STATUS OF THESIS

Title of thesis	Optimization of Rapid Thermal Processing in 0.13 micron CMOS Poly Silicon Gate

Hereby allow my thesis to be placed at the Information Resource Center (IRC) of Universiti Teknologi PETRONAS (UTP) with the following conditions:

1. The thesis becomes the property of UTP
2. The IRC of UTP may make copies of the thesis for academic purposes only.
3. This thesis is classified as

☐ Confidential

☒ Non-confidential

If this thesis is confidential, please state the reason:

____

____

____

The contents of the thesis will remain confidential for __________ years.

Remarks on disclosure:

____

____

____

Endorsed by

Signature of Author

Signature of Supervisor

Permanent address : 1873 Kampung Lima Kongsi,
14200 Sungai Jawi, Penang

Date : 24th May 04

Name of Supervisor
Assoc Prof. Dr Norani Muti Mohamed

Date : 24th May 2004

APPROVAL PAGE

UNIVERSITI TEKNOLOGI PETRONAS

Approval by Supervisor (s)
The undersigned certify that they have read, and recommend to The Postgraduate Studies Programme for acceptance, a thesis entitled "**Optimization of Rapid Thermal Processing in 0.13 micron CMOS Poly-silicon Gates** " submitted by **(Chew Soon Aik)** for the fulfilment of the requirements for the degree of **masters of science in electrical and electronics engineering**.

24th May 04
Date

Signature :
Main Supervisor : Assoc. Prof. Dr Norani Muti Mohamed
Date : 24th May 2004.
Co-Supervisor 1 :

TITLE PAGE

UNIVERSITI TEKNOLOGI PETRONAS

Optimization of Rapid Thermal Processing in 0.13 micron CMOS Poly Silicon Gate

By

Chew Soon Aik

A THESIS

SUBMITTED TO THE POSTGRADUATE STUDIES PROGRAMME

AS A REQUIREMENT FOR THE

DEGREE OF MASTERS OF SCIENCE IN ELECTRICAL AND ELECTRONICS ENGINEERING

Electrical and Electronics Engineering

BANDAR SERI ISKANDAR,

PERAK

MONTH NOVEMBER, 2003

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations, which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTP or other institutions.

Signature :

Name : Chew Soon Aik

Date : 24th May 04

ACKNOWLEDGEMENT

First of all, I would like to acknowledge my supervisor, Assoc Prof Dr. Norani Muti Mohamed. She has been a great mentor to me with her management skills and enthusiasm. She has also been an exceptional role model of life. She has done her utmost to help relieve concerns that her advisees have. It has been memorable in that she had encouraged and supported me to apply for Silterra Postgraduate Internship while I was in deep depression.

I also would like to thank my co-supervisor, Mr Albert Victor Kordesch, Senior Manager of Device Modeling Department, Silterra. Throughout my association with him, he gave me strong motivation in doing my research. I consider him to be a living encyclopedia that answers almost all my queries. Besides that, his encouragements also make me feel more comfortable and confidence while carrying out my research in Silterra.

I am indebted to the internal and external examiner who will be evaluating my thesis and oral exam. I am also grateful to Assoc Prof Ir. Dr.Ahmad Fadzil Mohamad Hani who has agreed to be the chairman of the UTP oral committee in spite of his busy schedule.

I wish to thank the Silterra (M) Sdn Bhd, pioneer of Malaysia semiconductor industry for the support of this work through postgraduate internship program. Silterra provides every technical support including financial, equipment, research environment and technical advice for my research work. I would like to acknowledge the Device Modeling Dept (Mr. Yusman Mohd Yusof, Ms. Norhafizah Che May, Mr. Philip Tan, Mr. Gilbert See, Mr. Nor Razman Md Zin and Mr. Mohd Fahmi Muhsain), Mr. CO Lim, Mr. Anuar Fadzil Ahmad, Mr. Daniel Oh, Mr. Mohamad Zainizan Zakaria, Ms. Su-Ann Heng, and Mr. KJ Koh. Special appreciation also goes to Mr. Palanivel Balasubramaniam for his advice and approval, Dr HC How and Ms. PL Teh for SIMS profile measurements and FIB scanning, and Ms. KS Lim for her advice on process and procedure. Ms Lena Tan, Mrs. Reha and Mr. Abdullah bin Lin for the non-technical support in Silterra. Lastly, many thanks to Silterra management for their support on the managing of the postgraduate internship program.

ABSTRACT

Complementary Metal-Oxide-Silicon (CMOS) technology has gained a great deal of attention in the decade of VLSI (Very Large Scale Integration) in the 1980s. As the CMOS devices continue to scale down, many problems such as poly depletion effect, stress induced leakage current etc. have arised leading to the reduction in the device performance. The objective of this work is to reduce the poly depletion effect occurring in deep submicron CMOS devices and at the same time suppresses boron penetration through the thin gate oxide. The study focused on optimising the spike annealing process on nitrided gate oxide of the 0.13-µm CMOS devices.

The work began with the simulation on the fabrication process of CMOS devices using different parameters of implant dosage, implant energy, implant angle and heat treatment. After successful simulated outcomes have been obtained, experiments were conducted based on the parameters used in the simulation work. The Source Drain Rapid Thermal processing (SD-RTP) using the spike annealing technique was adopted as the heat treatment process on the polysilicon layer. The samples were then characterized electrically using four-point probe and structurally using Secondary Ion Mass Spectrometry (SIMS) and Focused Ion Beam (FIB).

Results showed that peak temperature has a bigger influence in controlling the dopant activation in the poly gate compared to ramp-up rate. As evident in SIMS profile, boron concentration in the polysilicon that had been subjected to spike annealing at the peak temperature of 1080°C was found to be uniform. Uniform distribution of boron will reduce the poly depletion effect. FIB image showed that polysilicon treated to a higher peak temperature or longer annealing duration, has larger grain size. It is believed that with larger grain size, boron activation is increased due to the reduced ability to absorb interstitials. Moreover, dopant segregation at the grain boundaries is also decreased. As a conclusion for 0.13-µm CMOS devices, spike annealing at an optimum peak temperature of 1080°C is recommended to replace the existing RTP process to reduce the poly depletion effect and at the same time, suppressing the boron penetration.

ABSTRAK

Teknologi CMOS telah beroleh banyak perhatian pada dekad VLSI pada 1980an. Apabila saiz peranti CMOS semakin mengecil, banyak masalah seperti kesan susut poli, arus bocor terangsang dan lain-lain timbul menyebabkan penurunan dalam prestasi peranti. Objektif penyelidikan ini ialah untuk mengurangkan kesan susut poli yang berlaku dalam peranti submikron CMOS dan pada masa yang sama menahan penembusan boron melepasi get oksida yang nipis. Kajian memfokus kepada mengoptimumkan proses sepuhlindap puncak keatas oksida get nitrida bagi 0.13-µm peranti CMOS.

Penyelidikan bermula dengan simulasi keatas proses fabrikasi peranti CMOS dengan menggunakan parameter berbeza bagi dos penanaman, tenaga penanaman, dan sudut penanaman dan rawatan haba. Selepas berjaya memperolehi keputusan simulasi, eksperimen dijalankan berdasarkan kepada parameter yang digunakan dalam kerja simulasi. Pemprosesan terma kejut sumber parit (SD-RTP) dengan teknik sepuhlindap puncak digunakan sebagai rawatan haba keatas lapisan polisilikon. Sampel kemudiannya dicirikan secara elektrik dengan penduga-empat-titik dan secara struktur dengan spektrometer jisim ion sekunder (SIMS) dan bim ion fokus (FIB).

Keputusan menunjukkan bahawa suhu puncak mempunyai pengaruh yang lebih besar keatas pengaktifan pendop dalam get poli berbanding dengan kadar kenaikan sepuhlindap. Terbukti dalam profil SIMS, kepekatan boron dalam polisilikon yang dikenakan sepuhlindap pada suhu 1080°C telah didapati seragam. Taburan boron yang seragam akan mengurangkan kesan susutan poli. Imej FIB menunjukan bahawa polisilikon yang dirawat pada suhu puncak yang lebih tinggi dan tempoh sepuhlindap yang lebih lama, mempunyai saiz butiran yang lebih besar. Dipercayai bahawa dengan saiz butiran yang lebih besar, pengaktifan boron meningkat disebabkan oleh penurunan keupayaan untuk resapan pendop secara selitan. Tambahan pula, pengasingan pendop di sempadan butiran akan turut berkurang. Sebagai kesimpulan untuk peranti CMOS 0.13-µm, sepuhlindah kejut pada suhu puncak optimum 1080°C dicadangkan untuk mengganti proses RTP yang tersedia bagi mengurangkan kesan susutan poli dan pada masa yang sama, menahan penembusan boron.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background of CMOS Devices

Complementary Metal-Oxide-Silicon (hereafter abbreviated as CMOS) technology has acquired a tremendous amount of interest in the decade of Very Large Scaled Integration (hereafter abbreviated as VLSI) in the 1980s. The great interest is because CMOS, when compared with N-type MOSFET (hereafter abbreviated as NMOS), has much lower power, comparable speed, slightly poorer density but better reliability. The major advantage of using CMOS is the low power dissipation. In a VLSI circuit, too many devices exist and they run very fast. As a result, more power is dissipated from each device and even more power is dissipated from a chip. Current packaging technology limits power dissipation at a few watts per chip to avoid high temperature induced reliability problems. Sophisticated packages are being developed, but at a great expense. Even CMOS, with its inherent low power characteristics is now pushing the limits of thermal packaging (Kamins, 1998). The cross-sectional view of the simple CMOS device is shown in Figure 1.1.

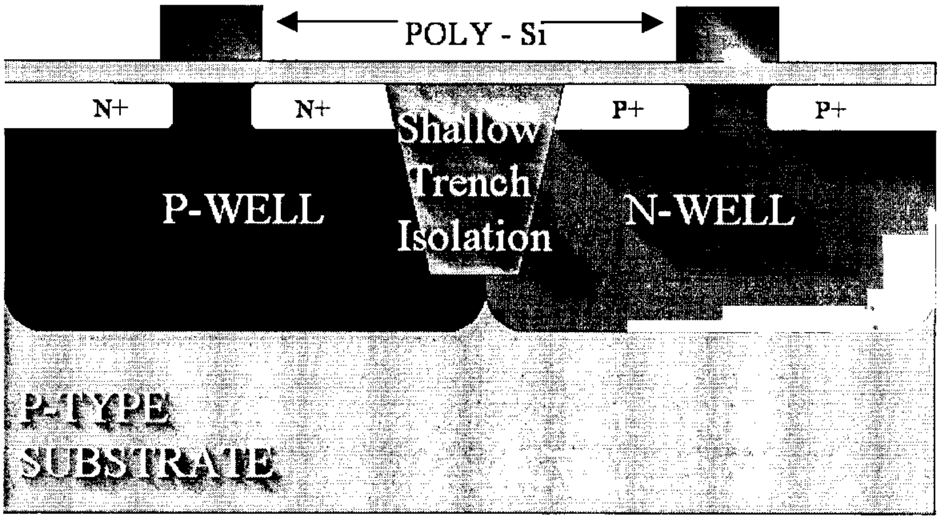


Figure 1.1. Simple CMOS device cross-sectional view.

In the early development, most gate electrodes for MOS were made from aluminum, which was deposited after the source, and drain regions were doped. However, in this case aluminum gate must overlap the source and drain region by at least one alignment tolerance. This can cause a significant Miller feedback capacitance between the gate and the drain resulting in a reduction in the circuit speed. The Miller effect is a way of dealing with a situation where the voltages at both ends of a capacitor change at the same time, either independently or dependently. In certain circuits, instead of taking into account what actually happens with the voltage, it is equivalent to having a larger capacitor. Polysilicon gates have much smaller Miller effect because the source and drain are self-aligned. So, polysilicon is used only after the Chemical Vapor Deposition (hereafter abbreviated as CVD) method has been developed. The ability of polysilicon to be oxidized and its compatibility with further high-temperature processing has allowed development of complex device structures employing several layer of polysilicon. As a conductor, polysilicon can be used to interconnect devices, allowing more flexible layout than obtained with only aluminum interconnections. Polysilicon can also be used for high-value resistors, thus further reducing the chip area (Sze, 1985).

1.2 Objective

As the critical dimensions of semiconductor MOS transistor are continually scaled down to deep-submicron for higher levels of integration and performance, poly-depletion effect and boron penetration through thin gate oxide have became more severe. In order to reduce poly-depletion effect and to suppress boron penetration, researchers have attempted several methods however none had claimed highly successful in reducing these two effects (Bin Yu *et al.*, 1998). On the other hand, having a heavily doped gate can lead to boron penetration resulting in large threshold voltage shift, large charge trapping rate, degradation of p-channel subthreshold slope and poor reliability of the devices (Lurng Shehng Lee *et al.*, 1998). Boron penetration is caused by thermal diffusion of boron through the gate oxide. To reduce boron penetration, source drain rapid thermal processing (hereafter abbreviated as SD-RTP) duration and temperature need to be reduced further. However if the time and temperature is reduced too much, low activation (less than the solid solubility) can arise. Spike annealing is widely used now in the wafer fabrication industries as a replacement for SD-RTP. Spike annealing is a special RTP process with faster ramp-up rate and zero dwell time.

Here, the goal of the research work is to optimize the spike annealing process parameters for the 0.13-µm micrometer channel CMOS devices in Silterra (M) Sdn Bhd. In Silterra, the wafer fabrication factory mainly fabricates chips designed by customers. There is a need to study activation of the dopant in the polysilicon after the heat treatments because higher concentration of activated dopant in the polysilicon can actually reduce the polydepletion effect. At the same time, boron penetration through the thin insulator gate oxide layer has to be prevented.

1.3 Scope of the Work

The work is mainly focused on optimizing the process recipe for 0.13-µm CMOS devices to reduce the polydepletion effect and suppress boron penetration.

In this work, process simulation software is used to get the optimized process recipe for 0.13-μm CMOS devices. The software comprises of 2 types of tools known as TSUPREM 4 for the process simulation and MEDICI for device characterization. An example of the simulation window is shown in Figure 1.2. TSUPREM 4 has been adopted from the SUPREM 3, which was developed at Stanford University in the 1980s. MEDICI is a 2-Dimensional finite element Poisson equation solver developed at Stanford University in the 1980s. This physically based process simulator can predict the outcome that result from specified process sequences. Simulation has been very important for the reason that it is always quicker and cheaper than performing experiments. It also provides information that is difficult or impossible to measure. The simulation software is also based on mathematical calculations such that finer grid (Figure 1.3) is needed for accurate result. After obtaining a better grid, an experiment can be planned and executed. Effect of different process condition in ion implantation, annealing, deposition and etching was studied by using TSUPREM 4. Figure 1.4 shows an example of the editing the implantation process. MEDICI was used to do the device electrical characterization, namely Current-Voltage (hereafter abbreviated as I-V) and Capacitance-Voltage (hereafter abbreviated as C-V) characterization. In simulation, different implantation parameters such as dopant species, implant doses, implant energy, implant tilt and rotation angle can be changed. Different thickness of deposition can also be done, as well as different annealing conditions. Simulation software also have the function to check the cross-sectional view of the devices (Figure 1.5), dopant distribution profile, junction depth, well isolation, sheet resistance and etc. MEDICI can calculate the threshold voltage and drain current. It can also check the electrical flow lines, leakage region, electric field, plot the I-V curve, plot the C-V curve and many more.

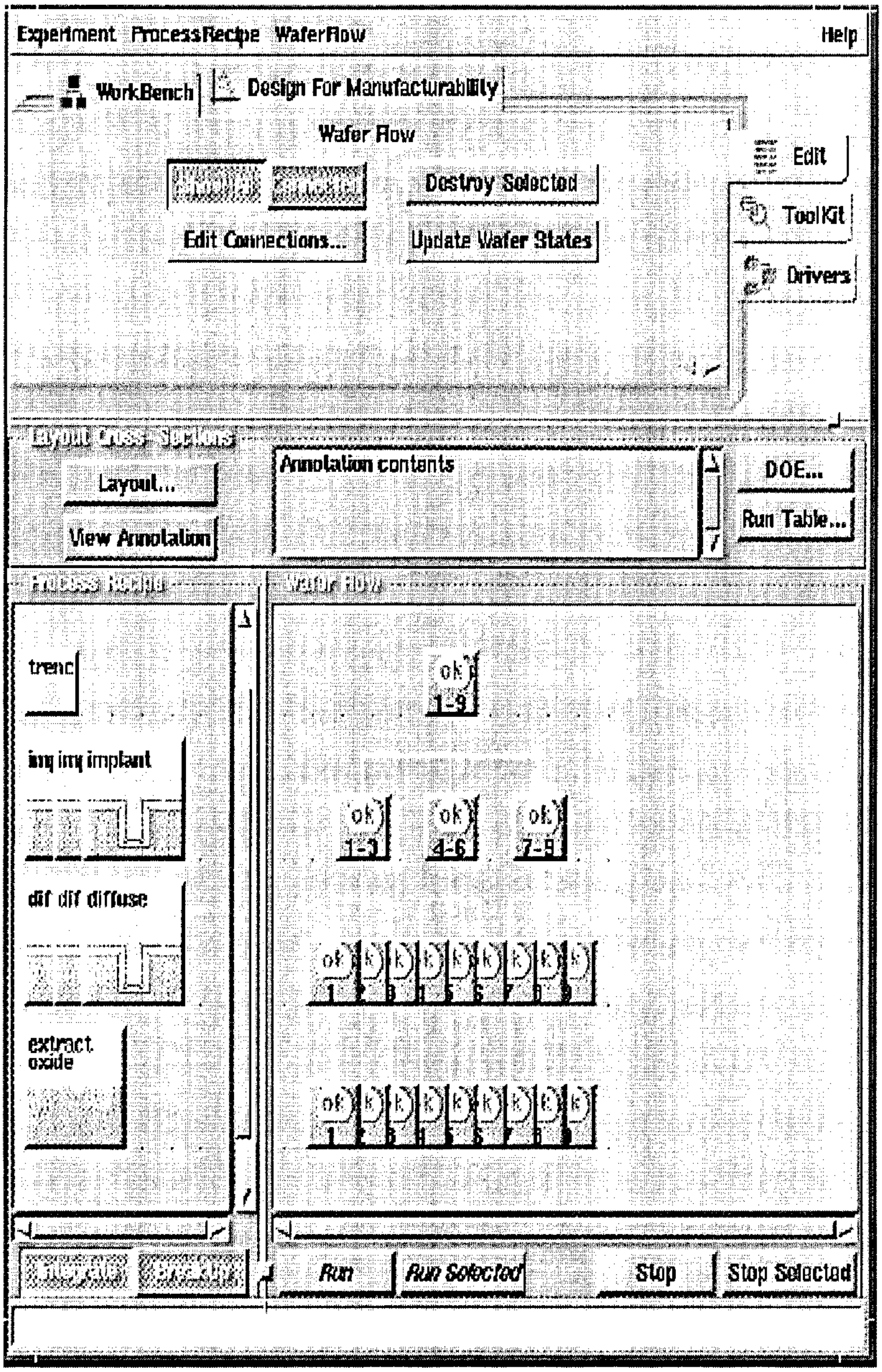


Figure 1.2. An example of simulation window of TSUPREM4.

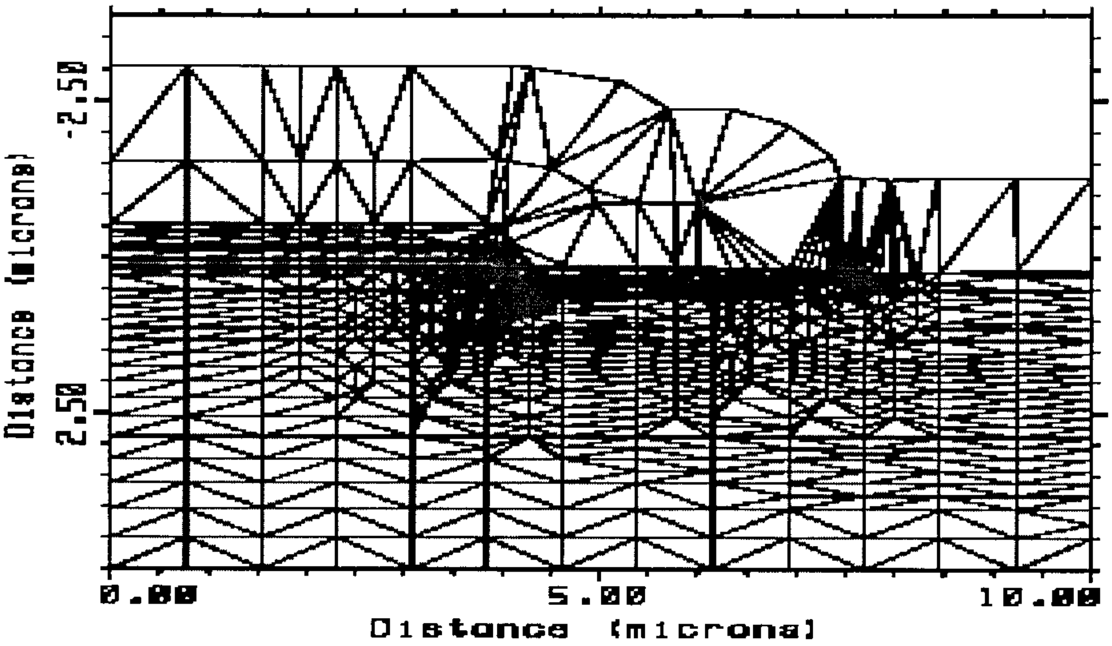
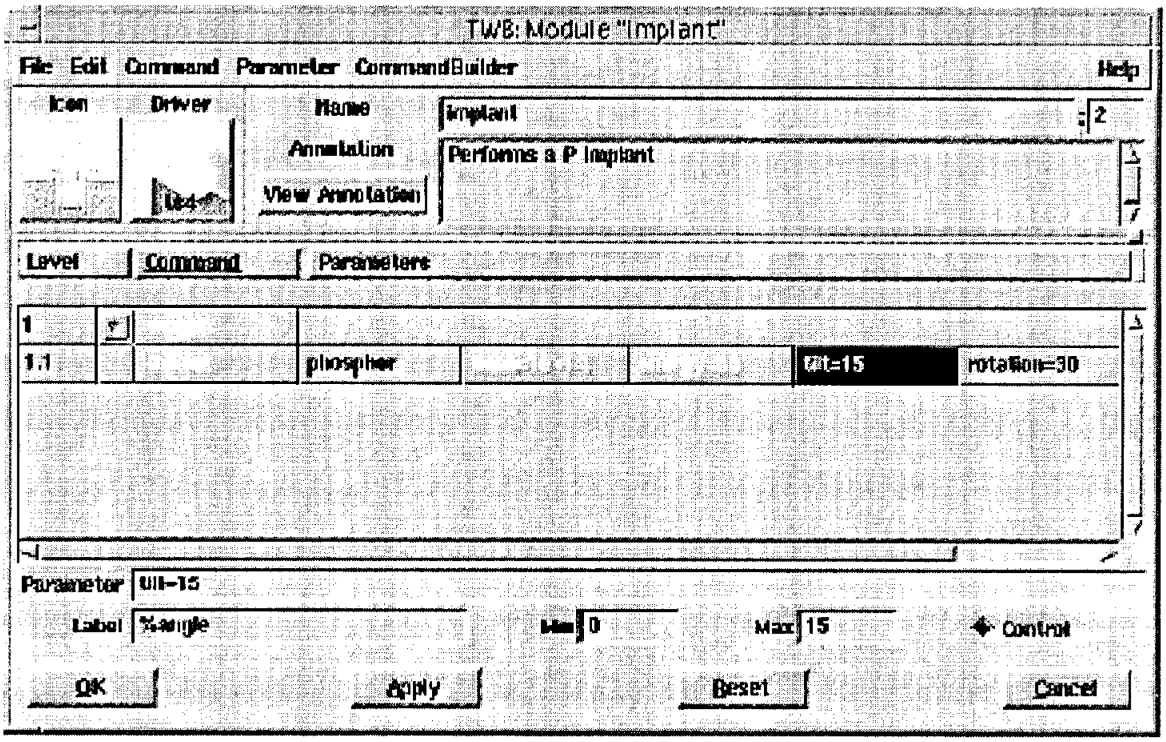


Figure 1.3. An example of finer grid for the model.



