after the etching has been performed. Thus, when choosing a wet etch process, in addition to selecting an etchant, a suitable masking material must be picked to have good adhesion to the underlying films, good coating integrity and ability to withstand attack by etchant. Photoresist is the most commonly encountered masking layer, but sometimes it falls short in this role. Problems involved include loss of adhesion at the edge of the mask-film interface due to etchant attack, and large area failure of the resist. Large area failures of the resist are usually due to differential stress buildups in the substrate and mask layers. Also bubble formation during etching process can lead to poor pattern definition, particularly at the pattern edges.

E.3 Dry Etching

Wet etching processes are typically isotropic, therefore if the thickness of the film being etched is comparable to the minimum pattern dimension, undercutting due to isotropic etching becomes intolerable. One alternative pattern transfer method that offers the capability of non isotropic(or anisotropic) etching is "dry etching". As a result, considerable effort has been expended to develop dry etch processes as replacements for wet etch processes.

The overall goal of an etch process, as mentioned earlier, is to be able to reproduce the features on the mask with fidelity. This should be achievable together with control of following aspects of etched features:

- the slope of the feature sidewalls(e.g. the slope of the sidewalls of the etched feature should have the desired angle, in some cases vertical)
- the degree of undercutting(i.e. usually the less undercutting the better)

There are a variety of dry etch processes. The mechanism of etching in each type of process can have a physical bias(e.g. glow-discharge sputtering), a chemical bias(e.g. plasma etching), or a combination of the two(e.g. reactive ion etching, RIE, and reactive ion beam etching RIBE).

In processes that rely predominantly on the physical mechanism of sputtering(including RIBE), the strongly directional nature of the incident energetic ions allows substrate material to be removed in a highly anisotropic manner(i.e. essentially vertical etch profiles are produced). Unfortunately such material mechanisms are non selective against both masking material and materials underlying the layers being etched. That is, the selectivity depends largely on sputter yield differences between materials. On the other hand purely chemical mechanisms for etching can exhibit very high selectivity against both mask and underlying substrate material. Such purely chemical etching mechanisms, however, typically etch in an isotropic fashion.

By adding a physical component to a purely chemical etching mechanism, however the shortcomings of both sputter based and purely chemical dry etching process can be surmounted. Plasma etching process is a purely chemical process and reactive ion etching processes is a physical/chemical process.

The basic concept of plasma etching is rather direct. A glow discharge is utilized to produce chemically reactive species from a relatively inert molecular gas. The etching gas is chosen so as to generate species, which react chemically with the material to be etched, and whose reaction product is volatile. An ideal dry etch process based solely on chemical mechanisms for material removal, can thus be broken down into six steps:

- reactive species are generated in a plasma

- these species diffuse to the surface of the material being etched
- the species are adsorbed on the surface
- a chemical reaction occurs with the formation of a volatile by product
- the by product is desorbed from the surface
- the desorbed species diffuse into the bulk of the gas

If any of these steps fail to occur, the overall etch process ceases. Many reactive species can react rapidly with a solid surface, but unless the product has a reasonable vapor pressure so that desorption occurs, no etching takes place. Reactive ion etching as described before is an anisotropic etching technique. After the lithographic step, windows are formed on the photoresist layer. The silicon nitride is exposed in these windows. RIE is carried out to remove this silicon nitride and expose the underlying silicon substrate.

E.4 Description of the Reactor for RIE

Plasma etching systems consist of several components: a) an etching chamber, that is evacuated to reduced pressures; b) a pumping system for establishing and maintaining the reduced pressure; c) pressure gauges to monitor the pressure in the chamber; d) a variable conductance between the pump and etching chamber so that the pressure and flow rate in the chamber can be controlled independently; e) an RF power supply to create the glow discharge; f) a gas handling capability to meter and control the flow of reactant gases; and g) electrodes. There are several types of commercially available etching systems. They include

1. barrel reactors
2. "downstream" etchers

3. parallel-electrode(planar) reactor etchers
4. stacked parallel-electrode etchers
5. hexode batch etchers
6. magnetron ion etchers

The RIE system in the class 10 clean room, where the fabrication of the V grooves was carried out, is a stacked parallel electrode etching system.

The stacked parallel electrode etcher is a small batch machine capable of handling 6 wafers at a time. Its unique design provides an individual pair of electrodes for each wafer thereby combining some of the advantages of a single wafer and batch etchers. Operating chamber pressures and RF power densities can be kept in the ranges between those of low pressure, low power density hexode batch etchers, and high pressure.

APPENDIX F

SIMULATION

F.1 Pspice Simulation

The following is the code written for the simulation of the shift register circuit using Pspice.

```
* shift register circuit
vin 1 0 pulse (0v 5v 0us .001us .001us .05ms .1ms)
vdd 9 0 dc 5.0volt
vsw1 2 0 pulse (0v 5v 0us .001us .001us .005ms .01ms)
vsw2 11 0 pulse (5v 0v 0us .001us .001us .005ms .01ms)
c1 3 0 .015p
c2 12 0 .015p
m1 3 2 1 0 nmos1 l=4u w=20u
m2 5 3 0 0 nmos1 l=4u w=20u
m3 9 9 5 0 nmos2 l=8u w=4u
m4 12 11 5 0 nmos1 l=4u w=20u
m5 15 12 0 0 nmos1 l=4u w=20u
m6 9 9 15 0 nmos2 l=8u w=4u
.model nmos1 nmos (level=2 vto=1.1)
.model nmos2 nmos (level=2 vto=1.1)
```

```
.tran .05m 1m
.op
.plot tran v(1) v(15)
.op
.probe v(1) v(15)
.end
```

F.2 Supreme Simulation

The following is the code used to simulate the value of threshold voltage using Supreme simulation.

```
Title SUPREM-III Example 1. NMOS Silicon Gate
Comment        Active Device region initial processing.
$ File S3EX1A
Comment Initialize silicon substrate.
Initialize <100> Silicon, Boron Concentration=1e16
  + Thickness=15 dx=.005 xdx=.02 Spaces=150
Comment Grow sacrificial oxide.
Diffusion Temperature=950 Time=30 Weto2
Print Layer
Comment Implant boron to shift the threshold voltage.
Implant Boron Dose=2e12 Energy=60
Comment Etch sacri oxide
```

```
Etch oxide all

Comment Grow gate oxide.

Diffusion Temperature=950 Time=45 Dryo2 HCL%=3

Print Layer

Comment Deposit Polysilicon.

Deposit Polysilicon Thickness=0.5 Temperature=600

Comment Heavily dope the polysilicon using POC13

Diffusion Temperature=1000 Time=25 dTmin=.3

  + Phosphorous Solidsolubility

Print Layer

Comment Calculating Vt now

V.Threshold V.sub1=0 V.sub2=2 DV.Sub=0.1 Bulkconc=1E18 P-type Q.F=1E11

END.Electrical

Plot Chemical Phosphorous Xmax=1.5 ^Clear ^Axis Linetype=3

Plot Chemical Net Xmax=1.5 ^Clear Axis

Print layer

Comment Save the structure at this point.  The simulation runs

$ are split for the gate and source/drain regions.

save Structure File=s3elas

Stop End of Suprem-III Example 1.
```

REFERENCES

1 S. Wolf and R.N. Tauber, *Silicon Processing for the VLSI Era Volume 1-Process Technology* Lattice Press, Sunset Beach, California, 1986.

2 S. Wolf and R.N. Tauber, *Silicon Processing for the VLSI Era Volume 2-Process Integration* Lattice Press, Sunset Beach, California, 1986.

3 Badih El-Kareh, *Fundamentals of Semiconductor Processing Technologies* Kulwar Academic Publishers, Norwell Massachusetts, 1995.

4 Neil Weste, Kamran Eshraghian, *Principles of CMOS VLSI Design* Addison-Wesley Publishing Company, Reading, Massachusetts, 1988.

5 J. Mavor, M.A. Jack and P.B. Denyer, *Introduction to MOS LSI Design* Addison-Wesley Publishing Company, Reading, Massachusetts, 1983.

6 Stanley G. Burns, Paul R. Bond, *Principles of Electronic Circuits* West Publishing Company, St. Paul, Minnesota, 1987.