Name: Implant 2

Annotation: Performs a P Implant

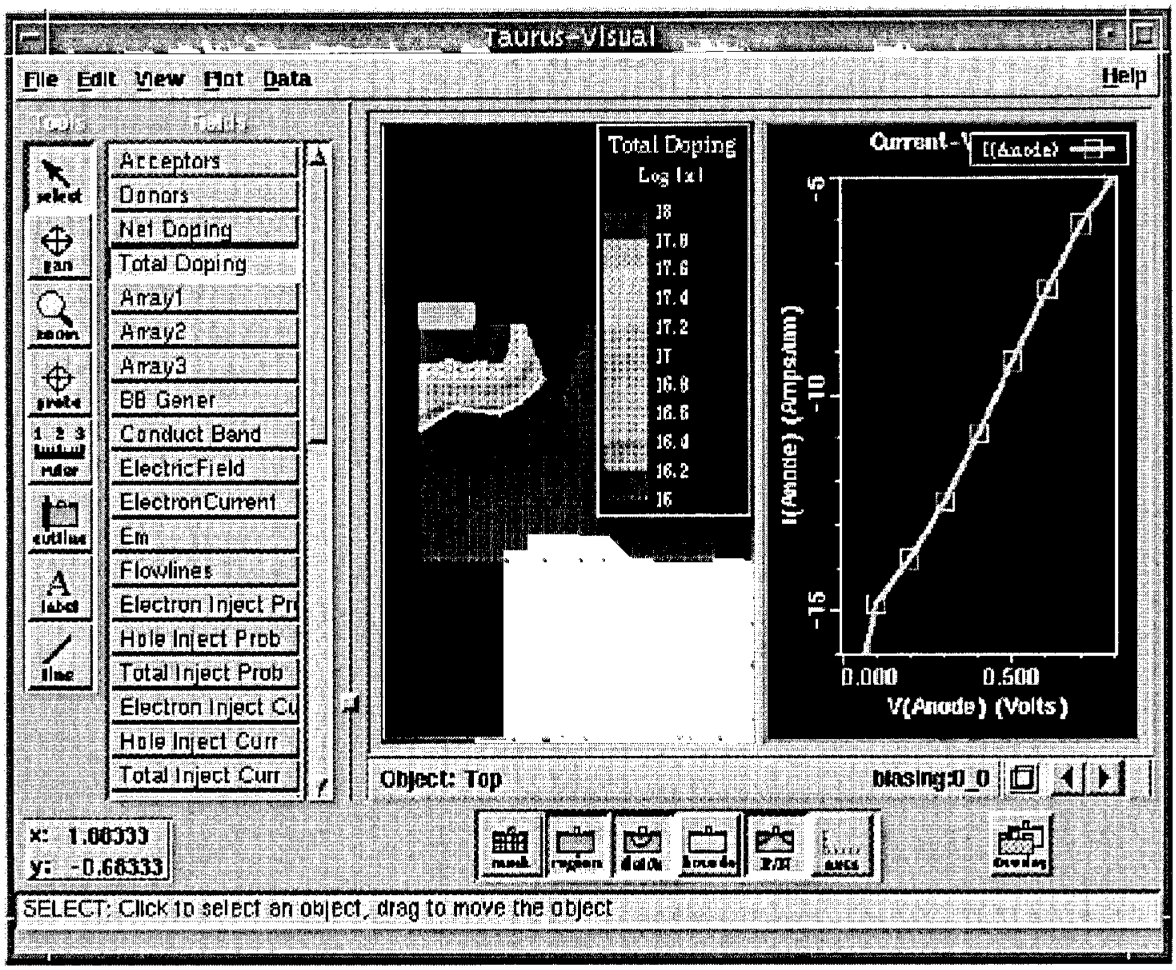
1

1.1 phospher tilt=15 rotation=30

Parameter: tilt=15

Label: %angle Min: 0 Max: 15

Figure 1.4. An example of process editing in implantation step.



x: 1.68333
y: -0.68333

SELECT: Click to select an object, drag to move the object

Figure 1.5. Example of the simulation visual result using Taurus Visual.

After the simulation, an experiment with the optimized process parameters from the results was planned and executed to optimize the process parameters with the objective to reduce polydepletion effect and suppress boron penetration. In this experiment, wafers were run through the processes such as oxidation, implantation, deposition and annealing. Through the experiment, we can get the exact result because some of the new process recipes such as spike annealing, grain boundary diffusion rate and etc have not yet included in the simulation calculation. After that, structural and electrical characterizations were carried out to study the effect of the new process recipe.

For a process recipe to be industrially accepted, a physical experiment needs to be conducted by introducing the new recipe into the process flow of the device fabrication. This involved measurement on 0.13-μm Metal Oxide Silicon Capacitor (hereafter abbreviated as MOSCAP) where C-V can be extracted and depletion thickness can then be calculated.

1.4 Organization of the Thesis

This thesis is organized into five chapters.

Chapter 1 provides an introduction to the CMOS devices, motivation of this work, and an outline of the thesis.

Chapter 2 focuses on the problems encountered during the process of downscaling CMOS. A review of other works in reducing the polydepletion effect and boron penetration is also given in this chapter.

Chapter 3 describes the theory of the processes and characterization techniques. Processes such as oxidation, ion implantation, diffusion/annealing and deposition are explained. Characterization techniques like Secondary Ion Mass Spectrometry (hereafter abbreviated as SIMS), Focused Ion Beam (hereafter abbreviated as FIB) and Four-Point Probe (FPP) are the three characterization techniques used in this work.

In chapter 4, the experiment procedure for this work has been elaborated. This chapter will cover all the experimental processes involved, equipments utilized and characterization test conducted.

Chapter 5 focuses on the result, discussion and conclusion. In this chapter, the simulation and experiment results are given. Here, interpretation is made based on the observation of the structural and electrical analysis. Conclusions of this work and recommendation process parameter made to Silterra are also described in this chapter.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1 Gate Poly Depletion and Boron Penetration

Present deep submicron CMOS technologies including the 0.13-μm modes always use polysilicon as gate material. The polysilicon should be degenerately doped; however, due to processing circumstances, insufficient doping concentration and/or insufficient activation can lead to the formation of a depletion layer in the gate, causing a non-negligible voltage drop. This voltage drop in the gate reduces the effective gate voltage available to create inversion in the channel, and therefore reduces the device performance.

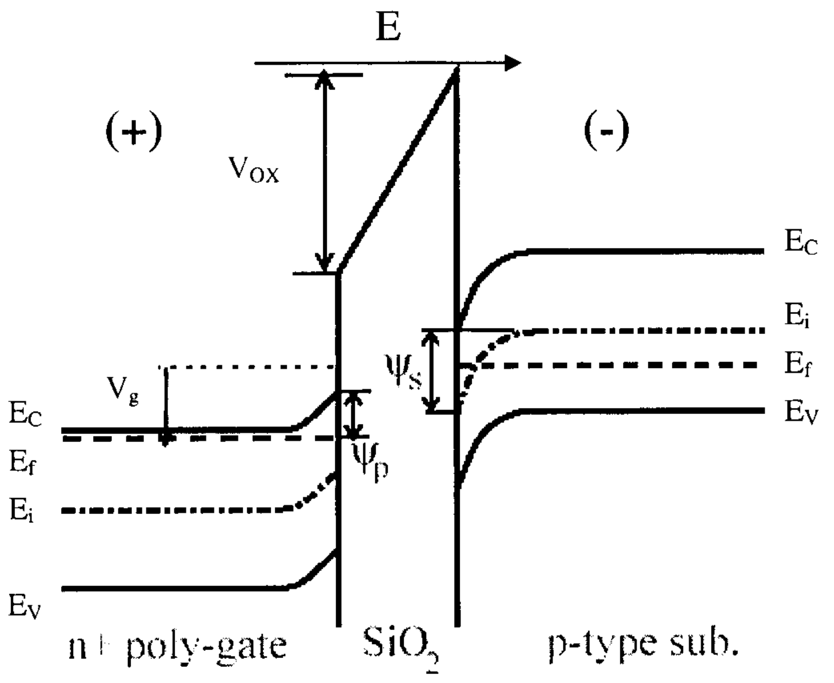


Figure 2.1. Band diagram of a poysilicon-gate depleted N-channel MOS capacitor (Reproduced from Choi Changhoon, 2002)

All device characteristics, i.e. charge/transcapacitance and drain current/transconductance are affected especially when using thin oxide thickness. Consider the band diagram of an n$^+$-polysilicon gated n-channel MOS structure biased into inversion as shown in Figure 2.1. Since the oxide field is in the direction, which attracts negative charge toward the gate, the bands in the n$^+$-polysilicon bend slightly upward near the surface. The field depletes the surface of electrons and forms a thin space-charge region in the polysilicon layer, which lower the total capacitance as given by:

$$\frac{1}{C} = \frac{1}{C_{OX}} + \frac{1}{C_{Si}} + \frac{1}{C_{dep}} \qquad \text{(Eq 2.1)}$$

Where C_{OX} = Oxide capacitance

C_{Si} = Silicon substrate capacitance

C_{dep} = Depletion capacitance

Typical low-frequency C-V curves in the presence of this gate depletion effect are shown in Figure 2.2. The maximum inversion capacitance is less than C_{OX}, and the higher the polysilicon doping concentration the less pronounced is the gate depletion effect.

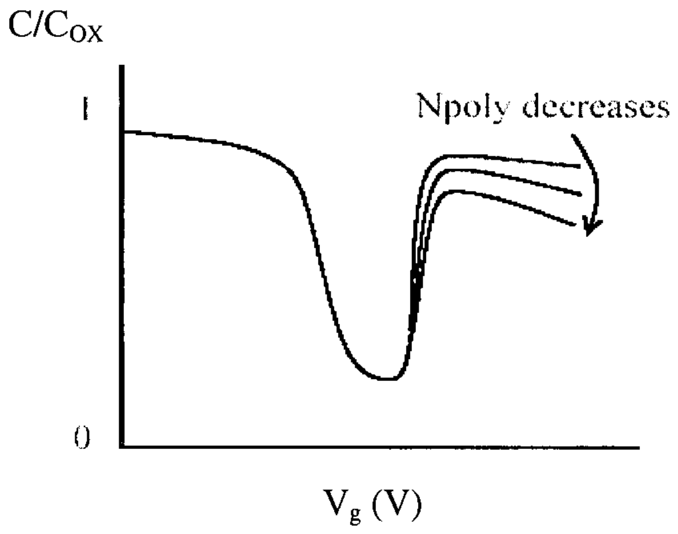


Figure 2.2. C-V curves of N-channel MOS capacitors in the presence of polysilicon depletion effects. (Reproduced from Choi Changhoon, 2002)

It has been understood that increasing the poly doping will reduce the poly-depletion effect, but the doping is limited by other considerations, such as boron penetration through the thin gate oxide and dopant uniformity. The work reported by Choi Chang-Hoon (2002) showed how the dopant profile and gate geometric influence the polydepletion effect.

The potential drop (V_p) in the presence of the poly-depletion effect is represented in Figure 2.3. When the gate is doped by ion implantation, the dopant distribution as a function of position from the top of gate to the gate oxide direction can be represented as in Figure 2.4. Since the dopant density and carrier concentration vary with position, a built-in electric field exists leading to built-in depletion. Under thermal equilibrium, since there is no current flow, the built-in field (E_x) in a p-type poly-gate can be expressed as (Muller and Kamins, 1986):

$$E_x = -\frac{d\phi}{dx} \approx -\frac{kT}{q}\frac{1}{N_A}\frac{dN_A}{dx} \qquad d\phi = \frac{kT}{q}\frac{dN_A}{N_A} \qquad \text{(Eq. 2.2)}$$

where N_A = active dopant concentration
q = electron charge
k = Boltzman constant
T = absolute temperature

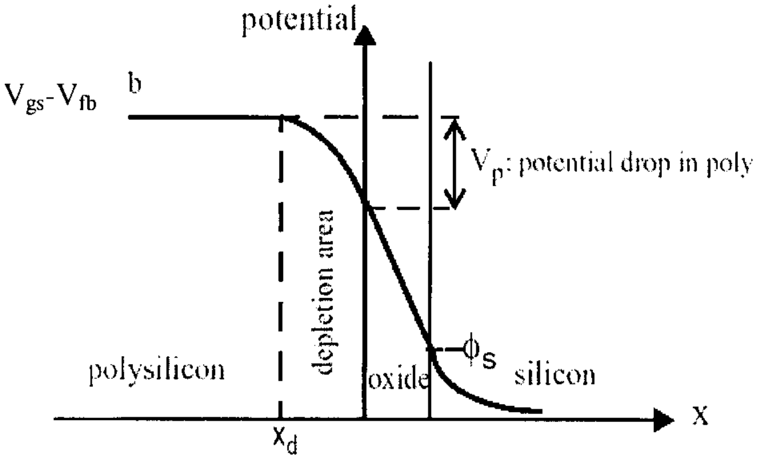


Figure 2.3. Potential distribution of MOS from the top of the poly-gate to the silicon substrate considering poly-depletion effects. (Reproduced from Choi Changhoon, 2002)

A potential drop (ΔV_{pl}) is established at the interface due to the graded dopant distribution between x_1 and x_2, which can be approximated as:

$$\Delta V_{pl} = \phi_1 - \phi_2 = \frac{kT}{q}\ln\frac{N_{A1}}{N_{A2}} \qquad \text{(Eq. 2.3)}$$

When N_A changes from 10^{20} to 10^{18} cm^{-3} (from SIMS profile for spike anneal at peak temperature 1000°C), the potential drop is about 0.12 V from equation (2.3) even without any gate bias, which is on the order of the threshold voltage for sub-micron MOSFETs. This additional potential drop across the polysilicon gate (ΔV_{pl}) does not contribute to gate bias, so it reduces transistor current.

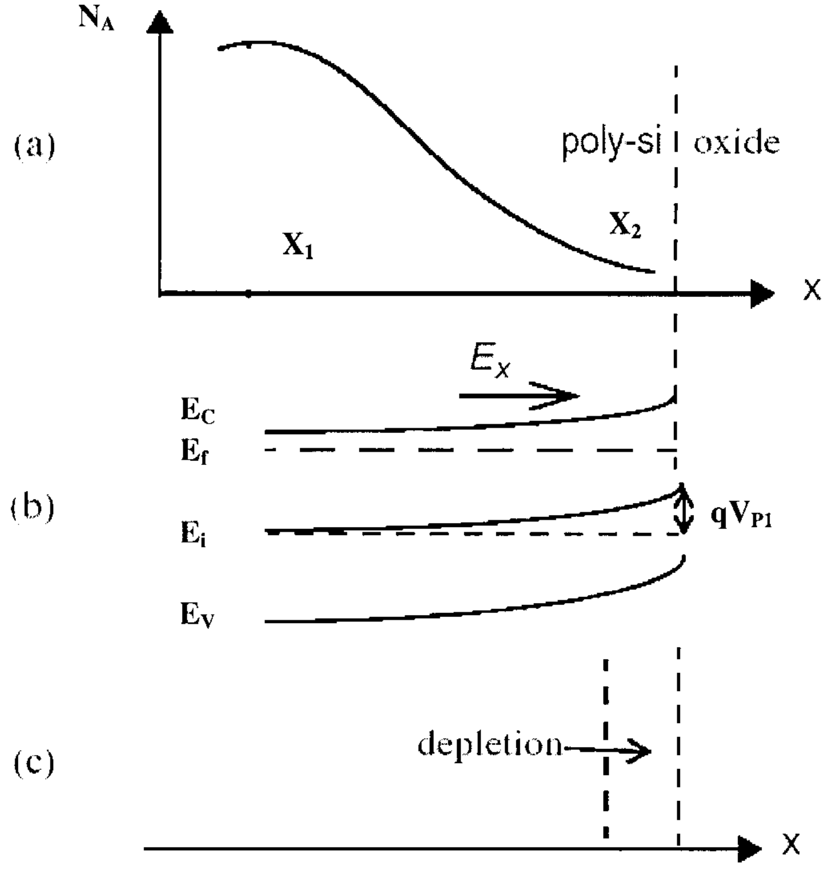


Figure 2.4. Non-uniform, graded impurity distribution and corresponding energy-band diagram (Reproduced from Choi Changhoon, 2002)

For boron-implanted polysilicon subjected to RTP at 960-1150°C, the carrier mobility is higher than after furnace annealing at the same temperatures. This finding is explained by a nonuniform impurity distribution inside grains and by the related decrease of carrier scattering at the grain boundaries. At the lowest temperature (840°C) the impurity peak does not change from the position in the as-implanted process, a process sandwiched between ion implantation and annealing. At higher temperatures, diffusion of impurity atoms into polysilicon would be initiated. At 1080°C, impurity profiles become flat throughout the polycrystalline layer. The uniform doping level is retained up to 1200°C. This is the temperature range where the experimental carrier concentration fits the clustering model. It also means that at that temperature range, the average grain size is large enough to minimize segregation at grain boundaries (Borisenko *et. al*, 1997).

The dopant distribution in the poly is complicated by the implant and heat treatment steps, and by the poly grain size. So, various studies are needed to understand this problem.

Boron penetration becomes a major issue in the fabrication of deep submicron PMOS device with P^+ poly doping and thin gate oxide. The boron penetration can cause the shifts of flatband voltage (V_{FB}), the distortion of C-V curves, the increase of subthreshold swing (S) and leakage current, and the deterioration of the gate oxide quality. Therefore, boron penetration should be eliminated to keep the desired performance of the surface-channel device. The boron penetration effect is characterized by the amount of flat band shift in the C-V curves of a MOS capacitor. Figure 2.5 shows the flat band voltage increases (shifts to the right side) when boron penetration increases.

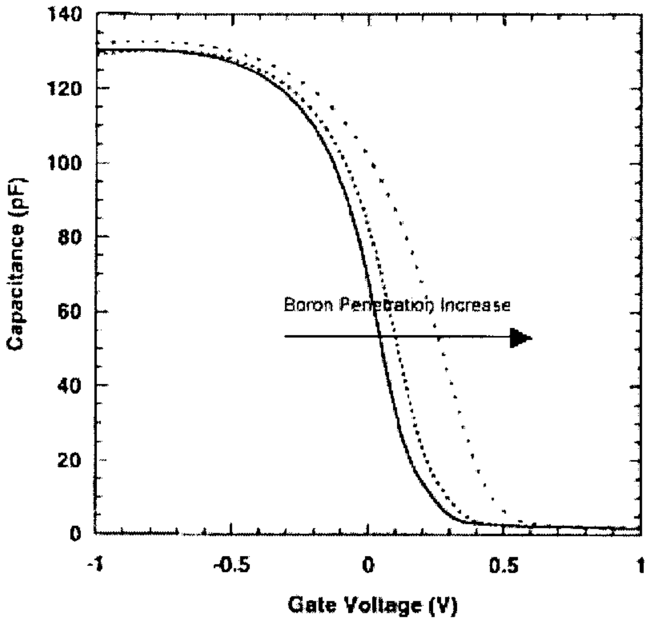


Figure 2.5. High-frequency C-V curves of a MOS Capacitor for different boron penetration level. (Reproduced from Cao Min *et. al*, 1998)

2.2 Methods Attempted by Other Researchers

As silicon devices are continuously scaled to smaller sizes, gate oxides becomes thinner, polydepletion effect and boron penetration effect have become more severe. Polysilicon, although is easy to manufacture, it also causes a lot of problems to CMOS devices. Researchers have been trying to find alternative gate materials to replace polysilicon. Poly-$Si_{1-x}Ge_x$ has been introduced into Thin Film Transistor (hereafter abbreviated as TFT) devices since early 90s. Lee Wen-Chin *et. al.* (1998) and Yu Bin *et. al.* (1998) mentioned that with the germanium content of ≈ 20%, both PMOS and NMOS have the least poly depletion effect. It is because of the higher active dopant concentration observed. In these papers, the authors used conventional Low Pressure Chemical Vapor Deposition (hereafter abbreviated as LPCVD) system to grow poly-$Si_{1-x}Ge_x$. The results obtained are as follows;

1. For both types of devices, the thinner the gate oxide, the more will be the capacitance reduction in the inversion region. This is because a higher electric field will cause more depletion near the poly/SiO_2 interface.

2. For NMOS devices, less polydepletion effect (determined by the capacitance reduction in inversion region) was observed for germanium content of 20%; while for PMOS, resistance to polydepletion effect increase as germanium mole fraction increases.
3. For NMOS devices, a threshold voltage (V_t) is not affected by germanium content whereas for PMOS, the more germanium contents, the less boron penetration, which means less V_t shift.

In the work reported by Ponomarev *et. al.* (2000), it was mentioned that boron diffusion in poly-SiGe is slower, while arsenic and phosphorus diffusions are faster than in poly-Si. These differences are increasing with germanium fraction. At the same time, boron activation is higher in poly-SiGe while arsenic is more difficult to activate in poly-SiGe, especially when the germanium percentage rises above ≈ 50%. These facts suggest that poly-SiGe allows for good dopant activation in p-type gates, while for higher germanium mole fractions n-type gate activation becomes problematic. In order to reduce this problem at NMOS, Rhee *et. al.* (2001) has introduced a Ge-redistributed poly-Si/SiGe stacked gate (hereafter abbreviated as GRPSG). GRPSG can improve the current performance of PMOS without the degradation of NMOS. In this method, thin poly-SiGe (20% - 40% germanium) and poly-Si capping layers were in-situ deposited by LPCVD using GeH_4 and SiH_4. Then, N^+ poly implantation and annealing for germanium distribution were performed. N-type dopants such as phosphorus and arsenic can enhance the germanium diffusion into poly-Si more than boron. After the annealing for germanium redistributed, different germanium concentration was obtained at the gate/oxide interface between NMOS and PMOS. The current performance of the NMOS with GRPSG with low germanium content (< 5%) was not degraded.

Lin Chih-Yung *et. al.* (1995) reported the use of amorphous/polysilicon gate electrode in BF_2-implanted poly-gated P-MOSFET's to suppress the boron penetration. SIMS analysis clearly shows that fluorine prefers to accumulate in the layer of amorphous silicon. The retardation of boron diffusion is therefore achieved by trapping of fluorine in the amorphous layer of stacked amorphous/polysilicon (hereafter abbreviated as SAP) p-type gate due to a lower diffusion rate of fluorine in the amorphous silicon layer. Polysilicon depletion effect did not become severe by introducing the amorphous silicon. In addition, using the gate structures does not degrade gate oxide reliability. But Yu Bin *et. al.* (1998) noticed that the depletion in amorphous silicon is slightly more severe than polysilicon. The result indicated that amorphous silicon reduces the gate dopant diffusion but increases gate depletion after rapid thermal annealing (hereafter abbreviated as RTA) more than 1000°C.

Besides using alternate gate materials and changing gate microstructure, researchers also did a lot of studies in gate doping. Sun W.T. *et. al.* (1995) reported that phosphorus or arsenic co-implant in p-poly-gate can prevent boron penetration. In the paper, a comprehensive study of the phosphorus dosage and annealing condition dependencies of boron penetration and poly-depletion is presented. The experimental results show that the boron penetration in BF_2 implanted poly-gate is significantly reduced as the dose of co-implanted phosphorus increases. The phosphorus dose about 1.5×10^{15} cm^{-2} in p^+-poly gate can effectively retard the penetration of boron in BF_2 3 to 5×10^{15} cm^{-2} doped poly gate under 900°C, 60 min annealing (30 min reflow anneal, and 30 min post-contact implant anneal). The performance of boron-penetration-free phosphorus co-implanted p^+-poly gate MOSFETs is also shown to be much better than the device with boron-penetration. In arsenic co-implanted p-poly gate, it also appears that arsenic co-implant would also retard boron penetrating through thin gate oxide. However, boron penetration becomes very severe when annealed at higher temperature for longer time.

In the work reported by Kuroi *et. al.* (1995), nitrogen implantation into polysilicon gate was introduced. In these papers, authors claim that nitrogen implantation can reduce the problem of boron penetration. It is because the nitrogen implanted into the polysilicon gate is segregated into the gate oxide during heat treatment after implantation. The nitrogen in the gate oxide film can act as a barrier layer for boron penetration and reduce the random failures of gate oxide films under highly doped polysilicon gates. Yu Bin *et. al.* (1997) mentioned that a nitrogen dose of $5 \times 10^{15}\,\text{cm}^{-2}$ is the optimum choice at implant energy of 40 KeV in terms of the overall performance of both p- and n-MOSFET's. The suppression of boron penetration is confirmed by the SIMS profiles to be attributed to the retardation effect in polysilicon with the presence of nitrogen. SIMS profile showed that nitrogen peak and boron peak is at the same location. It is believed that nitrogen peak in the bulk poly-Si is mainly responsible for the suppression, since the boron diffusion is retarded in the bulk. High nitrogen dose ($1 \times 10^{16}\,\text{cm}^{-2}$) results in polydepletion and increase of sheet resistance in both unsilicided and silicided p^+ poly, degrading the transistor performance. Chao T.S. *et. al.* (1997) is using nitrogen co-implanted with boron method to suppress the boron penetration. In this method, it was found that nitrogen combines with the boron to form a B-N complex. This B-N complex can retard the penetration of boron itself and it was identified by X-ray photoelectron spectroscopy (hereafter abbreviated as XPS). The optimum dosage of nitrogen should be $1 \times 10^{15}\,\text{cm}^{-2}$. Excessive amount of the nitrogen concentration will increase sheet resistance, R_S and polydepletion effect. In the work reported by Liu C. T. *et. al.* (1996), nitrogen was implanted into silicon substrate prior to the growth of gate oxides. From the SIMS profile, the nitrogen peak is observed to be in the gate oxide and boron concentration falls off right before the nitrogen peak. No boron is seen beyond the oxide layer.

Apart from implanting nitrogen into polysilicon or silicon substrate, annealing in nitrogen ambient can also be used to establish a nitrogen diffusion barrier at the interface of the poly-Si and oxide. Yu Bin *et. al.* (1998), Han L.K. *et. al.* (1995) and Chao T.S. *et. al.* (1997) mentioned that thermally nitrided oxide or reoxidized nitrided oxide in NO, N_2O or NH_3 ambient can act as a diffusion barrier for boron. However, H atom residues during the NH_3 nitridation reduce the resistance to hot carrier stressing, introduces a large number of interface traps and fixed oxides charge. This contributes to 1 by Flickers noise. An additional reoxidation process needs to be done to reduce the H atom. Hence, the NH_3 nitridation process is too complex to implement in a real production line. Therefore, a nitridation process on the oxide surface without H incorporation has started to be used. But between N_2O and NO, NO is the preferred ambient. It is because N_2O based oxides require a much higher thermal budget for sufficient nitrogen concentration. Han L. K. *et. al.* (1995) reported the use of rapid thermal NO-nitrided SiO_2 as an alternate oxynitride with enhanced boron diffusion barrier properties while maintaining a low thermal budget and H-free nature of the processing ambient. In this method, thermal oxides are grown at 1050°C in pure oxygen followed by in-situ nitridation at 1000°C in pure NO ambient, finally received an in-situ nitrogen anneal at 1050°C. Results show that NO-nitrided SiO_2, as compared to conventional thermal SiO_2, exhibits much higher resistance to boron penetration and also reduce polydepletion effect. It is believed to be due to the incorporation of nitrogen in the dielectric during NO nitridation. In the work of Xiang Qi *et. al.* (2000), ultra-thin nitride/oxynitride stack gate dielectric was formed by the thermally growing of oxynitride followed by Si_3N_4 deposition. Pre-doped gates for both N- and P-MOS devices (P-doped for NMOS and N-doped for PMOS) are employed to minimize poly depletion and maximize drive currents. The result shows that the stack gate dielectric has strong boron penetration resistance, which allows pre-doped for both type devices to minimize poly depletion and improve device performance.

Boron diffusion and activation is normally controlled by the annealing process after the implantation. To prevent boron penetrating through thin gate oxide, one can optimize the annealing parameters to get the optimize boron activation and diffusion in polysilicon gate. Fiory *et. al.* (2000) reported that when choosing ion implants and thermal budget for the polysilicon gate engineering in deep submicron devices, the following conditions must be met:

(i) The activation of implanted dopants (usually boron for PMOS and arsenic or phosphorus for NMOS devices) should be high across the gate in order to minimize the poly-depletion effect and to maximize the device speed;

(ii) Boron penetration through the thin gate oxides (which causes threshold voltage shifts in PMOS devices) should be minimal; and

(iii) Lateral diffusion (which cause cross-doping in connected gates) should be minimal.

These goals are difficult to achieve, since electrical activation of dopants in poly-Si entails rapid grain boundary diffusion and slower diffusion within grains. Fiory *et. al.* (2000) used spike annealing to avoid boron penetration through gate oxide. It is because of sufficient rapid grain boundary assisted diffusion of boron in polycrystalline silicon films where high carrier concentration could be achieved by utilizing high temperature annealing, since the concentration of boron in solid solution increases with temperature. It is concluded that spike annealing can increase effective diffusivity of boron in polysilicon of up to 200 times greater than in crystal silicon. The activation of boron also increases based on the lower sheet resistance of the polysilicon after spike annealing. In the work of Yu Bin *et. al.* (1999), laser thermal process (hereafter abbreviated as LTP) was introduced. LTP has 3 attractive features:

1. "Zero" thermal budget. Laser pulses last a few nanoseconds, which is about 8 orders of magnitude shorter than rapid thermal processing. Thermal diffusion of dopants is almost negligible.
2. Metastable process without dopant solid solubility limit. This allows active dopant concentrations larger than 10^{21} cm^{-3} to be achieved,
3. Local selective heating of specific regions of silicon wafer. No thermal budget is added to V_t/retrograde channel/halo profiles.

Although the results show the sheet resistance is lower, activation of the dopants occurs during LTP. But the gate-depletion effect is not significantly changed by LTP. Unfortunately, LTP could not be made into production process yet. The main problem faced is the non-uniform heating.

From the literature review work, the polydepletion effect and boron penetration effect are well understood. Based on the literature and equipment availability in Silterra, the spike annealing and NO gate oxide have been chosen as the main focus in this work.

CHAPTER 3

THEORY OF PROCESSES AND CHARACTERIZATION TECHNIQUES

3.1 Processes

In this experimental work, several processes were adopted. These processes are oxidation, chemical vapour deposition, ion implantation and annealing. The theory of the processes is discussed in the following subsections.

3.1.1 Oxidation

Oxidation of silicon is a thermal process to grow an oxide layer. This process is easily achieved by heating the silicon substrate to temperatures typically in the range of 900 - 1200°C. Oxidation can divided into wet or dry oxidation, depending on the atmosphere in the furnace. The atmosphere in the furnace can either contain pure oxygen or water vapour. Both of these molecules diffuse easily through the growing SiO_2 layer at these high temperatures. Oxygen arriving at the silicon surface will react with silicon to form silicon dioxide. The chemical reactions that take place are either

$$Si + O_2 \rightarrow SiO_2$$

for so-called dry oxidation or

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2$$

for wet oxidation.

Due to the stoichiometric relationships in these reactions and the difference between the densities of Si and SiO_2, about 46% of the silicon surface is "consumed" during oxidation. That is, for every 1 μm of SiO_2 grown, about 0.46 μm of silicon is consumed (see Figure 3.1).

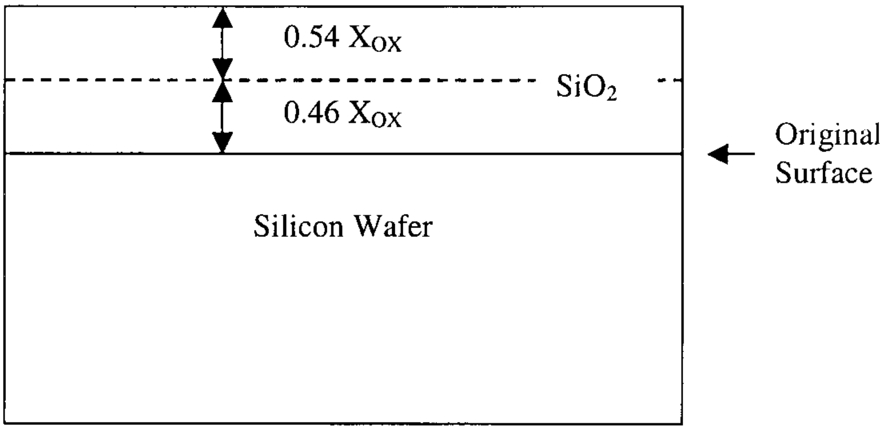


Figure 3.1. Oxidation process on silicon.

3.1.1.1 Oxide Growth Kinetics

The growth of silicon dioxide only involved surface reaction. However, after the layer of SiO_2 is getting thicker, the arriving oxygen molecules must diffuse through the growing SiO_2 layer to react with the silicon surface.

"Deal/Grove" model (Puchner, 1996) is used to explain the oxide growth kinetics. This model is generally valid for temperatures between 700 and 1300°C, partial pressures between 0.2 and 1.0 atmosphere, and oxide thickness between 0.03 and 2 μm for both wet and dry oxidation. To understand this model, please refer to Figure 3.2, and let:

C_g = concentration of oxidant molecules in the bulk gas
C_s = concentration of oxidant molecules immediately adjacent to the oxide surface
C_o = equilibrium concentration of oxidant molecules at the oxide surface
C_i = concentration of oxidant molecules at the Si/SiO_2 interface

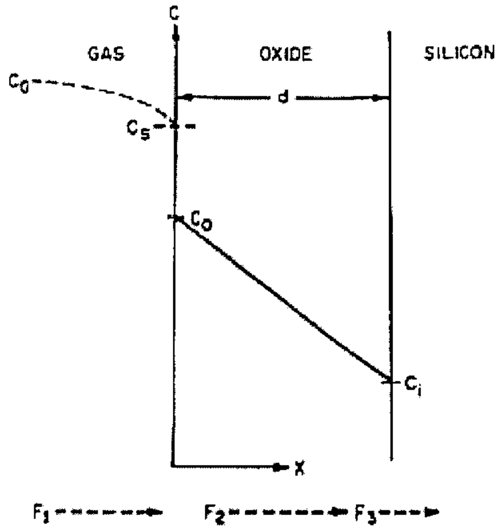


Figure 3.2. Deal/grove model for oxidation.

Note that: 1) $C_g > C_s$ (due to depletion of the oxidant at the surface)

2) $C_s > C_o$ (due to the solubility limits of SiO_2)

The oxidizing species (oxygen or water vapour) are flowed from the bulk gas to the gas/oxide interface with flux F_1 (where flux is the number of molecules crossing a unit area per unit time). The species are flowed diffuse through the growing oxide toward the silicon surface with flux F_2, and react at the Si/SiO_2 interface with flux F_3. Mathematically:

F_1 = flux of oxidant from gas to surface

$$F_1 = H_g(C_g - C_s) \qquad \text{(Eq 3.1)}$$

where H_g = the gas phase mass transfer coefficient.

F_2 = flux through the oxide layer

$$F_2 = -D\frac{dC}{dx} \qquad \text{(Eq 3.2)}$$

where D is the diffusivity of the oxidant molecule in SiO_2 and C is the concentration.

Assumed a linear concentration gradient inside the oxide layer, then:

$$F_2 = D\frac{(C_o - C_i)}{y} \qquad \text{(Eq 3.3)}$$

where y is the current value of the oxide thickness.

Finally:

$$F_3 = K_s \times C_i \qquad \text{(Eq 3.4)}$$

where K_s = the rate constant for the surface chemical reaction.

Now let's look at F_1 more closely. By Henry's Law:

$$C_o = H \times P_s \quad \text{and}$$

$$C^* = H \times P_g \qquad \text{(Eq 3.5)}$$

where H is Henry's Law constant, C^* is the equilibrium concentration of oxidant molecules in the bulk SiO_2, and P_s and P_g are the partial pressures of the oxidant molecules adjacent to the SiO_2 surface and in the bulk gas, respectively. From the Ideal Gas Law:

$$C_g = \frac{P_g}{kT} \quad \text{and}$$

$$C_s = \frac{P_s}{kT} \qquad \text{(Eq 3.6)}$$

where k is Boltzman's constant and T is the temperature in degrees K. Therefore, F_1 can be re-written as:

$$\begin{aligned} F_1 &= H_g(C_g - C_s) = H_g\left(\frac{P_g}{kT} - \frac{P_s}{kT}\right) \\ F_1 &= \frac{H_g}{kT}(P_g - P_s) = \frac{H_g}{kT}\left(\frac{C^*}{H} - \frac{C_o}{H}\right) \qquad \text{(Eq 3.7)} \\ F_1 &= h(C^* - C_o) \end{aligned}$$

where $h = H_g/HkT$. At steady state, all three fluxes should be equal. In other words,

$F_1 = F_2 = F_3$ under these conditions:

$$C_i = \frac{C^*}{(1+\frac{K_s}{h}+\frac{K_s * y}{D})} \quad \text{and}$$

$$C_o = \frac{(1+\frac{K_s * y}{D})C^*}{(1+\frac{K_s}{h}+\frac{K_s * y}{D})} \qquad \text{(Eq 3.8)}$$

If the oxidation growth rate depends only on the supply of oxidant to the Si/SiO$_2$ interface, it is said to be diffusion controlled. Under this condition, D is close to zero. Therefore:

$$C_i \approx 0 \text{ and } C_o \approx C^*$$

If, on the other hand, there is plenty of oxidant at the interface, the growth rate depends only on the reaction rate. This situation is called "reaction-controlled." In this case, D approaches infinity, and:

$$C_i = C_o = \frac{C^*}{(1+\frac{K_s}{h})} \qquad \text{(Eq 3.9)}$$

To compute the growth rate itself, let N_1 be the number of oxidant molecules per cubic cm incorporated into the oxide layer. Then, the differential equation can be written as below:

$$\frac{d}{dt}N_1(y) = F_3 = K_s \times C_i = \frac{K_s C^*}{(1+\frac{K_s}{h}+\frac{K_s * y}{D})} \qquad \text{(Eq 3.10)}$$

Under the boundary condition that $y = 0$ when $t = 0$, first order differential equation can be solved as:

$$y^2 + Ay = B(t+T) \qquad \text{(Eq 3.11)}$$

Where

$$A = 2d(\frac{1}{K_s} + \frac{1}{h})$$

$$B = \frac{2DC^*}{N_1}$$

$$T = \frac{(yi^2 + Ayi)}{B}$$

and yi is the initial oxide thickness. For short times, the growth rate is reaction limited, and the oxide thickness is approximately:

$$y = \frac{B}{A}(t + T) \qquad \text{(Eq 3.12)}$$

For longer times, growth is diffusion-limited, and the approximation used is:

$$y = \sqrt{Bt} \qquad \text{(Eq 3.13)}$$

A and B are constants which correspond to the oxidation conditions (i.e. - temperature and wet or dry) (Puchner, 1996).

3.1.1.2 Oxidation Reactor

Figure 3.3 shows a simple oxidation reactor system. Here the wafers were held in a vertical holder with pure water being pumped through. The water vapour was made from the reaction of hydrogen and oxygen gases. The temperature of the furnace was increased to 750°C to let the reaction of the water vapour on the wafer surface to form surface oxygen and hydrogen gas. Nitrogen gas will purge into the oxidation reactor after the formation of SiO_2. The wafer will be annealed to let the nitrogen to react with oxygen layer, forming an oxynitride layer.

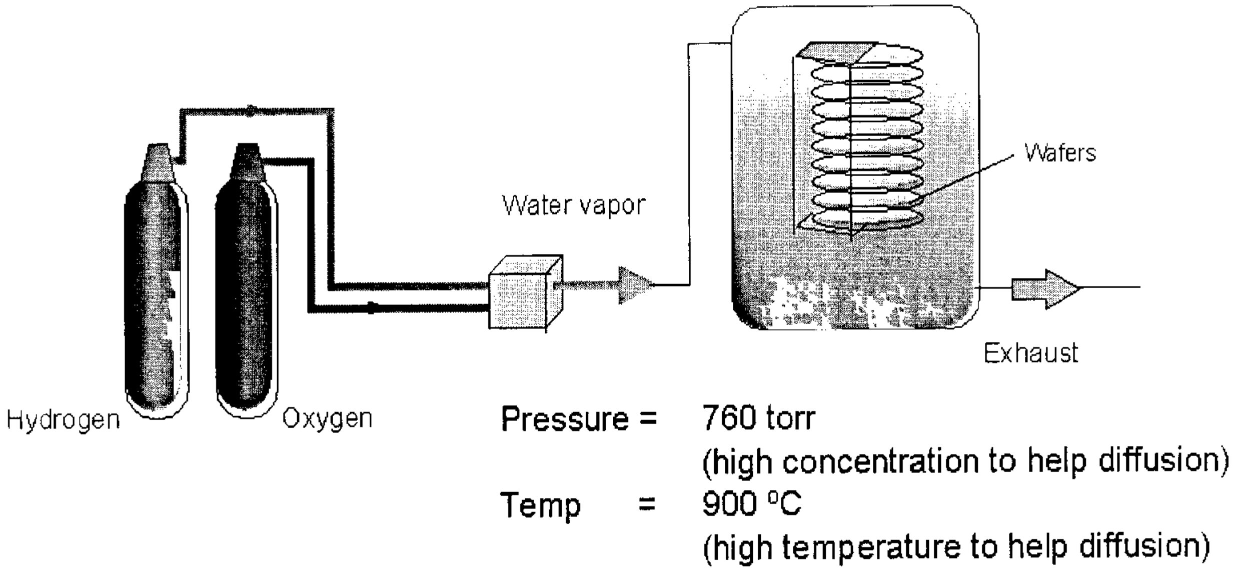


Figure 3.3. Simple diagram of oxidation reactor.

3.1.2 Chemical Vapour Deposition (CVD)

CVD is an extremely popular way to deposit thin film layers in integrated circuit manufacturing. The process is controlled by two factors namely the diffusion of gases to wafer surfaces and the reaction of the gases on the wafer surfaces. Typically the CVD process is designed in such a way that deposition is limited by only one of the two factors, and is independent of the other. The mathematics of gaseous diffusion and chemical reactivity in furnaces are very difficult and not fully understood. Therefore it is critical for integrated circuit manufacturers to develop standard, repeatable recipes (gas flows, temperatures, deposition time, etc.) and avoid deviating from them.

For better uniformity, LPCVD is used. Deposition rate in an LPCVD process is limited by gaseous diffusion. For uniformity, it is critical to have uniform gas flow to all wafers in an LPCVD furnace. LPCVD typically operates at a much higher temperature (up to 1000°C) than atmospheric pressure CVD (hereafter abbreviated as

APCVD). Polysilicon and most dielectrics are deposited by LPCVD. Typical LPCVD process pressures are from 0.1 to 1.0 torr.

3.1.2.1 Polysilicon Deposition Mechanism

Polysilicon is deposited by pyrolyzing silane between 575°C and 650°C in a low-pressure reaction:

$$SiH_4(g) \rightarrow Si(s) + 2H_2(g)$$

Either pure silane or 20 to 30% silane in nitrogen is bled into the LPCVD system at pressure of 0.2 to 1.0 torr. For practical use, a deposition rate of about 10 to 20 nm/min is required. The properties of the LPCVD polysilicon films are determined by the deposition pressure, silane concentration, deposition temperature and dopant content. The deposition rate is defined as below

$$\text{Deposition Rate} = \frac{\text{Thickness of Film}}{\text{Time to grow}}\left(\frac{\text{Å}}{\text{Min}}\right)$$

3.1.2.2 LPCVD Furnace

An LPCVD furnace is shown in Figure 3.4. A typical LPCVD system composes of a chamber and vacuum pump, source gas inlet controls, heating and temperature controls. The source gas is depleted (consumed) as it travels the length of the tube, often the temperature is increased toward the rear of the tube to help compensate for the decreased gas concentration.

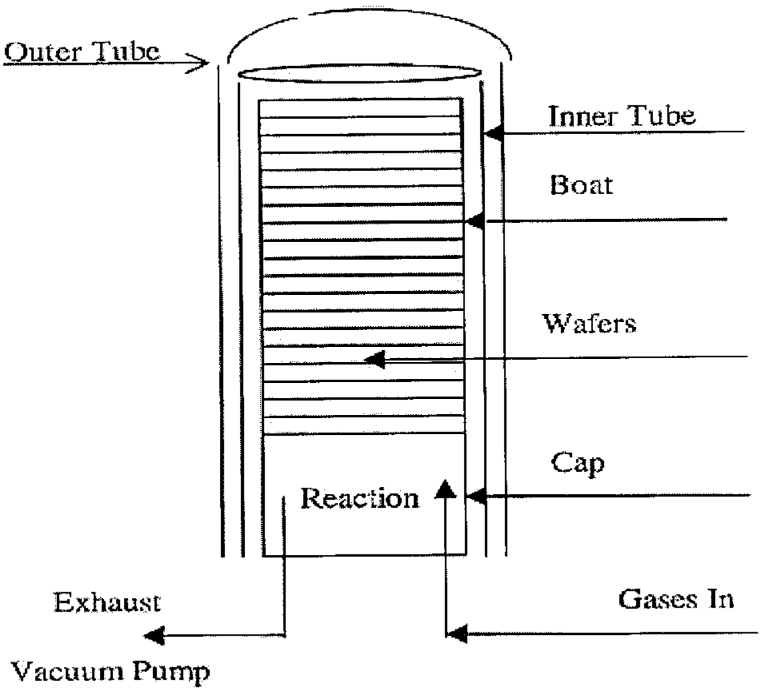


Figure 3.4. LPCVD Furnace for polysilicon deposition.

3.1.3 Ion Implantation

Ion implantation is the introduction of ionized dopant atoms into a substrate with enough energy to penetrate beyond the surface. The most common application is substrate doping. The common dopants that are used in the ion implantation are shown in Table 3.1. The use of 3 to 500 KeV implant energy for dopant ions is sufficient to implant the ions from 100 to 10,000 Å below the silicon surface. These depths place the atoms beyond any surface layers of 30 Å native SiO_2, and therefore any barrier effect of the surface oxides during impurity introduction is avoided. The depth of implantation, which is proportional to the ion energy, can be selected to meet a particular application.

The major advantage of ion implantation technology is the capability of precisely controlling the number of implanted dopant atoms. Upon annealing the target (heating to elevated temperatures of approximately 600-1000°C), precise dopant concentrations between 10 x 10^{14} to 10 x 10^{21} atoms/cm^3 in the silicon are obtained. Furthermore, the dopant's depth distribution profile can be well controlled. Some common dopants for silicon are listed in the Table 3.1 below.

Table 3.1: List of dopants commonly used in semiconductor doping.

Donors (Electron-increasing Dopants) Group V elements	Acceptors (Hole-increasing Dopants) Group III elements
P As Sb	**B** Ga In Al

Silicon, being an element from Group IV of the Periodic Table of the Elements (Appendix A), has four valence electrons. In an undoped silicon crystal, each atom shares these electrons with its four nearest neighbours via covalent bonding (see Figure 3.5). Dopant atoms can be either as "donors" or "acceptors." Donors normally elements from Group V increase the electron concentration in the silicon, whereas acceptors normally elements from Group III increase the hole concentration.

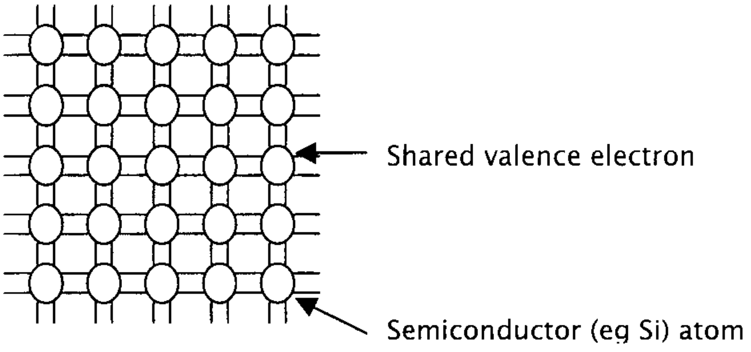


Figure 3.5. Atoms bonding diagram for an undoped silicon crystal.

When Group V elements such as arsenic or phosphorus are introduced into the silicon crystal lattice, four of the five valence electrons fit into the silicon bonding structure.

The fifth electron, however, is weakly bound, and at room temperature is free to move through the crystal (Figure 3.6). It is thus said to be "donated" to the semiconductor.

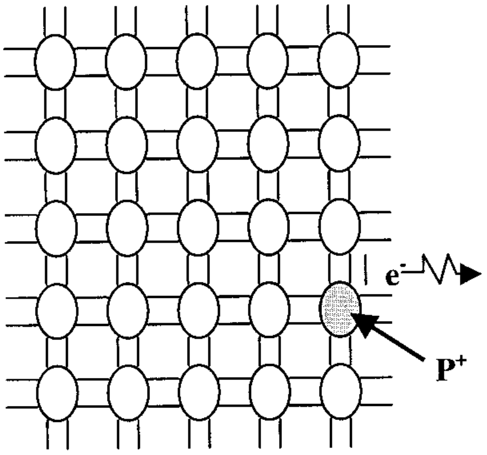


Figure 3.6. N-type dopant in the silicon crystal causes an excessive electron in the lattices.

Likewise, Group III elements such as boron have only three valence electrons, and therefore cannot complete all four of the available bonds when substituted for a silicon atom. These elements, however, readily "accept" an electron from a nearby silicon-silicon bond. This creates a hole, which is also free to wander through the crystal lattice at room temperature (Figure 3.7).

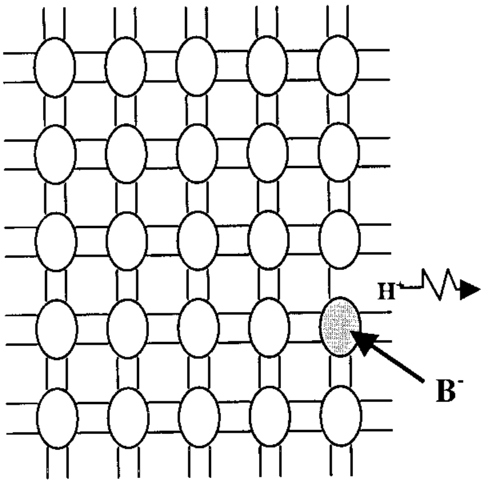


Figure 3.7. P-type dopant in the silicon crystal cause a hole in the lattice.

3.1.3.1 Ion Implanter

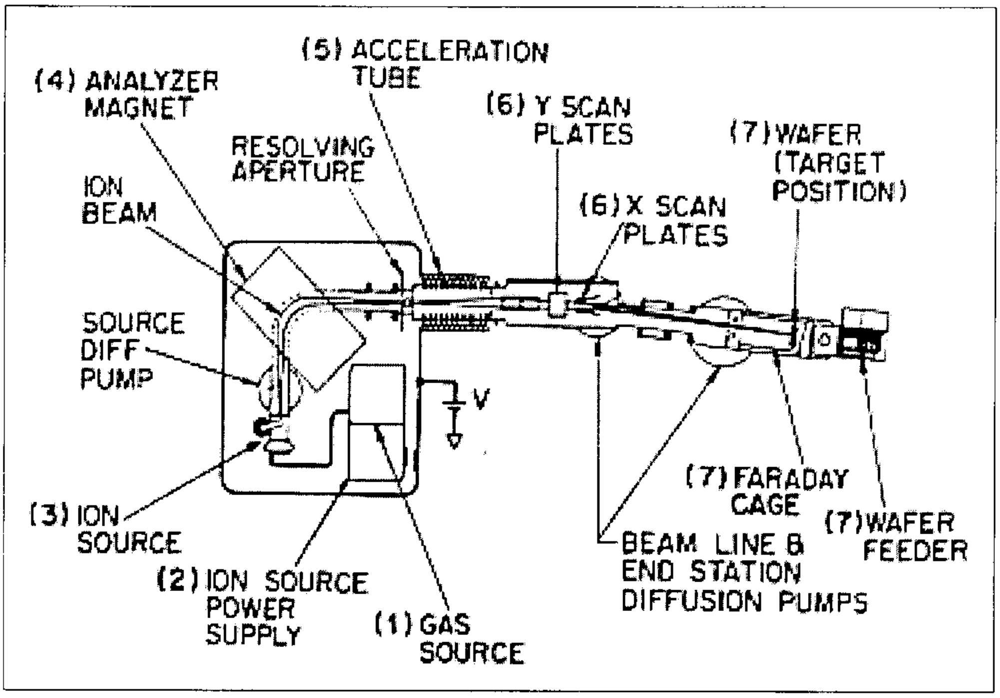


Figure 3.8. Basic diagram of ion implanter. (Reproduced from Silterra, 2000)

Figure 3.8 shows a simplified diagram indicating the major elements of a medium-current ion implanter system. In the ion implanter system, the ion source will start with selected molecular species from the gas source and converts it into ion. The ions are accelerated, and then enter the analyser magnet for ion selection. The exit beam of desired implant ions are chosen based on the charge-to-mass ratio of the ions, and the analyser is generally sensitive enough to discriminate against adjacent mass numbers. The ions are then given a final acceleration at the acceleration tube, after which the ion beam will be slightly electrostatically deflected to separate it from any neutral atoms that may have formed.

The beam is then scanned over the wafer surface, either electrostatically, mechanically, or by a combination of the two controlled X- and Y-scan plates. In addition, an electron source may be near the wafer to "flood" the surface with electrons and prevent a charge build-up on insulating surfaces such as SiO_2 or silicon nitride. If not eliminated, this kind of charge build-up can be severe enough to cause gate oxide failure because of electrical breakdown from the gate to the substrate through the oxide.

3.1.4 Annealing and Diffusion

Basically, annealing in the CMOS processes is for the purpose of activation and dopant diffusion. After a dopant is implanted in silicon an annealing step has to be performed to activate the introduced dopant atoms and to repair the damage caused by the ion implantation. In modern process, the technology of RTA processes is the method of choice with typical annealing times of several seconds. Diffusion can be defined as the random walk of an ensemble of particles from regions of high concentration to regions of lower concentration. In integrated circuit fabrication, diffusion is used to introduce dopants in controlled amounts into the semiconductor substrate.

According to the First Law of Diffusion (Puchner, 1996), the transfer of solute atoms per unit area in a one-dimensional flow can be described by the following equation:

$$J = -D\frac{\partial C(x,t)}{\partial x} \qquad \text{(Eq 3.14)}$$

where J is the particle flux, C is the concentration of the solute, D is the diffusion coefficient, x is the distance into the substrate, and t is the diffusion time. The negative sign indicates that the diffusing mass flows in the direction of decreasing concentration.

From the Conservation of Mass, we also know that:

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x} \qquad \text{(Eq 3.15)}$$

If we combine this relationship with the 1st Law of Diffusion, then we have derived the 2nd Law of Diffusion (otherwise known as Fick's Law), which states:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \qquad \text{(Eq 3.16)}$$

In order to solve Fick's Law, one initial condition and two boundary conditions are required. Two solutions to Fick's Law are generally encountered in Integrated Circuit (hereafter abbreviated as IC) fabrication: infinite-source and limited-source diffusion. Both solutions are described below.

3.1.4.1 Infinite-source Diffusion (Pre-deposition)

Infinite-source diffusion requires a constant surface concentration of diffusing atoms. This generally corresponds to the process step known as "pre-deposition." In this case, the initial condition and boundary conditions are:

$$\begin{aligned} C(x,0) &= 0 \\ C(0,t) &= C_s \\ C(\infty,t) &= 0 \end{aligned} \qquad \text{(Eq 3.17)}$$

where C_s is the surface concentration. The solution to Fick's Law (Puchner, 1996) under these conditions is:

$$C(x,t) = C_s \quad erfc\left[\frac{x}{2\sqrt{Dt}}\right] \qquad \text{(Eq 3.18)}$$

where "erfc" is the complementary error function.

3.1.4.2 Limited-source Diffusion (Drive-in)

Limited-source diffusion requires a constant amount of total dopant per unit area of the diffusing surface. This corresponds to the process step known as "drive-in" or any subsequent heat cycles. In this case, the initial condition and boundary conditions are:

$$C(x,0) = 0$$
$$\int C(x,t)dx = S \qquad \text{(Eq 3.19)}$$
$$C(x,\infty) = 0$$

where S is called the "dose." The solution to Fick's Law under these conditions is:

$$C(x,t) = \frac{S}{\sqrt{\pi Dt}}\exp\left[\frac{-x^2}{4Dt}\right] \qquad \text{(Eq 3.20)}$$

3.1.4.3 Diffusion Mechanism

From the past literatures (Sze, 1988, Fahey *et. al.*, 1989, Taylor *et. al.*, 1993), three possible atomic diffusion mechanisms have been established. The first one is the direct mechanism. In this mechanism, the impurities, which have small ionic radii, can travel directly interstitial. (see Figure 3.9).

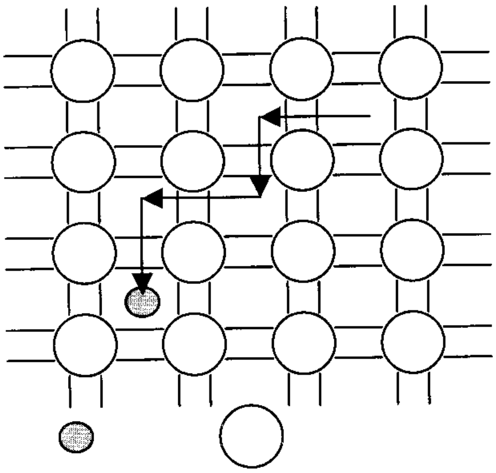


Figure 3.9. Direct diffusion mechanism within crystal targets. The diffuser can travel exclusively on interstitial sites.

The second possible mechanism is vacancy mechanism. In this mechanism, the substitutional dopants migrate through the host lattice by moving via adjacent vacant lattice sites. (see Figure 3.10*a*). To avoid oscillations of this exchange procedure, the vacancy must move at least to a third-neighbour site away from the dopant. Thus, the vacancy or another vacancy within the diffusion regime can return on a different path within the silicon's diamond lattice. Then long-range migration of the dopant will take place (see Figure 3.10*b*). Energetic details of this complicated exchange can be found in Fahey *et. al*, 1989.

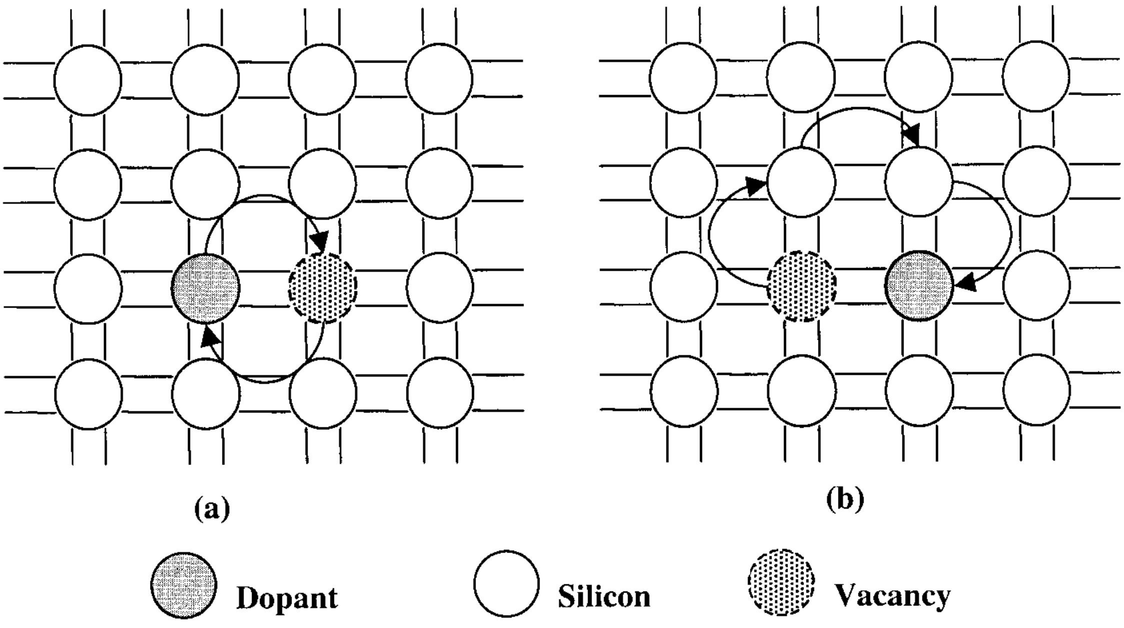

(a) **(b)**

Figure 3.10. Dopant diffusion by vacancy mechanism.

The third one is the interstitial mechanism (see Figure 3.11). In this mechanism, dopant diffusion takes place when a dopant at a substitutional site is approached by a silicon interstitial (Figure 3.11a). The dopant is kicked out by the interstitial to reside at interstitial position (Figure 3.11b), while the original self-interstitial has disappeared by occupying the regular lattice site. Now this interstitial dopant is able to move towards an adjacent lattice site to re-form a silicon interstitial by the same "kick-out" process, like the interstitial dopant itself was generated (Figure 3.11c).

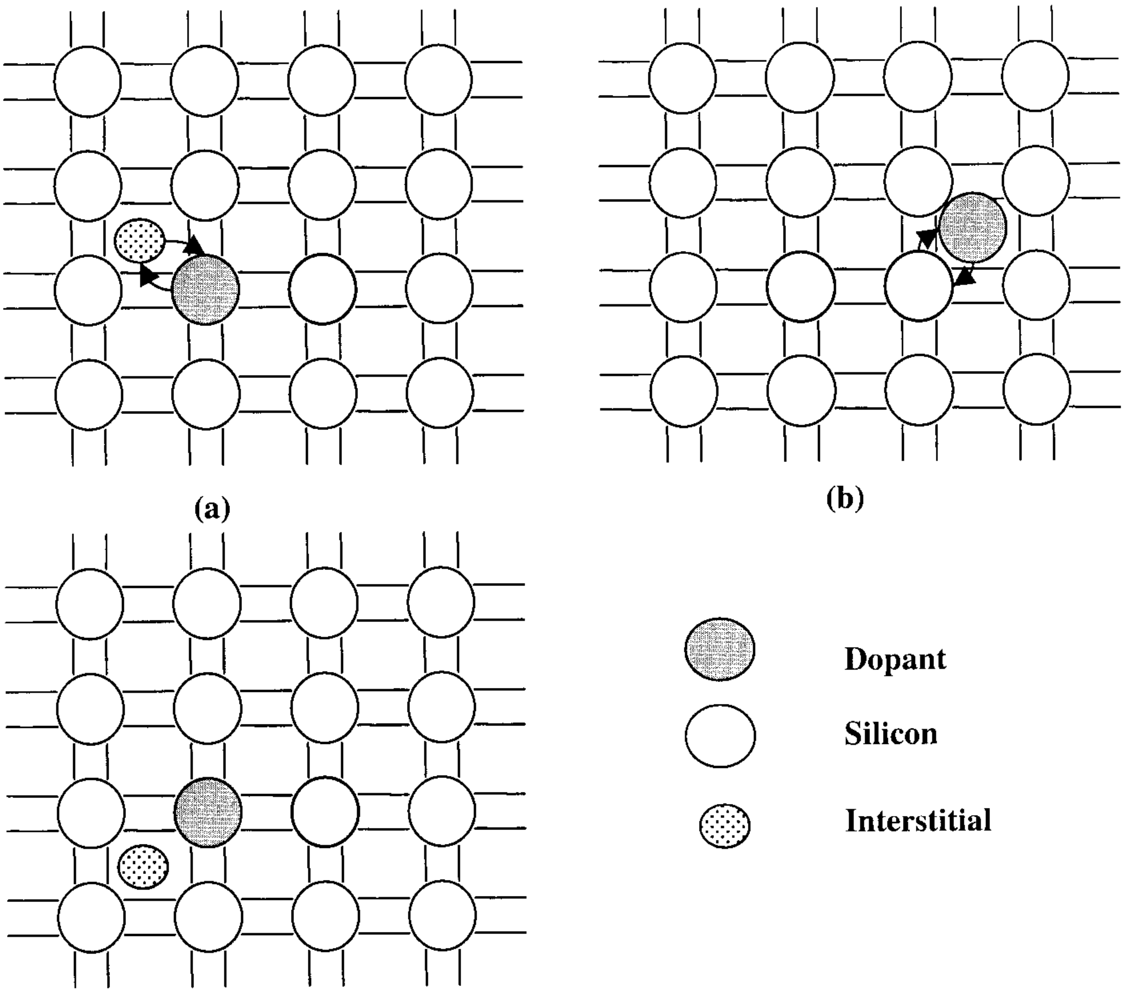


Figure 3.11. Dopant diffusion via a substitutional-interstitialcy interchange.

3.1.4.4 Diffusion in Polysilicon

Polysilicon shows extraordinary high diffusivity for dopants. This fact allows the dopants to diffuse even longer distances within the polysilicon material in reasonable short time periods.

Dopant transport within polysilicon involves four major mechanisms:

- fast diffusion in grain boundaries

- segregation between grain interior and grain boundaries
- grain boundary motion
- grain interior diffusion

A crystal structure of a grain bulk/grain boundary network can be used to understand the dopant/grain boundary system and the according diffusion mechanisms. From the tetrahedral bonding network for two polysilicon grains separated by a grain boundary (see Figure 3.12), it can be seen that the number of bonds crossing from one grain into the neighbouring grain is reduced at the interface. This can cause a corresponding high density of dangling bonds due to the fracture of the crystal along the interface. Therefore, this boundary area may be a preferred low energy local minimum for dopants. This outstanding energetic properties in combination with the irregularity of the tetrahedral bonding at the interface makes the grain boundary to be the fast diffusion path for dopants.

The second transport mechanism is segregation between grain interior and grain boundaries. This mechanism happened if the dopant and the grain boundary are a closed system and there is enough space in the grain boundary that had occupied by other dopant. The segregation of dopants into the grain boundary can be described by trapping and emission mechanisms. The capture and emission rates depend on the number of occupied and free states in the grain boundary, where the total number of states is limited by the grain boundary area.

The third transport mechanism is related to the grain growth phenomenon. In the grain growth process, the grain boundaries are moving. During moving of the grain boundaries, the dopants are incorporated in the grain boundaries. This movement results in a net dopant transport.

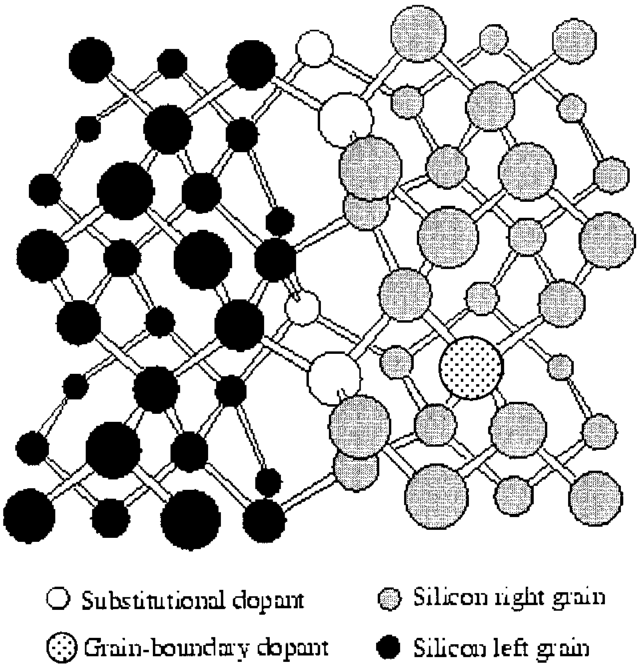


Figure 3.12. Three-dimensional perspective drawing of a grain/grain boundary network. (Reproduced from Puchner, 1996)

If the crystal structure of the polysilicon grain bulk region shows a regular silicon lattice, the diffusion of dopants within the grain interior regions will be treated like normal diffusion in silicon. Moreover, effects like dopant activation and clustering play an important role in the grain bulk and have to be considered for a complete description of the polysilicon diffusion problem.

3.1.4.5 Annealing Profile

Annealing profiles in this experiment are both spike annealing and normal RTA. Figure 3.13 shows an example of the spike annealing profile. The process started with slow ramp at 25°C/s from 450°C to 550°C. The temperature is then stabilized for 2 seconds before the high ramp up. There is zero dwell time for spike anneal. So, after reaching the peak temperature, the wafer is treated with high ramp down rate at 75°C. Spike annealing has high peak temperature that can maximize the dopant activation, aggressive ramp-up that can minimize the dopant diffusion and aggressive ramp-down that can minimize the dopant deactivation.

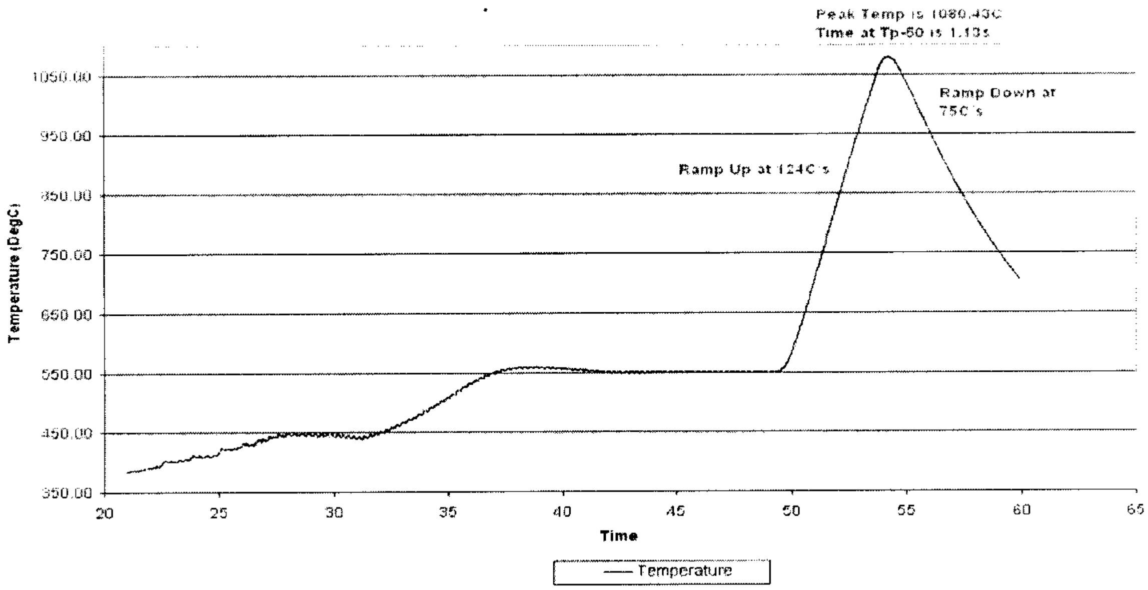


Figure 3.13. Spike annealing profile.

3.2 Characterization Techniques

For characterization, a four-point probe was used to measure the sheet resistance in the polysilicon, Secondary Ion Mass Spectrometry (SIMS) to get the dopant depth profile in order to measure the distribution of the dopants in the polysilicon and Focused Ion Beam (FIB) to analyse the surface structurally.

3.2.1 Four-Point Probe

The four-point probe is a simple, non-destructive way to determine resistivity of a diffused layer in a silicon substrate, or a deposited conducting film (e.g. Al). The four-point probe contains four thin collinearly placed tungsten wires probes, which are made to contact the sample under test, shown schematically in Figure 3.14. Current, *I* is made to flow between the outer probes, and voltage, *V* is measured between the two inner probes, ideally without drawing any current (Puchner, 1996).

The sheet resistance value, R_S is calculated by equation below:

$R_S = 4.53\left(\frac{V}{I}\right)$ in ohms/square or (Ω/□).

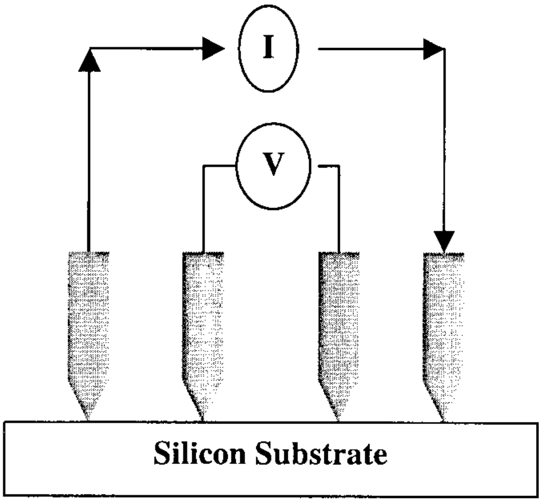


Figure 3.14. Four-point probe schematic diagram.

3.2.2 Secondary Ion Mass Spectrometry

The technique of Secondary Ion Mass Spectrometry (SIMS) is the most sensitive of all the commonly employed surface analytical techniques. This is because of the inherent sensitivity associated with mass spectrometric-based techniques.

There are 3 different variants of the technique, namely statics SIMS, dynamic SIMS and imaging SIMS. Static SIMS is used for sub-monolayer elemental analysis. Dynamic SIMS is used for obtaining compositional information as a function of depth below the surface. Imaging SIMS is used for spatially resolved elemental analysis.

All of these variations on the technique are based on the same basic physical process and it is this process, which is discussed here, together with a brief introduction to the field of static SIMS.

In SIMS (Figure 3.15) the surface of the sample is subjected to bombardment by high-energy ions leading to the ejection (or sputtering) of both neutral and charged (+/-) species from the surface. The ejected species may include atoms, clusters of atoms and molecular fragments.

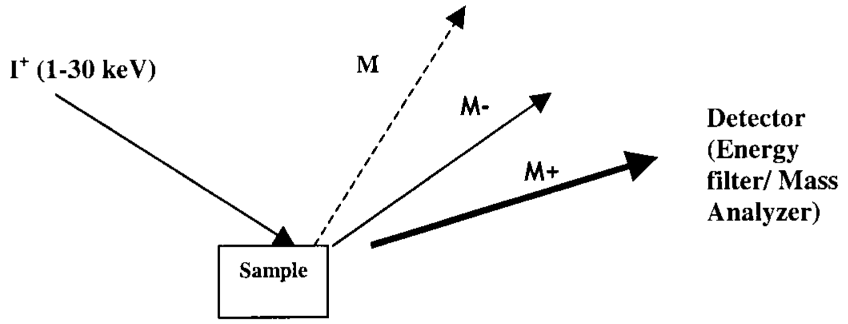


Figure 3.15. Extraction of dopant profile from the bombardment of ions.

In conventional SIMS, it is only the positive ions that are mass analysed. This is primarily for practical ease but it does lead to problems with quantifying the compositional data since the positive ions are but a small, non-representative fraction of the total sputtered species. It should be further noted that the displaced ions have to be energy filtered before they are mass analysed (i.e. only ions with kinetic energies within a limited range are mass analysed).

The most commonly employed incident ions (denoted by I^+ in the above diagram) used for bombarding the sample are argon ions (Ar^+) but other ions (e.g. alkali metal ions, Ga^+) are preferred for some applications.

The mass analyser is typically a quadruple analyser with unit mass resolution, but high specification time-of-flight (hereafter abbreviated as TOF) analysers are also used. The TOF analyser provides substantially higher sensitivity and a much greater mass range (although at a higher cost).

3.2.3 Focused Ion Beam

A focused ion beam (FIB) system is similar to a scanning electron microscope with an exception of using ion beam instead of electron beam to scan across the sample. The ion beam is ejected from a liquid metal ion source (usually Ga), with a spot size on modern systems of < 10 nm. Imaging using secondary electrons provides surface information with similar resolution to that obtainable from an SEM; however the main applications arise from the use of ions as the scanned species. These include compositional imaging via secondary ions, direct etching of material in selected regions for in-situ sectioning and imaging, microfabrication, transmission electron microscopy specimen preparation, and localized deposition and implantation of metal and insulator structures. Main application is in nanofabrication of devices. The unique combination of 10 nm resolution imaging with the ability both to remove and to deposit material in selected areas provides a means of performing materials studies or device fabrication processes which would otherwise be impossible or unreasonably time-consuming.

CHAPTER 4

METHODOLOGY PROCEDURES

This chapter describes the methodology procedures of this research. It is divided into two main sections, dealing with simulation and experimental aspect of the research work.

4.1 Simulation

This section discussed the simulation works that have done in this research. Here, process simulation tools, TSUPREM4 and MEDICI were used to get an optimised process parameter for 0.13-μm CMOS transistor.

4.1.1 Process Simulation

Before any simulation on any technology, a list of the specification of the required parameters needs to be prepared. These parameters include the junction depth, poly-gate thickness, channel length, threshold voltage, saturation drain current, well sheet resistance, poly gate sheet resistance, gate oxide thickness and junction capacitance. For a normal practice, the 0.13-μm CMOS process recipe is modified from the current 0.18-μm CMOS process recipe.

In order to get an accurate result, a simulation model has to have a finer grid. Simulation is based on the mathematical calculation where in order to get an accurate result; a finer meshing model has to be created. This means that the simulation model will have more calculation points to plot the results. Once the mesh had been created, the models used in the simulation will have to be defined.

In the poly depletion study, the simulation works will only concentrates on the process flow in between gate oxide growth and source drain anneal. In between this, the process flow consists of poly silicon gate deposition; LDD implant, pocket implant and source drain implant. In order to get an optimize implant recipe for the 0.13-μm, the full simulation work has to adopt a 0.13-μm CMOS complete process flow. This work included tuning the implantation condition, optimizing the annealing parameters and changing the deposition thickness. The simulation by the name of MEDICI was then used to simulate the electrical parameters such as, threshold voltage, saturation drain current, sheet resistance values and the junction capacitance in order to meet with the desired value. For the junction depth, Taurus visual was used to plot the cross sectional view of the device. Taurus visual can be also used to plot the electrical flow and electrical field, simulated by MEDICI.

Having obtained the desired process parameters, the simulation will now concentrate on the investigation of poly depletion study. First and foremost, a finer grid of the devices has to be created. Because the investigation is in the poly silicon gate, the finer grid has to be created at the interface of the poly silicon gate and the gate oxide. This is done during the deposition step of the poly silicon. Besides defining a finer grid in the poly silicon region, the diffusion model for poly silicon is also defined. This is because different diffusion model for different materials like single crystal silicon, poly crystal silicon and metal have different value that represent different diffusion properties. The process is then continued by applying the LDD implantation, pocket implantation and source drain implantation. The implantation recipes for these 4 different implantations are shown in Table 4.1. In the implantation, Monte Carlo model was used for the defect calculation in order to get a better accuracy. After that, the source drain anneal (split into many different heat cycle) is applied. followed by simulation with MEDICI to get the poly depletion thickness and the sheet resistance value. This is done by plotting the cross sectional view of the devices using Taurus Visual. The poly depletion thickness and the dopant distribution can be seen in the 2-Dimensional view of the device cross section. To analyze the dopant profile, a one-dimensional plot of concentration versus depth of the device can be shown indicating a dopant cut line at the desired region . From this plot, the dopant concentration can be derived.

Table 4.1. Implantation recipe used in the simulation.

Steps	**Recipe details**					
	Ion species	**Dosage (cm^3)**	**Implant energy (keV)**	**Tilt angle (°)**	**Rotation/ Twist angle (°)**	**Quad implant**
LDD Implant	B11+	5x10^{14}(total)	0.6	5	0	Yes
Pocket Implant	As75	2x10^{13}(total)	100	30	22	Yes
SD1 Implant	B11+	3x10^{15}(total)	3	7	0	Yes
SD2 Implant	B11+	3x10^{13}(total)	12	7	0	Yes

4.2 Experiment

This section discusses the experimental procedures and characterization techniques employed in this research work. Here, blanket (unpatterned) wafers were used to study the activation and diffusion of boron in P-polysilicon layer under different SD-RTP split. The processes involved are oxidation, deposition, implantation and annealing. The characterization techniques used in this research work are sheet resistance measurement, secondary ion mass spectrometry and focused ion beam scanning.

4.2.1 Process Flow of The Experiment

Figure 4.1 shows the process flow involved in this experiment.

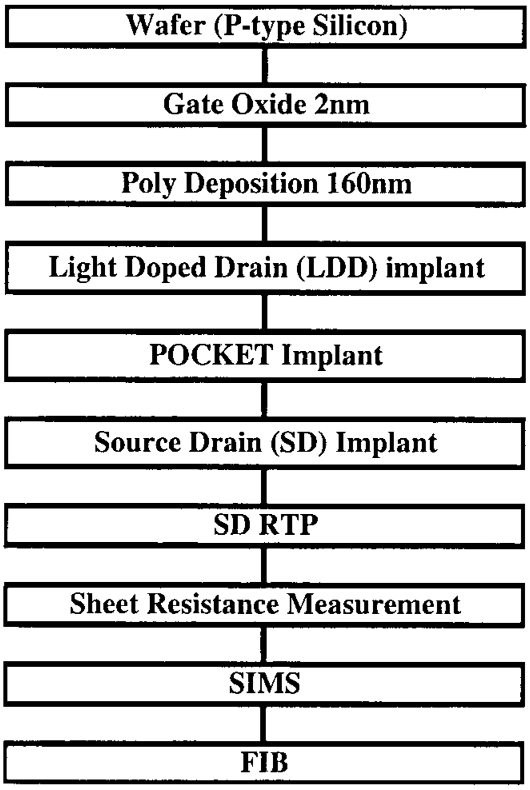


Figure 4.1. Process flow of the experiment.

4.2.2 Wafer

The starting material in the experiment is an 8-inch (200mm) silicon wafer with a thickness of 725 micrometer. The wafer used in this experiment is P-type (100) Czochralski wafer. The sheet resistance of the wafer is 10–20 Ω-cm. Prior to the first processing step; wafers are cleaned with Radio Corporation of America (hereafter abbreviated as RCA, please refer to Appendix B for the full recipe) solution in order to remove contaminants on the surface.

4.2.3 Gate Oxide Growth

This experiment begins with gate oxide growth step. Exposure to oxygen can cause a silicon surface to oxidize forming silicon dioxide (SiO_2). Native silicon dioxide is a high-quality electrical insulator and can be used as a barrier material during impurity implants or diffusion, for electrical isolation of semiconductor devices, as a component in MOS transistors, or as an interlayer dielectric in multilevel metallization structures such as multichip modules. The ability to form a native oxide is one of the primary processing considerations, which led to silicon becoming the dominant semiconductor material used in integrated circuits today. The thickness of the gate oxide in this experiment is only 2 nm. This thin gate oxide layer is the insulator layer for the sheet resistance measurement by using the four-point probe. Besides that, boron penetration effect can also be studied through the SIMS profile by checking the boron concentration at the oxide/silicon interface.

In every oxidation runs, the oxide furnace must contain 170 wafers. Wafers are loaded into the cassette storage unit vertically. Dummy wafers are used to complete the load pattern for every run if total test wafers are less than 170 wafers. All the dummy wafers must be loaded prior to any run. Temperature of the furnace is then increased to 650^{o}C with 10000 sccm flow rate of nitrogen and 100 sccm flow rate of oxygen purged into the furnace. Wafers are then loaded into the furnace. After the loading, the temperature is ramped-up to 750^{o}C. At temperature 750^{o}C, pressurized water vapour is purged into the furnace for wet oxidation reaction. The wet oxidation process acquired for 2 minutes will result in a desired thickness of 2 nm. Nitrogen is then purged into the furnace at the same temperature for gate oxide nitridation. This is a process of annealing called "NO annealing" which requires 30 minutes for gate nitridation reaction. After that, nitrogen annealing is carried out for 20 minutes to make sure the nitrided gate oxide is stable. Temperature is ramped down to 650^{o}C and the wafers are unloaded. Ellipsometer was used to check the thickness of the grown gate oxide.

4.2.4 Polysilicon Deposition

After the growth of gate oxide, the wafers will be deposited with a layer of polysilicon. In this experiment, the technology of 0.13-μm channel length transistor is used. The thickness of the deposited polysilicon was found to be about 160 nm. The deposition technique used for this purpose is LPCVD.

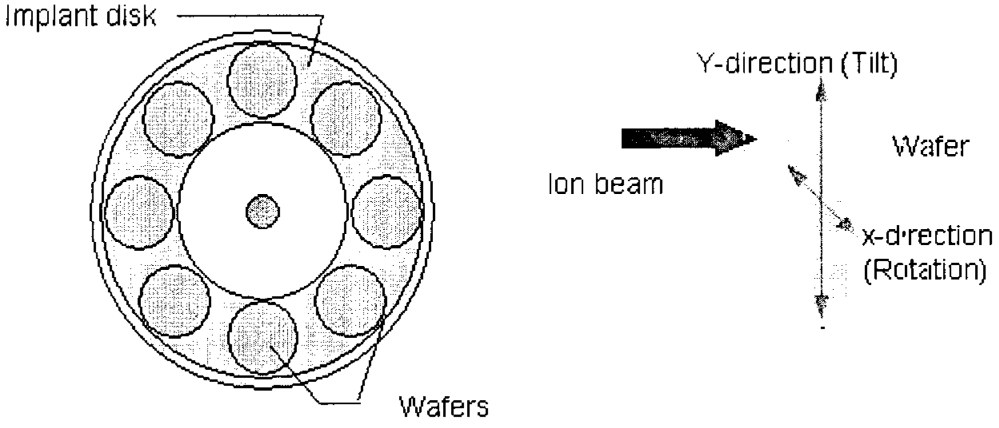
In every polysilicon deposition run, the LPCVD furnace must contain 175 wafers. Wafers were loaded into the cassette storage unit vertically. Dummy wafers are used to complete the load pattern for every run if the total test wafers are less than 170 wafers. All the dummy and test wafers must be loaded prior to the starting of any run. Temperature furnace is then raised to 620^{o}C with 2000 sccm nitrogen purged into the furnace. The temperature is maintained at 620^{o}C throughout the duration of the process. The wafers are loaded slowly into the furnace. SiH_4 is then purged into the furnace. The deposition step requires 16 minutes to deposit 160 nm poly silicon deposition. Lastly, the outlet of SiH_4 gas is shut off and the wafers are unloaded. The wafers are then sent to the ion implanter for the next processing.

4.2.5 Ion Implantation

Here, the deposited polysilicon will be implanted with several implantation steps. Before the start of the ion implantation process, a recipe needs to be created. The recipe will consist of parameters such as ion species, ion dosage, implant energy, implant angle (tilt and rotation control by implant disk, Figure 4.2 a & b) and selection of quad implant (quad implant referred to four, each 90^{o} rotation implant, total dosage needs to divide by four for quad implant). All the implantation processes are carried out in the process chamber maintain at vacuum. Wafers are first loaded into the process chamber. The wafers are held onto the implant disk as shown in Figure 4.2a. Recipe listed in the Table 4.1 is selected for different implant steps.

The first process is the Lightly Doped Drain (hereafter abbreviated as LDD) implant. LDD implant is used to reduce the high field in the drain junction of small junction geometry devices. In this step, the wafers are implanted by 5 x 10^{14}cm^3 dosage of ion Boron at 0.6keV implant energy. The implantation tilt angle is 5^o with 0^o rotation angle for quad implant. This is followed by the Pocket implant, where the wafers are implanted with 2x10^{13}cm^3 ion Arsenic at 100keV implant energy. This quad implant at rotation angle 22^o is tilted at an angle of 30^o. Pocket implant is known to suppress the short channel effects and gave high saturation currents and low off currents. Next process is the Source Drain (hereafter abbreviated as SD) implant for the formation of source drain region. There are two SD implants in this step. First is the SD1 implant with 3x10^{15}cm^3 ion Boron at 3kev implant energy. Next is the SD2 implant with 3x10^{13}cm^3 ion Boron at 12keV implant energy. Both SD implant steps are quad implanted at tilt angle 7^o and rotation angle 0^o.

Wafers are unloaded after the implantation and sent for the next processing at SD RTP.



(a) (b)

Figure 4.2. (a) Implant disc for ion implant tools. (b) Tilt and rotation angle for ion implantation.

4.2.6 Rapid Thermal Processing

RTP is a heat treatment process after the ion implantation. In RTP, the wafer is heated to repair the damage done to the lattice. The dopant ions will be activated and become part of the crystal lattice (activation). At the same time, the ions will also be distributed during the process. For this annealing process, wafer is loaded one by one into the furnace, with the preset process recipe.

Samples are divided into 10 Source Drain Rapid Thermal Processing (SD-RTP) recipes, which have been created earlier. . For each recipe, the temperatures are chosen based on the practical experience of the vendor and references from Applied Materials and literatures. In order to get higher activation of the dopant in the polysilicon and source drain region, the annealing temperature has to be higher than the current condition, which is 1000°C. From the literatures, it is also pointed out that the dopant profile is uniformly distributed after annealing temperature of 1080°C (Borisenko *et. al*, 1997). Hence, the temperatures of 1080oC and 1160oC have been chosen. For the ramp up rate, 75°C/s difference was chosen based on the past experience of Applied Materials. It is noticed that only the ramp-up rate of at least 75°C/s different, will there be a significant different in the dopant behaviour, namely diffusion and activation .

Each wafer is treated using different recipe as described in Table 4.2. Recipe SD 1000, non-spike annealing is selected as a reference to get the desired sheet resistance. Three different ramp-up rates can be found in recipe SD 1000. Fast ramp at 75°C/s from 650°C to 850°C is the first ramp up step. Followed by medium ramp at 50°C/s to reach to 900°C; slow ramp up at 25°C/s to peak temperature 1000°C. Wafer is annealed at peak temperature for 15 seconds. After the annealing, ramped down at 35°C/s from 1000°C to 600°C is carried out.

For spike annealing (recipe name that starts with SPK), the annealing profiles are as shown in the Figure 3.13. The process started with slow ramp at 25°C/s from 450°C to 550°C. The temperature is then stabilized for 2 seconds before the higher ramp up. There is zero dwell time for spike anneal. So, after reaching the peak temperature, high ramp down rate at 75°C/s is applied until the temperature reached 650°C. The wafer is then removed from the furnace. In this experiment, 9 spike annealing splits are created as listed in Table 4.2. It should be noted that during annealing, gas nitrogen is purged into furnace in order to prevent oxidation. For spike annealing ramp down, helium is also purged into the furnace with the purpose of increasing the ramp down rate.

In order to check the repeatability of the process, the experiment was conducted with two wafers treated at the same spike annealing condition for recipes SPK1080-175 and SPK1080-250.

Table 4.2. Spike annealing process recipe.

Recipe	Recipe details	
	Peak Temperature (°C)	Ramp up rate (°C/s)
SPK1000-100	1000	100
SPK1000-175	1000	175
SPK1000-250	1000	250
SPK1080-100	1080	100
SPK1080-175	1080	175
SPK1080-250	1080	250
SPK1160-100	1160	100
SPK1160-175	1160	175
SPK1160-250	1160	250

4.2.7 Sheet Resistance Measurement

In this experiment, four-point probe as shown in Figure 4.3 was used to measure the sheet resistance of the doped polysilicon. The measurement was done on 49 locations.

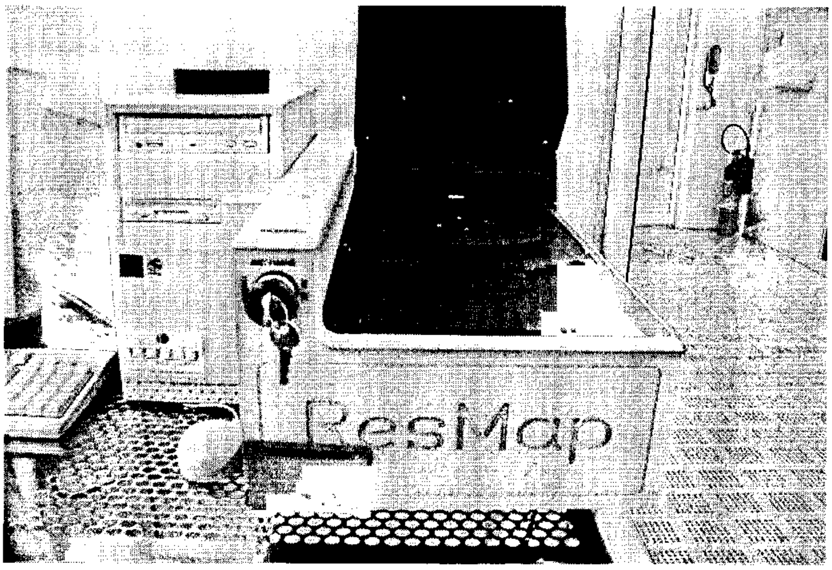
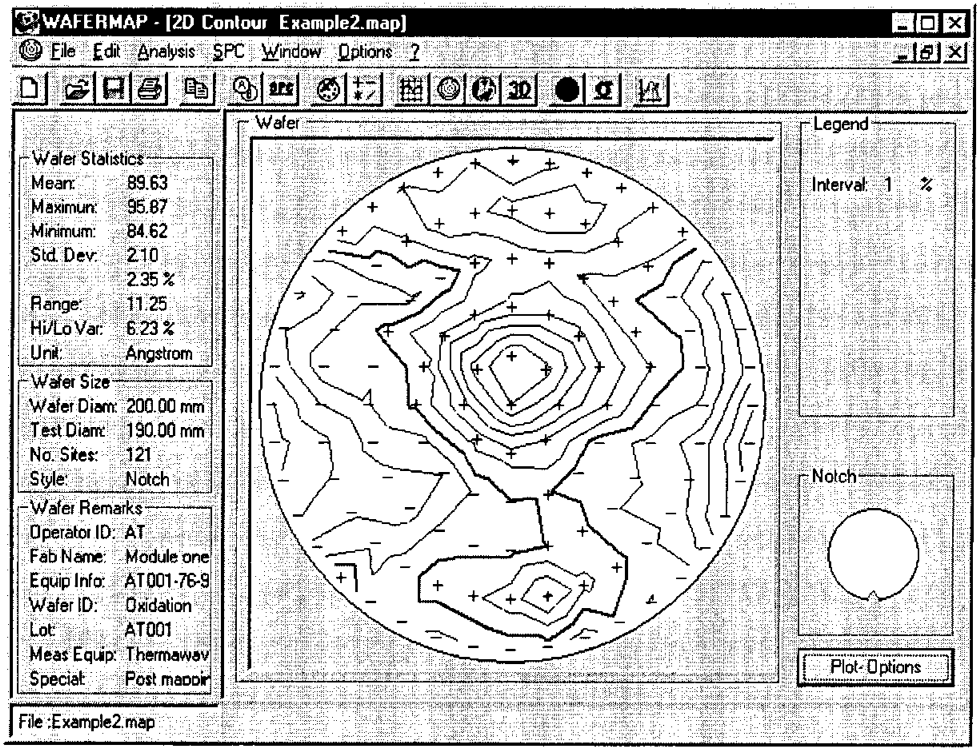


Figure 4.3. CDE ResMap four-point probe machine.

Wafer cassette is first loaded into the holder. The CDE ResMap four-point probe as shown in Figure 4.3 was used in this experiment. A wafer is transferred to the prober plate where 49 locations are tested. Result in contour map of the wafer sheet resistance value as shown in Figure 4.4 will appear in the monitor after the measurement. The contour map was plotted based on the sheet resistance value at the measured 49 points. The results in data sheet are then saved in the softcopy form.

4.2.8 Secondary Ion Mass Spectrometry

Silterra SIMS vendor, Charles & Evans, conducted SIMS profiling. Samples were identified and selected based on the sheet resistance results. Details such as dopant species, implant doses, implant energy, the expected dopant depth and full process of the samples were provided in order to get a more accurate result. Samples were also cut into small pieces at the centre of the wafer before delivered to the vendor in order to save the cost. The profiling takes 2 weeks to complete for normal price, extra charges will be needed if the duration of the profiling is shorter than 2 weeks. Results are available in both hard copy and soft copy (in the form of excel sheet).



Wafer Statistics
Mean: 89.63
Maximun: 95.87
Minimum: 84.62
Std. Dev: 2.10
2.35 %
Range: 11.25
Hi/Lo Var: 6.23 %
Unit: Angstrom

Wafer Size
Wafer Diam: 200.00 mm
Test Diam: 190.00 mm
No. Sites: 121
Style: Notch

Wafer Remarks
Operator ID: AT
Fab Name: Module one
Equip Info: AT001-76-9
Wafer ID: Oxidation
Lot: AT001
Meas Equip: Thermawav
Special: Post mappir

File :Example2.map

Figure 4.4. ResMap contour map after the sheet resistance measurement.

4.2.9 Focused Ion Beam

The remaining wafer pieces of the same samples that were sent for SIMS profiling were sent for FIB scanning. This is to further investigate the effect of spike annealing on the grain size of poly silicon. The FIB machine is switched on after the vacuum level has reached at 3.8 x 10^{-8} torr. Button "OK" at the machine panel is clicked to initialise the stage. Double-sided sticker is used on the broken wafer piece to attach it to the appropriate sample holder. All the sample materials must be able to withstand a high vacuum environment without out gassing. After the program has started, "UNLOAD" button has to be selected before the sample can be loaded. After the prepared sample has been loaded onto the load lock, the "LOAD" button is selected to have the sample transferred to the specimen chamber. The program is then switched to I-beam mode. The specific structure is moved to the centre of the display and a sharp beam image soon appears. "Couple Magnification" is selected to couple the magnification of the image. "Shift" knob is selected at the control panel in X and/or Y directions to centre the specific structure. The program is then switched to E-beam mode to obtain an image. "HV" button is selected to turn on the high voltage. E-beam button is selected to turn on the electron beam. Operation parameters, i.e. magnification, acceleration voltage, working distance, brightness contrast, astigmatism correction and focusing are adjusted to optimise the image. Lens alignment (if necessary) is performed to further enhance the image resolution. After scanning image has been obtained, the intercept technique is used to calculate the average grain size of the polysilicon.

CHAPTER 5

RESULT AND DISCUSSION

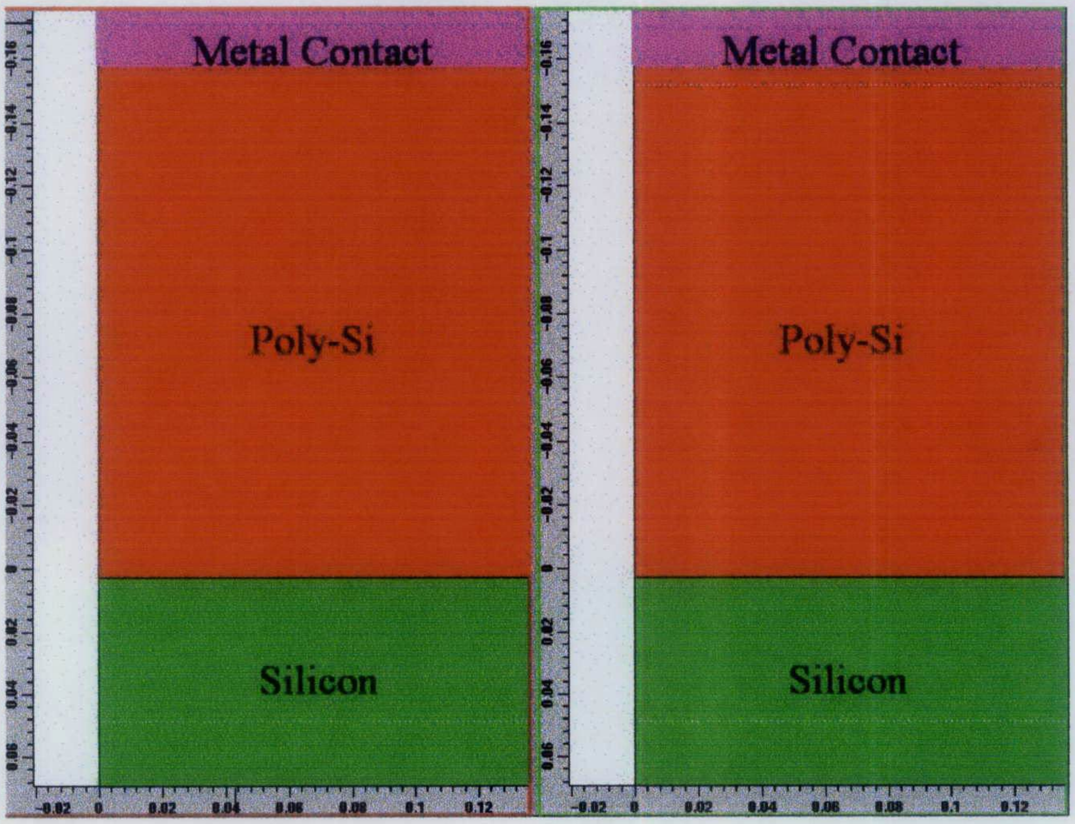
This chapter analyses and discusses the results obtained from both the simulation and experimental work. This includes a brief report on the industrial verification of the recommended process from the finding of this research work.

5.1 Simulation result

In the simulation work, cross-sectional view of the devices constructed will be displayed. The profile of the dopant distribution in the metal contact and polysilicon gate, dopant distributions as a function of the depth of the device had been observed and analysed. From this profile, the thickness of the depletion layer in the polysilicon region can also be derived. The simulation was conducted for two separate devices, PMOS and NMOS with the focus on two heat treatments namely standard SD-RTP and spike annealing. Spike annealing is a special RTP that does not have a dwell time at the peak temperature. The objective of adopting these two processes is to investigate the effect of spike annealing process on the distribution of the dopants. Here, comparison can be made on the distribution profiles of the dopants for both processes in the polysilicon layer. The other layer, i.e. silicon, metal contact (indicated as polysilicon2) can be neglected because the diffusion model used in this experiment is suitable for polysilicon material only. After the simulation, the cross-sectional view of the devices showing the dopant distributions were derived and displayed in the Taurus Visual. In Taurus Visual, various concentrations of impurity are indicated by different coloured region. The legend for various concentrations is indicated on the top right hand side of the plot. These different impurities are boron, phosphorus, arsenic and indium. Besides that, the boundary of the devices will be presented by the x- and y-axis, illustrating the width and length of the devices respectively. Different layer of the devices will be also labelled accordingly.

5.1.1 NMOS Simulation

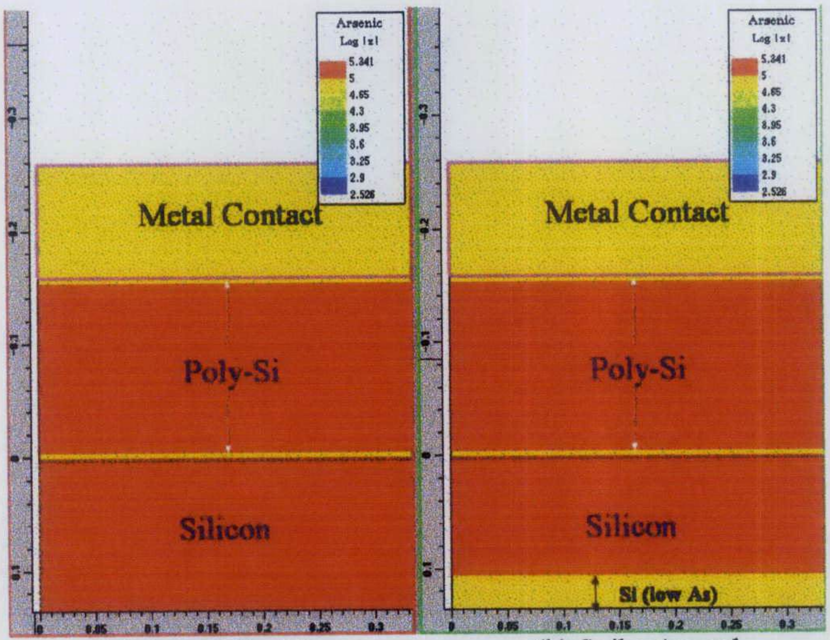
In NMOS simulation, the annealing conditions at SD-RTP are the standard RTP at 1000°C for 15 seconds and spike annealing at peak temperature 1080°C. In the software program of MEDICI, depletion region is obtained by applying an electric field to the device. From the cross-sectional view of NMOS device shown in the Figure 5.1, it is observed that no depletion region occurred at polysilicon and oxide interface for both processes. This is because as depicted in Figure 5.2 to 5.5, dopants (boron, arsenic, indium and phosphorus) in the polysilicon subjected to both processes are uniformly distributed with the same order of magnitude of concentration.



(a) Anneal for 15 sec

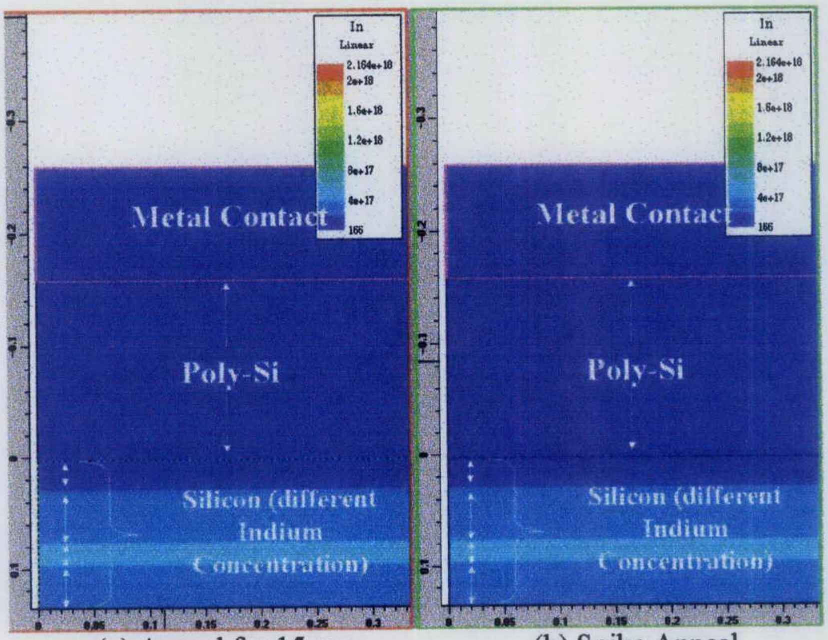
(b) Spike Anneal

Figure 5.1. Device cross-sectional view of NMOS from simulation for (a) RTP and (b) spike anneal.



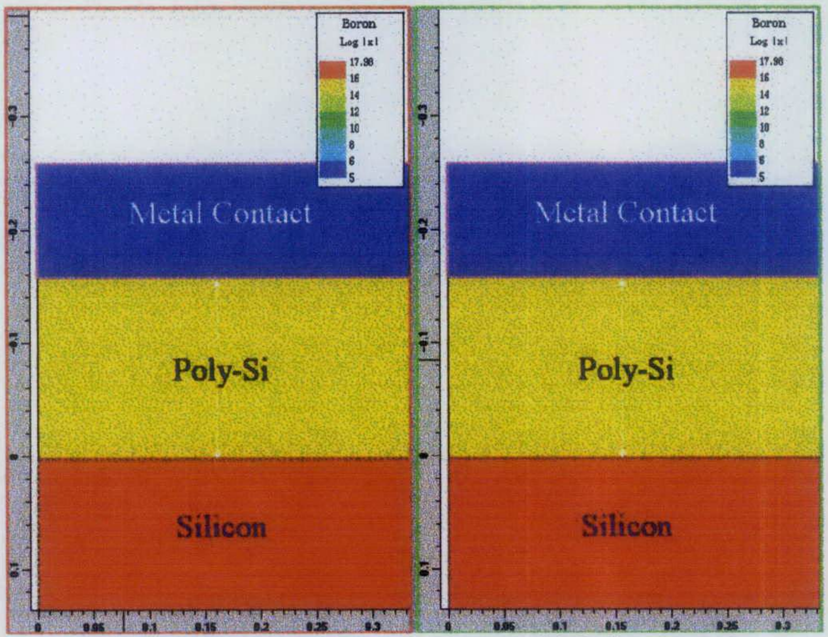
(a) Anneal for 15 sec (b) Spike Anneal

Figure 5.2. Distribution of arsenic in N-polysilicon for (a) Anneal for 15 sec and (b) Spike Anneal



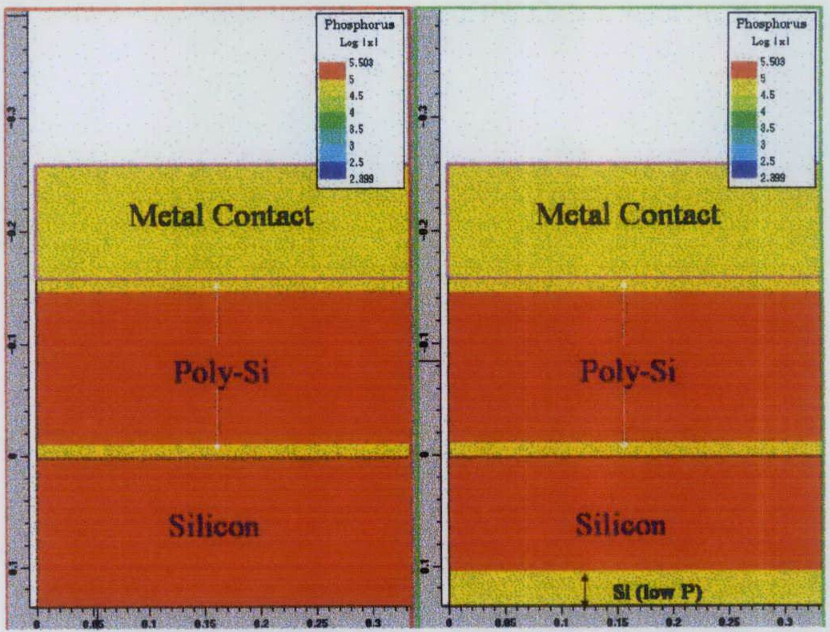
(a) Anneal for 15 sec (b) Spike Anneal

Figure 5.3. Distribution of indium in N-polysilicon for (a) Anneal for 15 sec and (b) Spike Anneal



(a) Anneal for 15 sec (b) Spike Anneal

Figure 5.4. Distribution of boron in N-polysilicon for (a) Anneal for 15 sec and (b) Spike Anneal



(a) Anneal for 15 sec (b) Spike Anneal

Figure 5.5. Distribution of phosphorus in N-polysilicon for (a) Anneal for 15 sec and (b) Spike Anneal

For a closer examination of the uniformity of the dopant concentration, a plot of concentration versus depth of the device was extracted from the cutline analysis conducted on the device cross-sectional output shown in Figure 5.1. Figure 5.6 and 5.7 show the plot for NMOS devices after the standard RTP treatment and spike annealing respectively.

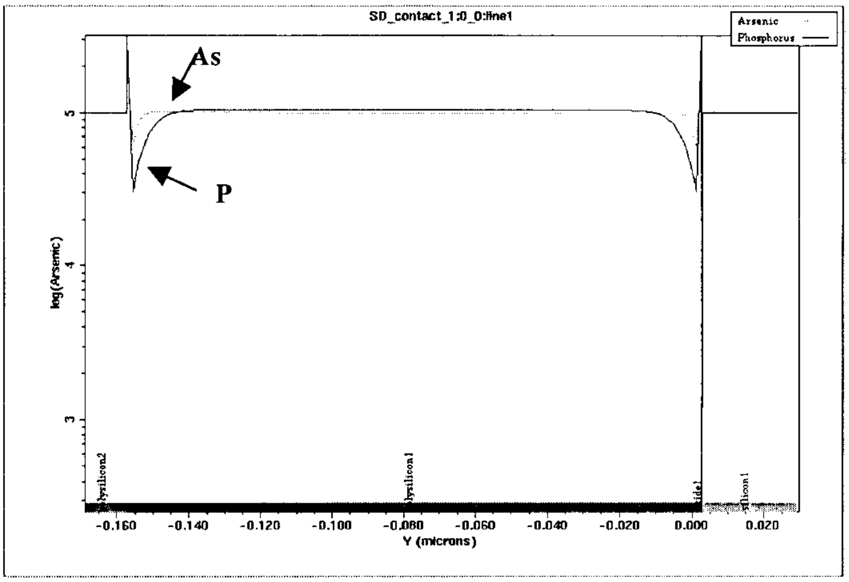


Figure 5.6. A plot of dopant concentration versus depth for NMOS after standard RTP.

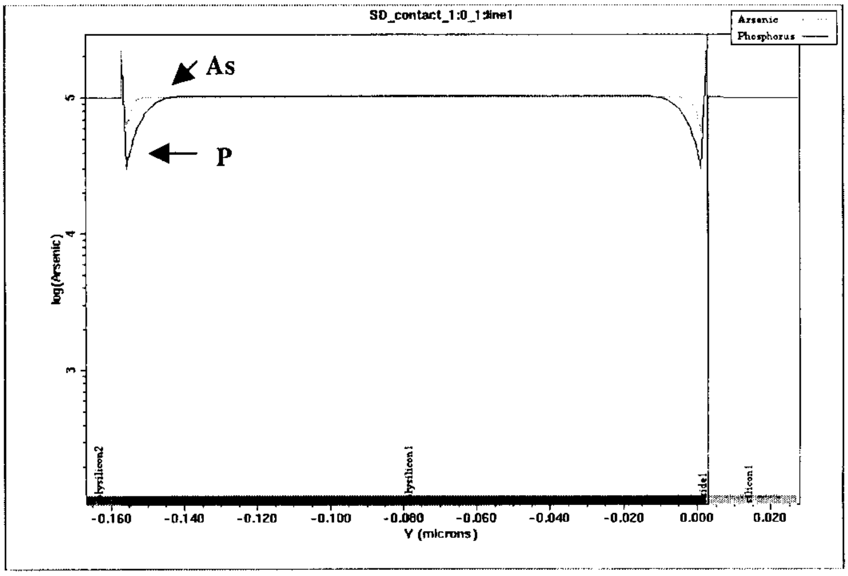


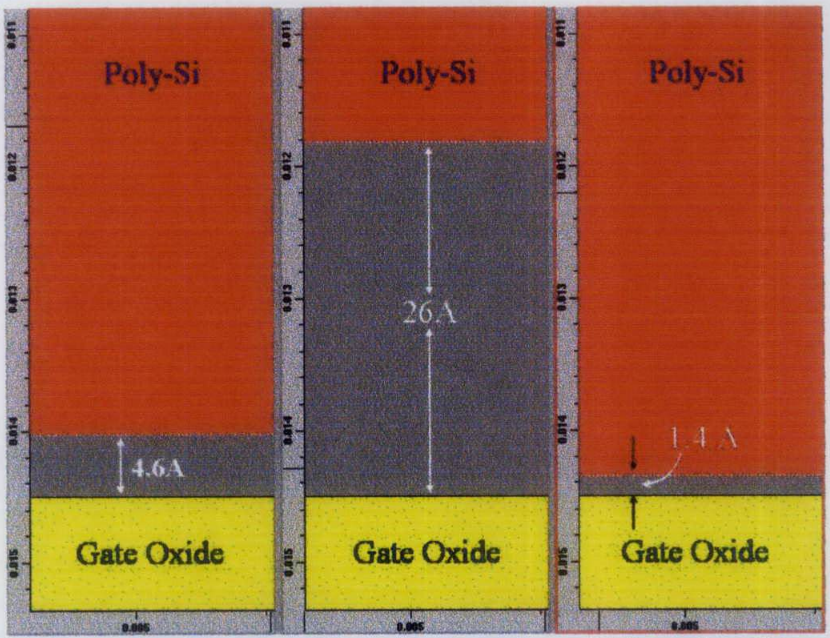
Figure 5.7. A plot of dopant concentration versus depth for NMOS after spike annealing

5.1.2 PMOS Simulation

In PMOS simulation, the annealing condition splits at SD-RTP are the standard RTP at 1000°C for 15 seconds, standard RTP at 1000°C for 30 seconds and spike annealing at peak temperature 1080°C. The purpose of adding one more annealing condition, standard RTP at 1000°C for 30 seconds is to investigate boron penetration problem at longer anneal time and to increase dopant activation.

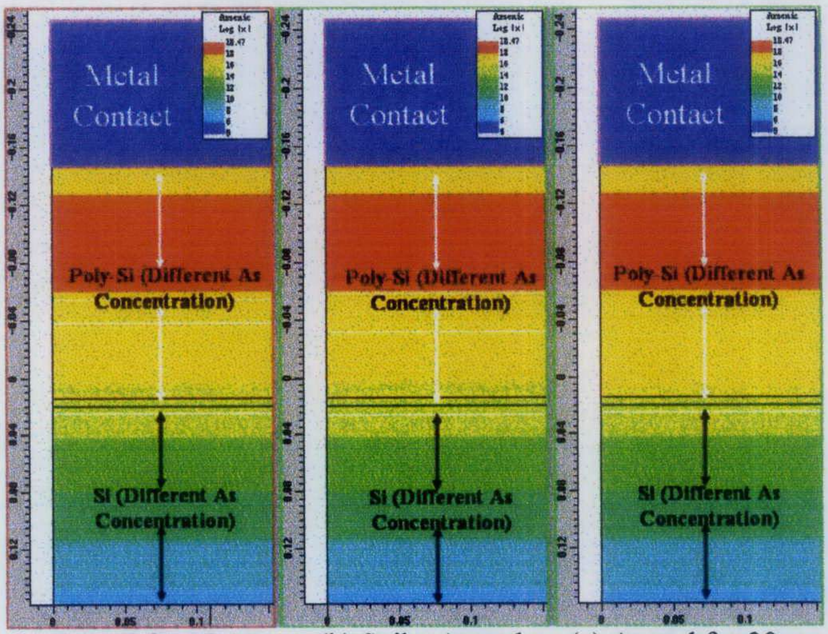
As in the previous case, depletion region can be obtained and displayed in the device cross-sectional view after an electric field has been applied to the device. It is observed that in the output of PMOS shown in Figure 5.8b, the thickness of the depletion region in polysilicon for spike annealing is 26 Å. For the standard RTP processes of 15 seconds and 30 seconds, the thicknesses of the depletion regions are 4.6 Å (Figure 5.8a) and 1.4 Å (Figure 5.8c), respectively.

In order to understand this observation, dopant distributions for each process were examined. It is shown clearly in Figure 5.9 that arsenic distribution in PMOS for the three processes is almost the same. The uniformity of arsenic distribution is to be expected because arsenic is considered heavier and has lower diffusion rate. Unlike arsenic profile distribution, boron and phosphorus have shown significant difference in their distributions for the three processes. The distribution profile for boron and phosphorus is depicted in Figure 5.10 and 5.11.



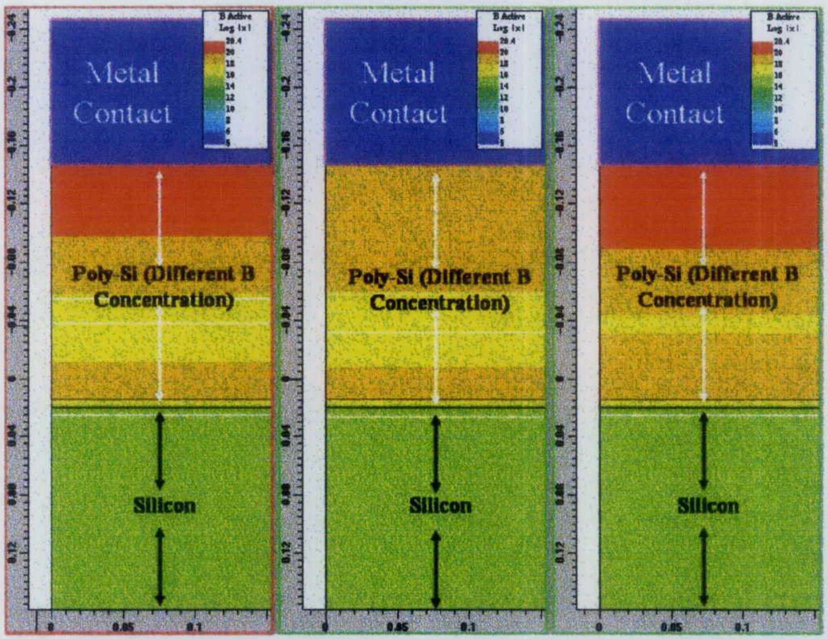
(a) Anneal for 15 sec (b) Spike Anneal (c)Anneal for 30 sec

Figure 5.8. Cross-sectional view of PMOS for (a) standard RTP for 15 seconds, (b) spike annealing and (c) standard RTP for 30 seconds.



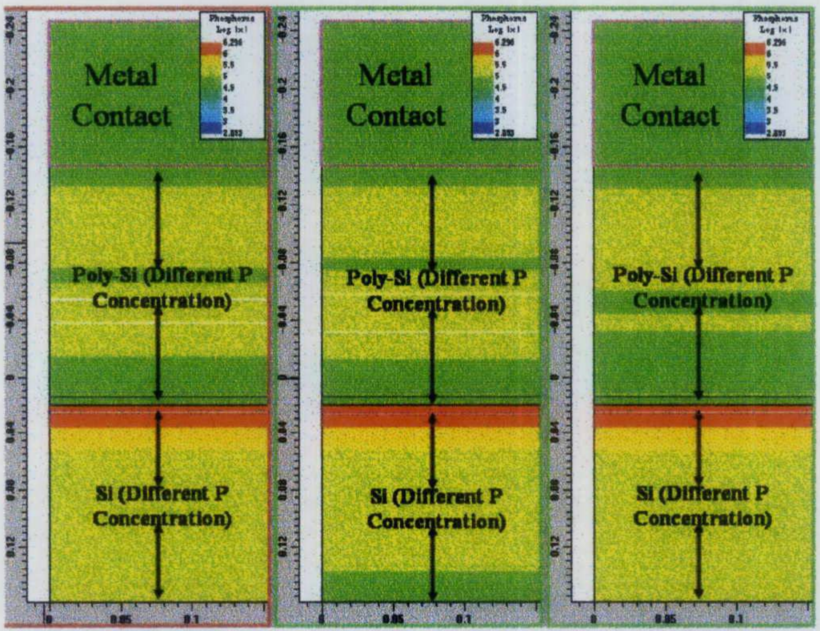
(a) Anneal for 15 sec (b) Spike Anneal (c) Anneal for 30 sec

Figure 5.9. Distribution of Arsenic in the P-polysilicon for (a) Anneal for 15 sec, (b) Spike Anneal and (c) Anneal for 30 sec.



(a) Anneal for 15 sec (b) Spike Anneal (c) Anneal for 30 sec

Figure 5.10. Distribution of boron in the P-polysilicon for (a) Anneal for 15 sec, (b) Spike Anneal and (c) Anneal for 30 sec.



(a) Anneal for 15 sec (b) Spike Anneal (c) Anneal for 30 sec

Figure 5.11. Distribution of phosphorus in the P-polysilicon for (a) Anneal for 15 sec, (b) Spike Anneal and (c) Anneal for 30 sec.

For a closer examination of the uniformity of the dopant concentration, a plot of concentration versus depth of the device was extracted from the cutline analysis conducted on the device cross-sectional output in Figure 5.9. Figure 5.12, Figure 5.13 and Figure 5.14 described the plot for PMOS devices after the standard RTP treatment for 15 seconds, standard RTP treatment for 30 seconds and spike annealing respectively.

From the plots, it is observed that boron concentration at polysilicon and oxide interface decreases when the duration of the annealing process is reduced. This results in a thicker polydepletion region as illustrated in Figure 5.9. It is noticed in Figure 5.13 that when boron concentration at the polysilicon and oxide interface reached 10^{20} cm^{-3}, the depletion region is reduced. Similar observation is also reported by other literature (Yu Bin, 1998).

Simulation output has shown that polydepletion problem occurred only in PMOS devices. It is also verified that uniform dopant distribution in the polysilicon and highly doped polysilicon will be able to reduce the polydepletion effect. In the simulation, the effect of the spike annealing treatment had been studied. The comparison result for different heat treatment had shown that spike annealing would reduce the activation and diffusion of the dopant in the polysilicon. This is considered to be the drawback of using the spike annealing process. However, due to the fact that spike annealing is still a new and special process, an experimental work need to be performed in order to verify the simulation result obtained here.

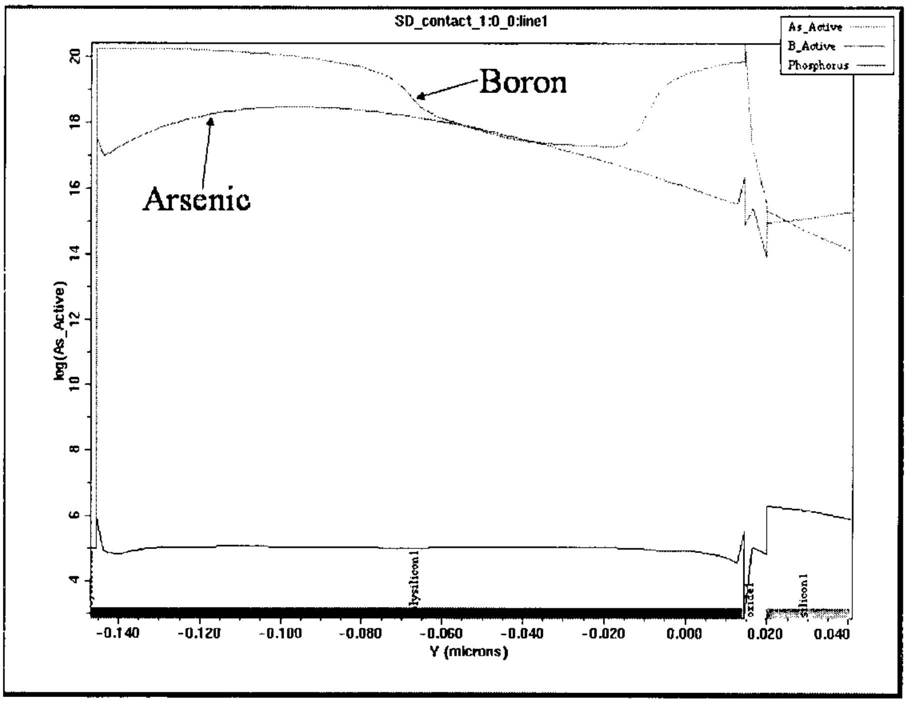


Figure 5.12. A plot of dopant concentration versus depth for PMOS after standard RTP for 15 seconds.

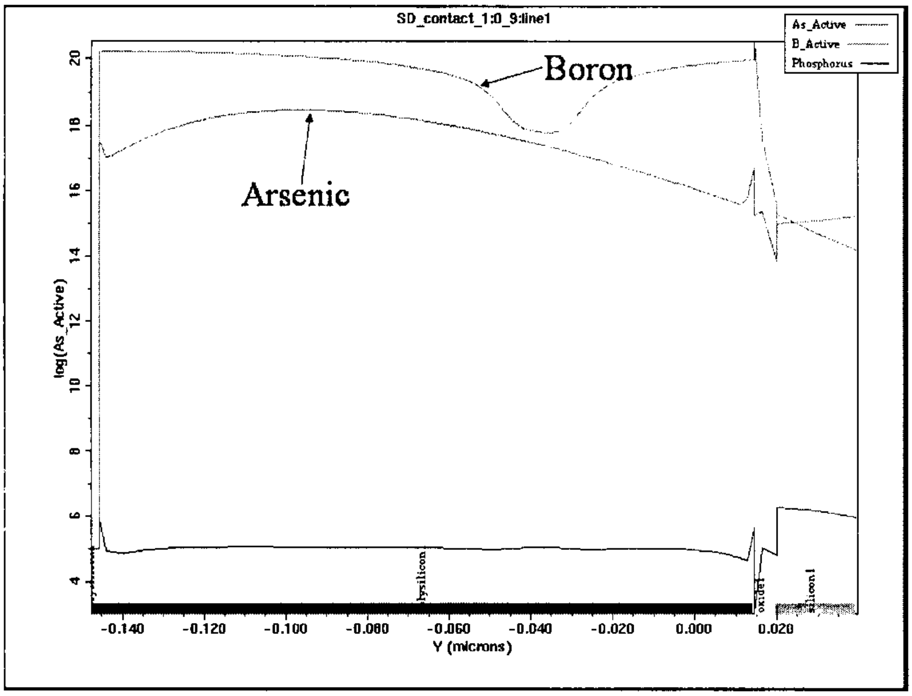


Figure 5.13. A plot of dopant concentration versus depth for PMOS after standard RTP 30 seconds.

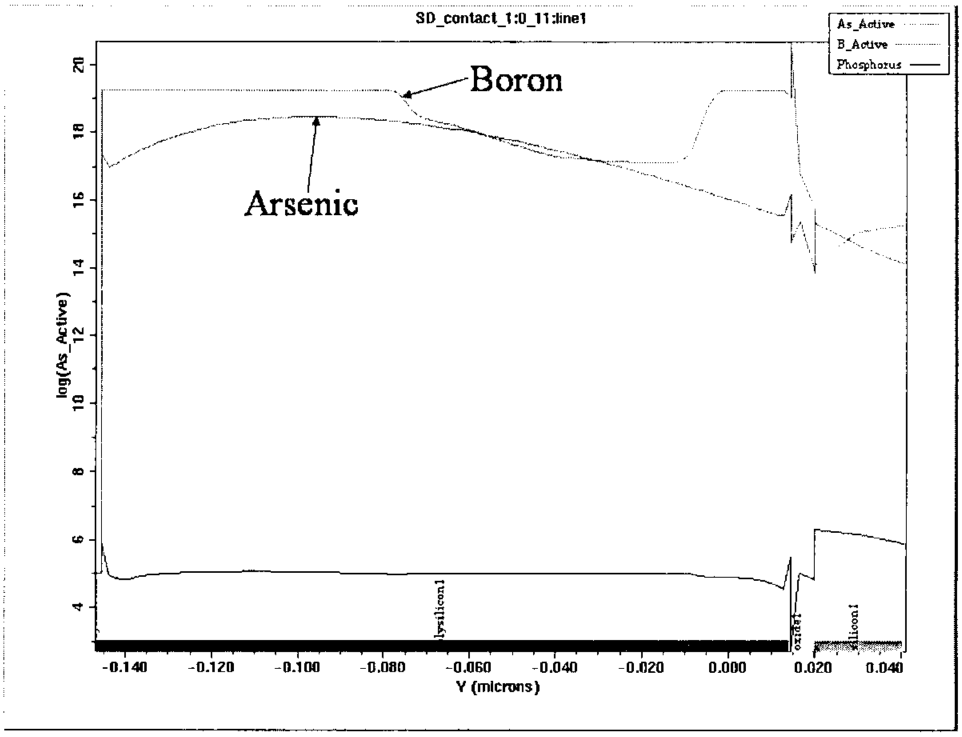


Figure 5.14. A plot of dopant concentration versus depth for PMOS after spike annealing

5.2 Experimental Result

In this experiment, the effect of spike annealing on boron activation and diffusion in P-polysilicon layer was studied. The experiment was conducted on the wafer, which have been divided into 10 splits corresponding to different conditions in SD RTP step. In this step, 9 splits were subjected to the spike annealing process with 3 ramp-up rates and 3 peak temperatures, and one to the standard RTP for reference. Sheet resistance measurement, SIMS and FIB were used to characterize the treated polysilicon for PMOS devices.

5.2.1 Sheet Resistance

The sheet resistance measurements had been carried out at room temperature on polysilicon surface at 49 locations by CDE ResMap four-point probe as recommended by the CDE ResMap vendor. The mean value of the sheet resistance for each wafer is shown in Table 5.1.

For IC layout, it is convenient to work with a parameter called sheet resistance R_S. For a region of length, L and width, W the sheet resistance is found from

$$R = \left(\frac{1}{q\mu_n N_d}\right)\frac{L}{Wt} = \frac{L}{W}\left(\frac{1}{q\mu_n N_d t}\right) = \frac{L}{W}R_s \qquad \text{(Eq 5.1)}$$

where L/W is the number of squares, N_d is the activated dopant concentration. Non-rectangular regions can be modeled by an effective number of squares (Sze, 1985).

Table 5.1. Sheet resistance mean value for different wafer with different spike anneal split.

Wafer number (VID)	Spike anneal split	Measured R_s (Mean)
14	SD1000 (reference, non spike anneal treatment)	282.039
15	SPK1160-250	163.221
16	SPK1160-175	161.730
17	SPK1160-100	164.230
18	SPK1080-250	253.588
19	SPK1080-250	253.959
20	SPK1080-175	246.509
21	SPK1080-175	246.327
22	SPK1080-100	246.112
23	SPK1000-250	559.837
24	SPK1000-175	529.239
25	SPK1000-100	520.351

From the equation 5.1, it is observed that the dopant activation is inversely proportional to sheet resistance; it means that higher dopant activation gives lower sheet resistance value. It is observed from the lower sheet resistance value in Table 5.1 that higher spike annealing peak temperature has higher dopant activation. However spike annealing at peak temperature of 1080°C is considered the most suitable heat treatment for 0.13-μm CMOS devices because the sheet resistance value is the closest to the reference value given in Figure 5.15. Besides that, the sheet resistance values for VID 18, 19, 20 and 21 also showed the good reproducibility of the experiment. Generally, the sheet resistance values shown can be categorized into four different groups, which are referred to as the reference point, spike 1000°C, spike 1080°C and spike 1160°C. It is observed that peak temperature has a much greater effect in dopant activation than ramp up rate. For a closer analysis of the dopant activation in polysilicon after the heat treatment, a plot of sheet resistance versus peak temperature and ramp-up rate for spike annealing wafers are shown in Figure 5.16 and 5.17 respectively.

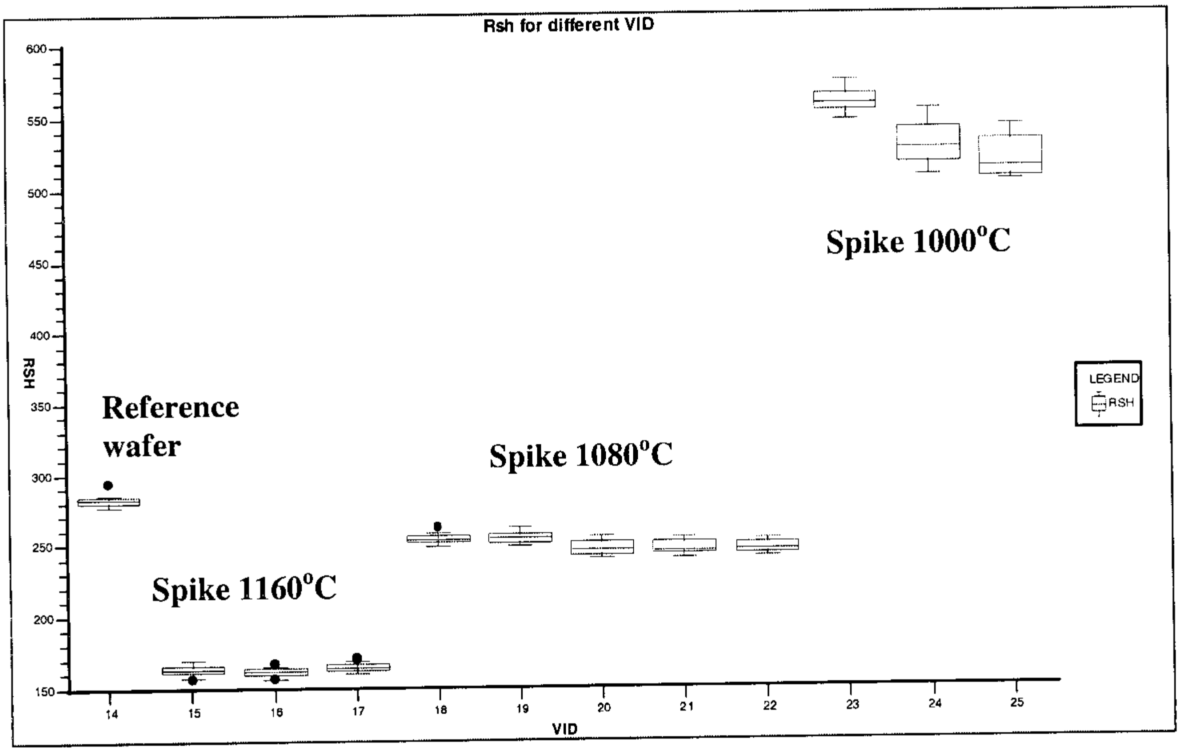


Figure 5.15. Box plot of sheet resistance for different wafer (VID).

In Figure 5.16, it is noted that there is no significant difference in R_s values with different ramp-up rate at different peak temperature. However, in Figure 5.17, R_s values for different peak temperatures at different ramp-up vary significantly. It is obvious that peak temperature has a much greater effect than ramp up rate. When the peak temperature increased, R_S decreased. This is believed to be due to the increase in the activation of boron at higher temperatures (Sze, 1985). Peak temperature of 1080°C is the best choice because the result is close to the targeted value of the reference wafer, which is 260 Ω/ .

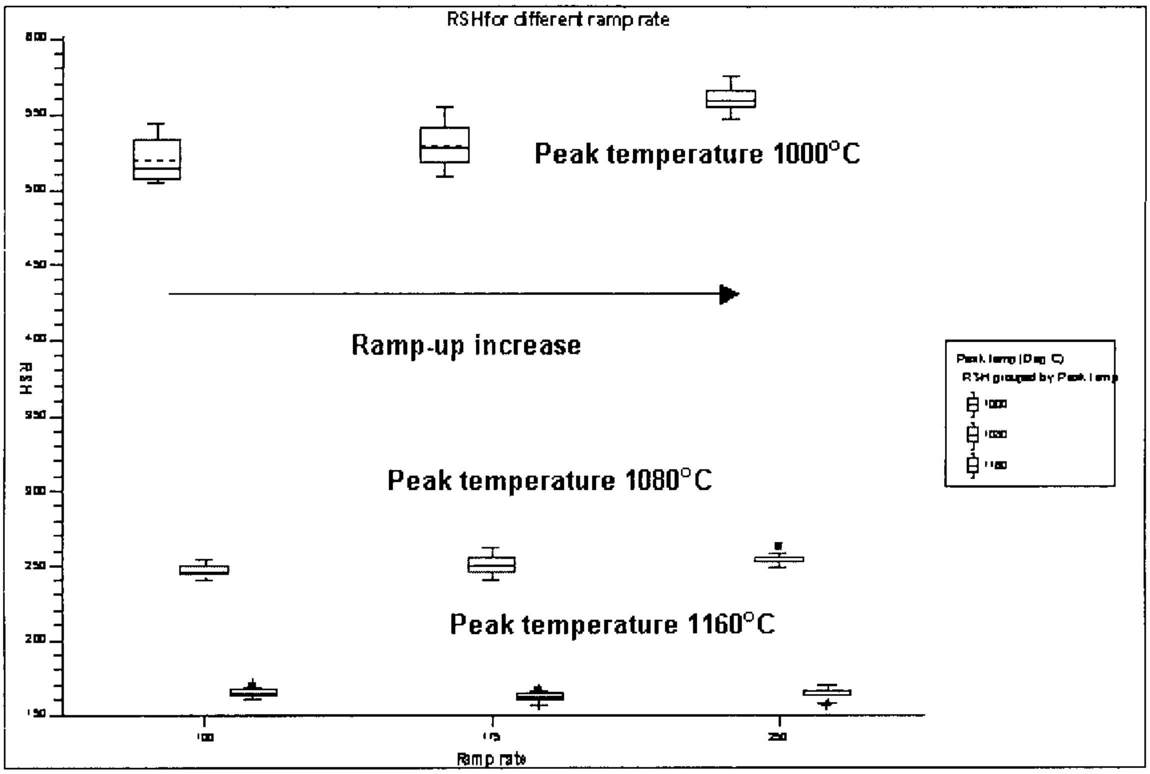


Figure 5.16. A plot of R_S against ramp-up rate for different peak temperature.

5.2.2 SIMS Profile

In SIMS profiles, it can be seen that both standard RTP 1000°C for 15 seconds (Figure 5.18) and spike anneal at 1080°C (Figure 5.20) has even distribution of boron in the polysilicon. Whilst for the spike anneal at peak temperature 1000°C (Figure

5.19), the boron profile is decreasing when the depth is increasing. Boron atoms are not evenly distributed in the polysilicon and the boron concentration at the polysilicon/oxide interface is less than 10^{20} cm^{-3}. This will consequently lead to poly depletion effect.

In SIMS profile, the boron penetration problem was also investigated. A comparison of Figure 5.18 and Figure 5.20 shows the drop in the boron concentration at the oxide/silicon interface when the sample was treated with spike annealing. The boron concentration at oxide/silicon interface has been reduced from 3.05 x 10^{17} cm^{-3} to 1.80 x 10^{17} cm^{-3} giving a 41% reduction of boron penetration if spike annealing is used in SD RTP to replace the standard RTP.

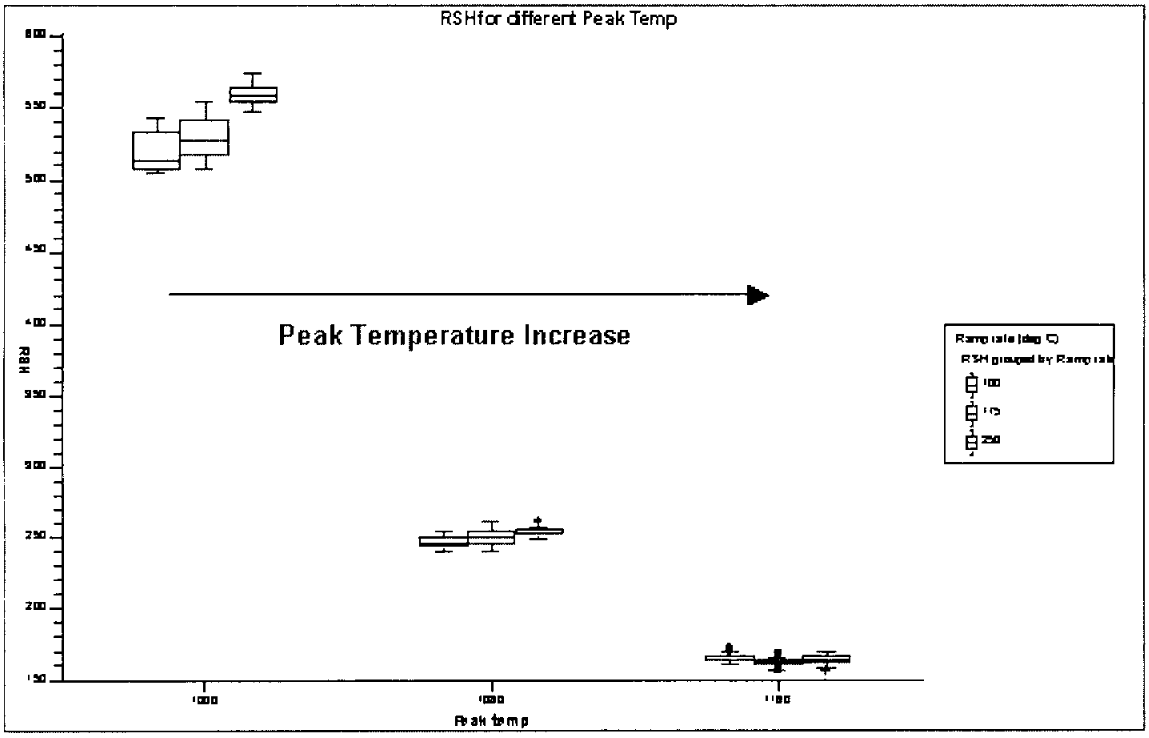


Figure 5.17. A plot of R_S against peak temperature for different ramp-up rate.

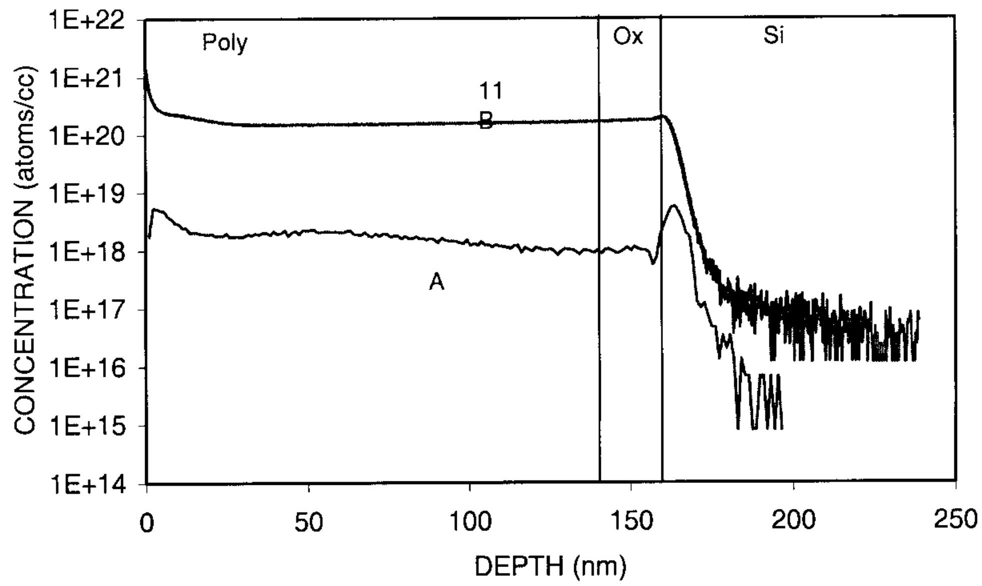


Figure 5.18. SIMS for boron and arsenic (from Pocket implant) in polysilicon after Standard RTP at 1000°C for 15 seconds.

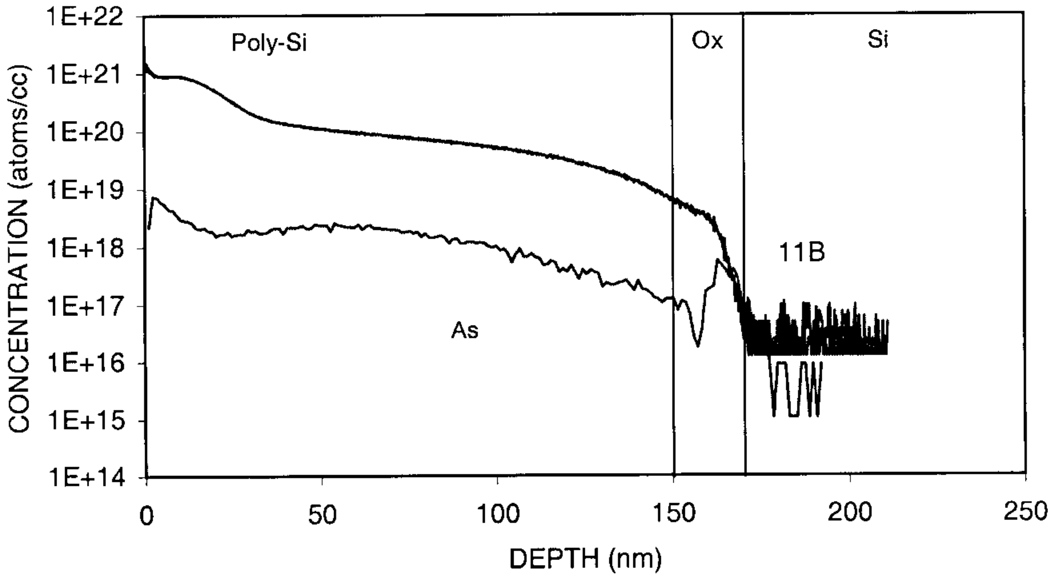


Figure 5.19. SIMS for boron and arsenic (from Pocket implant) in polysilicon after spike annealing at peak temperature 1000°C.

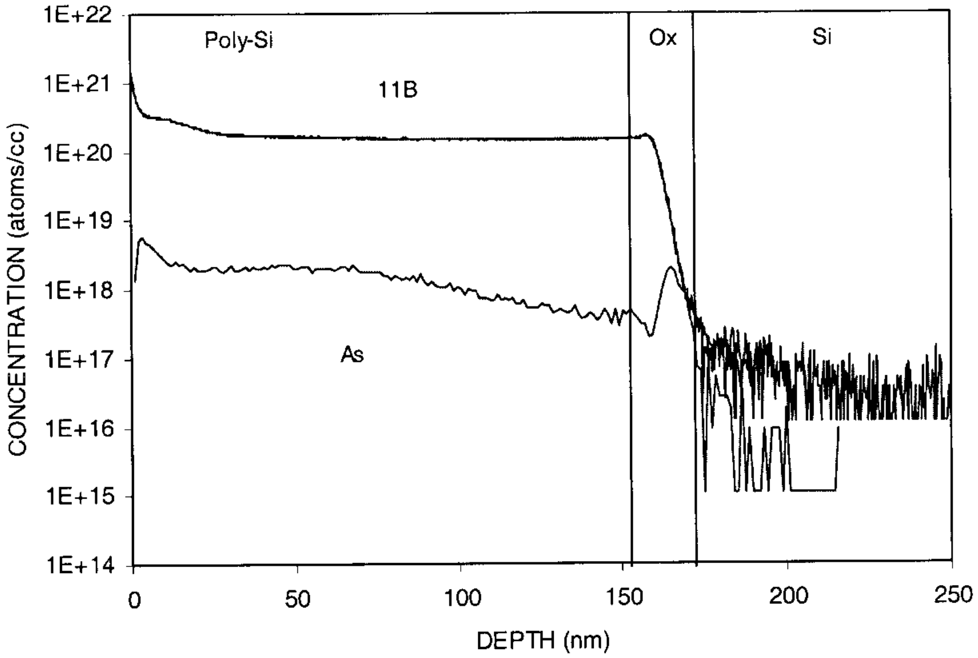


Figure 5.20. SIMS for boron and arsenic (from Pocket implant) in polysilicon after spike annealing at peak temperature 1080°C.

5.2.3 FIB Scanning

Figures 5.21 to 5.23 describe the FIB image on polysilicon surface after been treated with standard RTP and spike annealing. In Figure 5.21, 5.22 and 5.23, the average grain size was calculated by intercept method. The average grain size for standard RTP, spike annealing at peak temperature 1000°C and spike annealing at peak temperature 1080°C is 113.5 nm, 89.6 nm and 106.42 nm respectively. It can be deduced that higher peak temperature or longer anneal time could produce larger polysilicon grain. Sintering of the grains of the polysilicon can actually change the sheet resistance. Sintering would lead to increase of contact area between adjacent grains, thus, leading to mainly geometrical changes that do not affect significantly the physical properties of grain boundaries such as Temperature Coefficient of Grain Boundary Resistance (Hereafter abbreviated as TCR_{gb}) (Spoutai *et. al*, 1998).

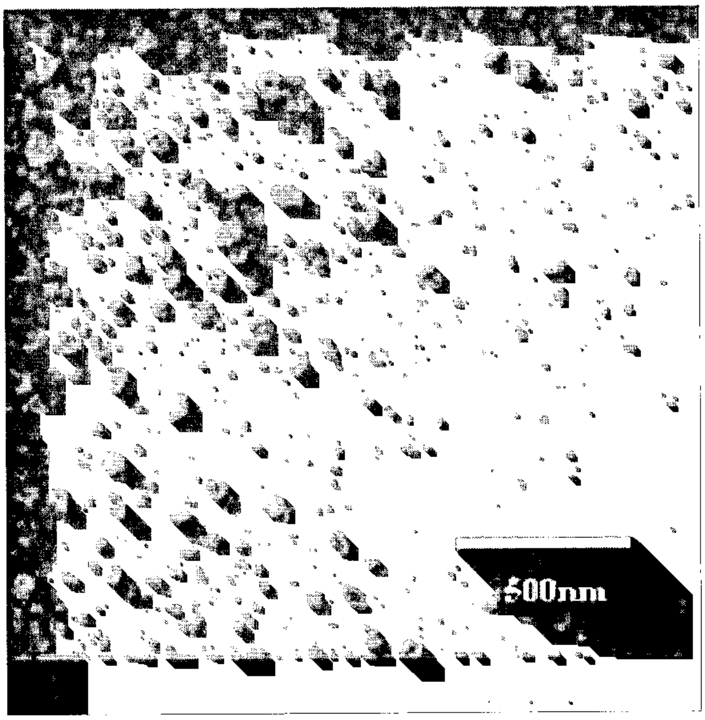


Figure 5.21. FIB image of polysilicon grain size after standard RTP for 15 seconds at 1000°C. d_{grain}=113.5 nm

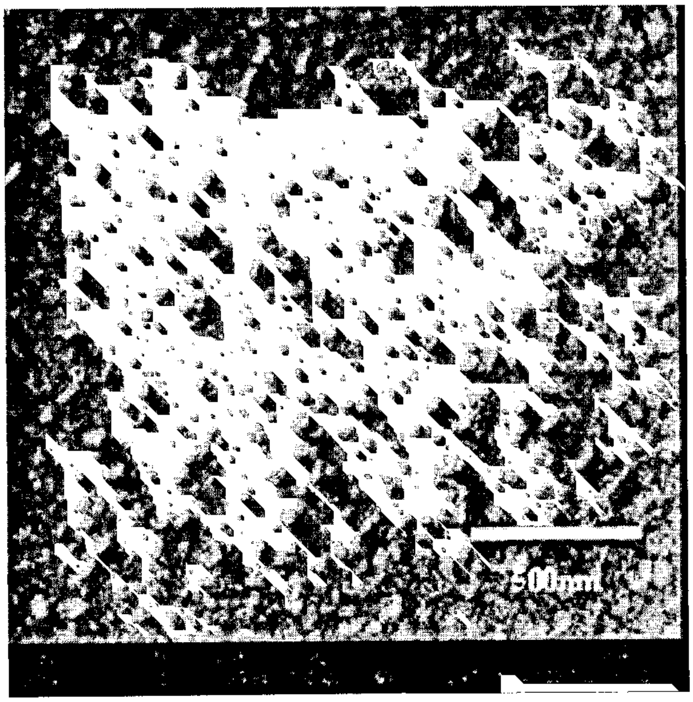


Figure 5.22. FIB image of polysilicon grain size after spike anneal at peak temperature 1000 °C. d_{grain}=89.6 nm

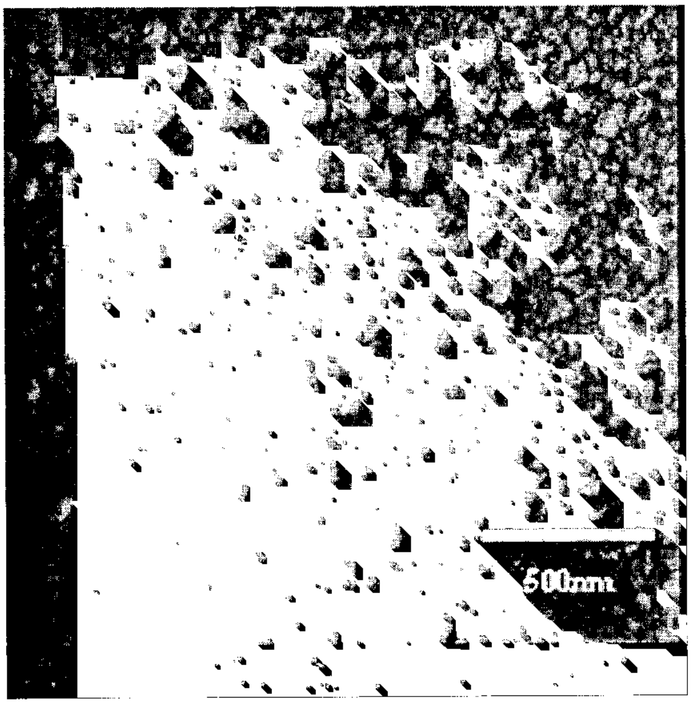


Figure 5.23. FIB image of polysilicon grain size after spike anneal at peak temperature 1080 °C. d_{grain}=106.42 nm

5.3 Industrial Verification

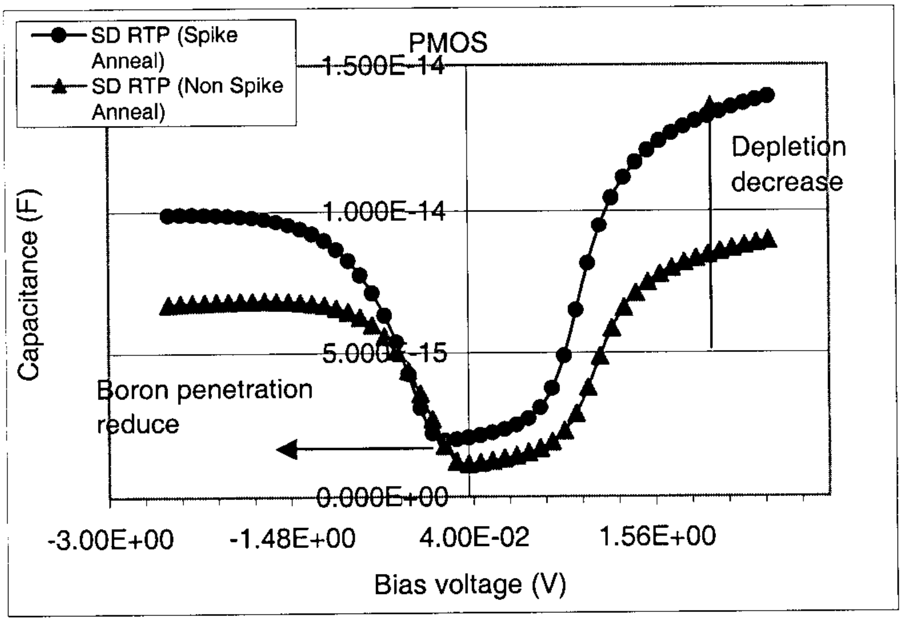


Figure 5.24. Comparison C-V curves between reference wafer (Non spike annealing) and wafer spike annealing at peak temperature 1080°C.

Spike annealing at peak temperature 1080°C and ramp-up rate 175°C/s had been proposed to Silterra (M) Sdn Bhd as the fabrication process for the polysilicon gate of 0.13-μm CMOS devices. The proposal had been adopted and fabrication process on the 0.13-μm CMOS devices was conducted. The fabricated device was sent for electrical characterization and the outcome is depicted in Figure 5.24. The C-V curve shows a significant increase in the capacitance of the PMOS device at the inversion region if treated with spike annealing. This would also imply that the thickness of the depletion region was reduced as discussed in chapter 2.1. Besides that, the threshold is also shifted to the negative side and it is understood from chapter 2.2 that boron penetration would be suppressed. Thus spike annealing is the best and proven RTP method to be used in the 0.13-μm CMOS device processing.

CHAPTER 6

SUMMARY, CONCLUSION AND RECOMMENDATION

This chapter is divided into 3 parts. The first part is the summary of the research findings. Second-part describes the conclusion of the research followed by the third, which deals with the recommendation for future research.

6.1 Summary

From the simulation results, it is noticed that polydepletion problem occurred only in PMOS devices. It is also shown that uniform dopant distribution in the polysilicon and highly doped polysilicon will reduce the polydepletion effect. In the simulation, the effect of the spike anneal had been investigated. The comparison result for different heat treatments had shown that spike annealing could actually reduce the activation and diffusion of the dopant in the polysilicon. However, due to the fact that spike annealing is still a new and a special process, an experiment needs to be conducted in order to verify the simulation result.

From the experiment results, it was found that spike annealing reduced the dopant diffusion rate and increased the activation of the dopant, which could reduced the polydepletion effect and suppressed the boron penetration. It can be shown in the sheet resistance value and SIMS profile. Ramp-up rate does not have much influence on dopant activation and diffusion. Higher ramp-up rate can be proposed to reduce the process cycle time and thermal budget for cost saving. The activation of the dopant under spike annealing had been proven by the sheet resistance measurement. The wafer was tested at 49 locations by using four-point probe (Figure 5.15 to 5.17). After the comparison of the sheet resistance value between the reference wafer and spike annealing at peak temperature 1080°C, the latter is found to be the best choice for the 0.13-μm CMOS devices. The decrease of sheet resistance was due to the increase of carrier mobility and activated carrier concentration. The sheet resistance of polysilicon film decreased significantly since carrier mobility and activated carrier concentration of poly-Si films increased significantly. (Kang *et. al*, 2003)

Further investigation by SIMS was carried out to analyze the dopant distribution in the polysilicon. Figures 5.18 and 5.20 showed dopant distributions after the processes. For both the reference sample and spike anneal at peak temperature 1080°C, boron atoms are evenly distributed while for spike anneal at peak temperature 1000°C as shown in Figure 5.19, the boron distribution is not uniform. It is mentioned in Borisenko *et. al*, (1997) which electrical activation occurs when impurity atoms migrated to a substitutional position inside the grains. Heat processing leads to point defect annealing and to intragrain recrystallisation in implanted layers. A diffusion length of several interatomic spacing is sufficient for a fraction of impurity atoms to reach the nearest grain boundary and segregate there. These processes are the fastest, occurring at processing times as short as 10^{-2} s when the induced temperature is 600-800°C. The electrical activation and segregation of the impurity atoms appear to be closely correlated.

From the experiment, it can be concluded that recrystallisation inside the grains takes place in the first heating. The movement of impurity atoms accompanies the recrystallisation movement to substitutional positions, where they display electrical activity. Thus, impurity diffusion and segregation are the processes that could achieve the balance.

Figure 5.21, 5.22 and 5.23 clearly shows that polysilicon grains grow at longer annealing duration or higher peak temperature. Borisenko *et. al*, (1997) mentioned that grain growth in substitutional impurity doped polysilicon is followed by an increase of carrier mobility and hence a material with lower resistivity will be obtained. Annealing would lead to an increase of contact area between adjacent grains, thus, leading to mainly geometrical changes that do not affect significantly the physical properties of grain boundaries such as TCR_{gb} (Spoutai *et. al*, 1998). With larger grain size, the dopant diffusion is faster compared to smaller grain size due to the fact that the ability to absorb interstitials decreases (Kamins, 1998). When the interstitial atom diffusion decreases, dopant activation will be higher. This is because interstitial atoms cannot be activated easily. It is also believed that dopant segregation at the grain boundaries also decreases when the average grain size is larger.

6.2 Conclusion

In this work, the effectiveness of the spike annealing in reducing the polydepletion effect and suppress the boron penetration with the combination of nitrided gate oxide was studied. Besides that, an optimum spike annealing conditions for 0.13-μm CMOS devices was found.

Before any experiment was started to optimize the process parameters, process simulation and device characterization tools (TSUPREM 4 and MEDICI) were used to simulate different process parameters to get a targeted value for next generation CMOS devices, in this case, the 0.13-μm CMOS devices. Several simulations with different implant doses and energy, and different heat treatments had been carried out before performing any physical experiment.

Through the experiment, the behavior of the dopant activation and diffusion in the polysilicon under different heat treatment cycles was studied. It was found that the annealing duration and temperature mainly affect the activation and diffusion of a dopant in polysilicon. Spike annealing ramp-up rate almost has no effect at all on the boron activation and diffusion. Spike annealing peak temperature is the key parameter that controls the activation and diffusion. This can be seen from the sheet resistance value measured using four-point probe and SIMS. From the literature and simulation result, it is shown that polydepletion could be reduced if the dopant concentration is high ($\approx 10^{20}$ cm^{-3}) and uniformly distributed in the polysilicon gate. In this study, both spike anneal at 1080°C and normal RTP at 1000°C for 15 seconds gave a similar SIMS profile. These processing conditions are considered to be the better choice for 0.13-μm CMOS devices. However, wafer subjected to normal RTP at 1000°C for 15 seconds has higher boron concentration compared to the wafer treated to spike anneal at 1080°C at gate oxide/silicon interface. A total of 41% reduction has been reported in the wafer subjected to spike annealing at 1080°C. FIB images have added another finding that higher spike anneal temperature as well as longer duration could actually increase the polysilicon grain growth. It is deduced that grain growth in substitutional impurity doped polysilicon will be followed by an increase of dopant mobility and hence producing material with a lower resistivity. Annealing would lead to an increase of contact area between adjacent grains, thus, leading to mainly geometrical changes that do not affect significantly the physical properties of grain boundaries. With larger grain size, dopant diffusion is faster compared to smaller grain size. This is attributed to the reduced ability to absorb interstitials. A decrease in the diffusion of interstitial atoms would also mean a decrease in the number of interstitial atoms. This would allow for more substitutional diffusion to occur thus increasing the dopant activation. It is also understood that dopant segregation at the grain boundaries would decrease if the average grain size is larger.

6.3 Recommendation

Spike annealing at peak temperature 1080°C and higher ramp-up rate 250°C/s with nitrided gate oxide is recommended to be adopted as the SD-RTP for the 0.13-µm CMOS devices. It is because significant improvement has been shown in the finding where both polydepletion effect and boron penetration effect have been reduced.

However to strengthen the finding, further investigation is required to study the doping in order to reduce the quantum mechanical effect and leakage problem that may occur. Achievement of this work can encourage process engineers to use this spike annealing technique as a fundamental study to design the processing of the next generation CMOS devices.

BIBLIOGRAPHY

1 Ang C.H., Tan S.S., Lek C.M., Lin W., Zheng Z.J., Chen T. and Cho B.J., "Suppression of Nitridation-Induced Interface Traps and Hole Mobility Degradation by Nitrogen Plasma Nitridation", Electrochemical and Solid-State Letters, 5(4), 2002, pp G26-G28.

2 ASTM Standard, "Vol E112-96el Standard Test Methods for Determining Average Grain Size", ASTM International, 2003.

3 Borisenko Victor E., Hesketh P. J., "Rapid thermal Processing of Semiconductors", Plenum Press, New York and London, 1997, chapter 3.

4 Cao Min, Voorde Paul Vande, Cox Mike and Greene Wayne, "Boron Diffusion and Penetration in Ultrathin Oxide with Poly-Si Gate", IEEE Electron Device Letters, Vol. 19, No. 8, August 1998, pp 291-293.

5 Carroll M.S., Chang C-L., Sturn J.C. and Buyuklimanli, "Complete Suppression of Boron Transient Enhanced Diffusion and Oxidation-enhanced Diffusion in Silicon Using Localized Substitutional Carbon Incorporation", Appl. Phys. Lett. 1998.

6 Chao T.S., Chien C.H., Hao C.P., Liaw M.C., Chu C.H., Chang C.Y., Lei T.F., Sun W.T. and Hsu C.H., "Suppression of Boron Penetration in P+-Poly-Si Gate Metal-Oxide-Semiconductor Transistor Using Nitrogen Implantation", Jpn. J. Appl. Phys. Vol. 36, 1997 pp 1364-1367.

7 Choi Changhoon, "Phd Thesis: Modeling of Nanoscale MOSFETs", Stanford University, 2002.

8 Choi Chang-Hoon, Chidambaram P.R., Khamankar Rajesh, Machala Charles F., Yu Zhiping and Dutton Robert W., "Dopant Profile and Gate Geometric

Effects on Polysilicon Gate Depletion in Scaled MOS", IEEE Transaction On Electron Devices, Vol. 49, No. 7, July 2002, pp 1227-1231.

9 Choi Chang-Hoon, Chidambaram P.R., Khamankar Rajesh, Machala Charles F., Yu Zhiping and Dutton Robert W., "Gate Length Dependent Polysilicon Depletion Effects", IEEE Electron Device Letters, Vol. 23, No. 4, April 2002, pp 224-226.

10 Chong Y.F., Pey K.L., Lu Y.F., Wee A.T.S., Osipowicz T., Seng H.L., See A. and Dai J.-Y., "Liquid-phase Epitaxial Growth of Amorphous Silicon During Laser Annealing of Ultrashallow p+/n Junctions", Applied Physics Letters, Volume 77, Number 19, November 2000, pp 2994-2996.

11 Chou A.I., Lin C., Kumar K., Chowdhury P., Gardner M., Gilmer M., Fulford J. and Lee J. C., "The Effect of Nitrogen Implant into Gate Electrode on the Characteristics of Dual-gate MOSFETs with Ultra-thin Oxide and Oxynitrides", IEEE, 1997, Page 174-177.

12 Fahey P.M., Griffin P.B., and Plummer J.D.. "Point Defects and Dopant Diffusion in Silicon". Review of Modern Physics, Vol. 61, No. 2, 1989, pp. 289-384.

13 Fang H., Krisch K. S., Gross B.J., Sodini Charles G., Chung J. and Antoniadis, "Low-Temperature Furnace-Grown Reoxidized Nitrided Oxide gate Dielectrics as a barrier to Boron Penetration", IEEE Electron Device Letters, Vol. 13, No. 4, April 1992, pp 217-219.

14 Fiory A.T. and Bourdelle K. K., "Activation of Implanted Poly Gates by Short Cycle Time Annealing", Mat. Res. Soc. Symp, Vol. 610, 2000, pp. B3.3.1-B3.3.6.

15 Fiory A.T., Bourdelle K.K., Roy P.K. and McCoy S.P., "Spike Annealing of Implanted PMOS gates", Proceedings of RTP 2000 Conference, 2000.

16 Han L.K., Wristers D., Yan J., Bhat M and Kwong D.L., "Highly Suppressed Boron Penetration in NO-Nitrided SiO2 for P+-Polysilicon Gated MOS Device Applications", IEEE Electron Device Letters, Vol. 16, No. 7, July 1995, pp 319-321.

17 ITRS, Front End Processes Report, http://public.itrs.net/files/1999_SIA_ Roadmap/FEP/BRM99%20Sept30.doc, ITRS, 1999.

18 Jacques J.M., Robertson L.S., and Jones K.S., Bennett Joe and Rendon Mike, "Effect of Fluorine on the Diffusion of Boron in Amorphous Silicon", Mat. Res. Soc. Symp. Proc. Vol. 717, 2002 pp C4.6.1-C4.6.6.

19 Joshi A.B., Ahn J. and Kwong D.L., "Oxynitride Gate Dielectrics for P+-Polysilicon Gate MOS Devices", IEEE Electron Devices Letters, Vol. 14, No. 12, December 1993, pp 560-562.

20 Josse E., Arnaud F., Wacquant F., Lenoble D., Menut O. and Robilliart E., "Spike Anneal Optimisation for Digital and Analogue High Performance 0.13mm CMOS platform", Proceedings ESSDERC 2002, 2002, Page 207-210.

21 Kamins T, "Polycrystalline Silicon for Integrated Circuits and Displays", Kluwer Academic Publishers, 1998.

22 Kanert W. , "SIMS Measurements of Polysilicon Out diffusion Experiments using Arsenic". SIEMENS, München, 1994.

23 Kang S.-K, Kim J.J., Kim B.G., Ko D.-H., Kang H.B. and Yang C.W., "N-Type Dopant Activation Behaviours in Poly Si1-xGex Films with Ge Contents and Activation Temperatures", J. ECS, 150 (3) 2003, G173-G176.

24 Kodate J. , Miyake M. , and Konake S. , "Diffusion for Uniforming Ion Implanted As Profile in Polysilicon". Fall Meeting, The Japanese Society of Applied Physics, Vol. 18a-ZT-1, 1992, pp. 718-719.

25 Krisch K.S., Green M.L., Baumann F.H., Brasen D., Feldman L.C. and Manchanda L., "Thickness Dependence of Boron Penetration Through O2- and N2O-Grown Gate Oxides and Its Impact on Threshold Voltage Variation", IEEE Transactions On Electron Devices, Vol. 43, No. 6, June 1996, pp 982-990.

26 Ku Ja-Hum, Choi C.-J., S. Song, Choi S., Fujihara K., Kang H.-K., Lee S.-I., Choi H.-G. and Ko D.-H., "High Performance pMOSFETs with Ni(SixGe1-x)/Poly-Si0.8Ge0.2 Gate", 2000 Symp. On VLSI Tech. Digest of Technical Papers, 2000, pp 114-115.

27 Kubicek S., Henson W.K., De Keersgieter A., Badenes G., Jansen P., Meer H. van, Kerr D., Naem A., Deferm L. and De Meyer K., "Investigation of Intrinsic Transistor Performance of Advanced CMOS Devices with 2.5nm NO Gate Oxides", IEDM 1999, 1999, pp 34.3.1-34.3.4.

28 Kuroi Takashi, Kobayashi Maiko, Shirahata Masayoshi, Okumura Yoshiki, Kusunoki Shigeru, Inuishi Masahide and Tsubouchi Natsuro, "The Impact of nitrogen Implantation into Highly Doped Polysilicon Gates for Highly Reliable and High-Performance Sub-Quarter-Micron Dual-Gate Complementary Metal Oxide Semiconductor", Jpn. J. Appl. Phys. Vol. 34, 1995, pp. 771-775.

29 Lee Lurng Shehng and Lee Chung Len, "Argon ion-Implantation on Polysilicon or Amorphous-Silicon for Boron Penetration Suppression in P+ PMOSFET", IEEE Transactions on Electron Devices, Vol. 45, No 8, August 1998, pp. 1737-1744.

30 Lee Wen-Chin, King Ya-Chin, King Tsu-Jae and Hu Chenming, "Investigation of Poly-Si1-xGex for Dual-gate CMOS Technology", IEEE Electron Device Letters, Vol. 19, No. 7, July 1998, pp 247-249.

31 Lin Chih-Yung, Chang Chun-Yen and Hsu Charles Ching-Hsing, "Suppression of Boron Penetration in BF2-Implanted P-type Gate MOSFET by trapping of Fluorines in Amorphous Gate", IEEE Transaction on Electron Devices, Vol. 42, No. 8, Aug 1995, pp 1503-1509.

32 Liu C.T., Ma Y., Cheung C.P., Fritzinger L., Becerro J., Luftman H., Vaidya H.M., Colonell J.J., Kamgar A., Minor J.F., Murray R.G., Lai W.Y.C., Pai C.S. and Hillenius S.J., "25Å Gate Oxide without Boron Penetration for 0.25 and 0.3-μm PMOSFETs", 1996 Symposium VLSI Technology Digest of Technical Papers, 1996, pp 2.4-2.2.

33 Ma Z.J., Chen J.C., Liu Z.H., Krick J.T., Cheng Y.C., Hu C. and Ko P.K., "Suppression of Boron Penetration in P+Polysilicon gate P-MOSFET's Using Low-Temperature Gate-Oxide N2O Anneal", IEEE Electron Device Letters, Vol. 15, No. 3, March 1994, pp 109-111.

34 Muller R. and Kamins T., "Device Electronics for Integrated Circuits", Wiley and Sons, Inc, 1986

35 Nam In-Ho, Sim Jae Sung, Hong Sung In, Park Byung-Gook, Lee Jong Duk, Lee Seung-Woo, Kang Man-Sug, Kim Young-Wug, Suh Kwang-Pyuk and Lee Won Seong, "Ultrathin Gate Oxide Grown on Nitrogen-Implanted Silicon for Deep Submicron CMOS Transistors", IEEE Transaction On Electron Devices, Vol. 48, No. 10, October 2001, pp 2310-2316.

36 Ponomarev Youri V., Stolk Peter A., Salm Cora, Schmitz Jurriaan and Woerlee Pierre H., "High-performance Deep Submicron CMOS Technologies with Polycrystalline-SiGe Gates", IEEE Transaction on Electron Devices, Vol. 47, No. 4, Apr 2000, pp 848-855.

37 Puchner H., "Phd Thesis: Advanced Process Modeling for VLSI Technology", Technical University Vienna, 1996.

38 Rhee H.S., Bae G.J., Choe T.H., Kim S.S., Song S., Lee N.I., Fujihara K., Kang H.K. and Moon J.T., "Ge-Redistributed Poly-Si/SiGe Stack Gate (GRPSG) for High-Performance CMOSFETs", 2001 Symposium on VLSI Technology Digest of Technical Papers, 2001, pp 6A-1-6A-2.

39 Schaber H. , Criegern R. V., and Weitzel I. , "Analysis of Polycrystalline Silicon Diffusion Sources by Secondary Ion Mass Spectrometry". J.Appl.Phys., Vol. 58, No. 11, 1985, pp. 4036-4042.

40 Silterra, “Silterra Training Materials”, 2000.

41 Spinella C. , Cacciato A. , Benyaich F. , Pannitteri S. , and Rimini E. , "Early Stages of the epitaxial Realignment of Poly-silicon Films onto Silicon Substrates: Integration and Size Distribution". In Proceedings: 1st International Rapid Thermal Processing Conference, 1993, pp. 141-155.

42 Spoutai Serguei, Chun Hui-Gon and Ahn Kang-Ho, "New Method for Estimation of Grain Boundaries Contribution to Resistance of Highly Doped Polysilicon," Japan J. Appl. Phys. December 1998, Pp. 6974-6976.

43 Stewart E. J., Carroll M.S. and Sturn James C., "Suppression of Boron Penetration in P-Channel MOSFETs Using Polycrystalline Si1-x-yGexCy Gate Layers", IEEE Electron Device Letters, Vol. 22, No. 12, December 2001, pp 574-576.

44 Sun W.T., Chen S.H., Lin C.J., Chao T.S. and Hsu C.C.-H., "Process optimization for Preventing Boro-Penetration Using P or As Co-Implant in P-Poly Gate of P-MOSFETs", VLSI, 1995, pp 40-42.

45 Sze S.M., "Semiconductor Devices, Physics and Technology", John Wiley & Sons, 1985.

46 Taylor W., Göselle U. , and Tan T.Y., "Present Understanding of Point Defect Parameters and Diffusion in Silicon: An Overview, In Proceedings: Process Physics and Modeling in Semiconductor Technology", The Electrochemical Society, 1993, pp. 3-19.

47 Williams J.D., "Epitaxial Alignment of Polycrystalline Silicon and its Implications for Analogue Bipolar Circuits. PhD thesis", University of Southampton, United Kingdom, 1992.

48 Xiang Qi, Joong Jeon, Sachdey Pinkish, Yu Bin, Saraswat Krishna C. and Lim Ming-Ren, "Very High Performance 40nm CMOS with Ultra-thin Nitrided/Oxynitride Stack Gate Dielectric and Pre-doped Dual Poly-Si Gate Electrodes", Proceedings of IEDM 2000, 2000, pp 10.8.1-10.8.3.

49 Yu Bin, Ju Dong-Hyuk, Kepler Nick and Hu Chenming, "Impact of Nitrogen (N14) Implantation into Polysilicon Gate on High-Performance Dual-Gate CMOS Transistors", IEEE Electron Device Letters, Vol. 18, No. 7, July 1997, pp 312-314.

50 Yu Bin, Ju Dong-Hyuk, lee Wen-Chin, Kepler Nick, King Tsu-Jae and Hu Chenming, "Gate Engineering for Deep-Submicron CMOS Transistors", IEEE Transaction on Electron Devices, Vol. 45, No. 6, June 1998, pp 1253-1262.

51 Yu Bin, Wan Yun, Wang Haihong, Xiang Qi, Riccobene Concetta, Talwar Somit, Lin Min-Ren, "70nm MOSFET with Ultra-Shallow, Abrupt, and Super-Doped S/D Extension Implemented by Laser Thermal Process (LTP)", IEDM 1999, 1999, pp 20.4.1-20.4.4.

PUBLICATIONS

1. Soon-Aik Chew, Norani Muti Mohamed, Albert Victor Kordesch, "Optimization of Spike Annealing on Boron Activation and Diffusion in P-Poly-Silicon Gate," The 3rd International Conference on Advances in Strategic Technologies (ICAST 2003), Kuala Lumpur, Malaysia, 12-14 August, 2003.
2. Soon-Aik Chew, Norani Muti Mohamed, Albert Victor Kordesch, "Effect of Spike Annealing on Poly Depletion and Boron Penetration in PMOSFETs," 2003 IEEE National Symposium on Microelectronics (NSM 2003), Perlis, Malaysia, 9-10 September, 2003.

APPENDIX A

Periodic Table of the Elements

Group**

Period	1 IA 1A	2 IIA 2A	3 IIIB 3B	4 IVB 4B	5 VB 5B	6 VIB 6B	7 VIIB 7B	8 -------	9 VIII --- 8	10 ----	11 IB 1B	12 IIB 2B	13 IIIA 3A	14 IVA 4A	15 VA 5A	16 VIA 6A	17 VIIA 7A	18 vIIIA 8A
1	1 H 1.008																	2 He 4.003
2	3 Li 6.941	4 Be 9.012											5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00	9 F 19.00	10 Ne 20.18
3	11 Na 22.99	12 Mg 24.31											13 Al 26.98	14 Si 28.09	15 P 30.97	16 S 32.07	17 Cl 35.45	18 Ar 39.95
4	19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.88	23 V 50.94	24 Cr 52.00	25 Mn 54.94	26 Fe 55.85	27 Co 58.47	28 Ni 58.69	29 Cu 63.55	30 Zn 65.39	31 Ga 69.72	32 Ge 72.59	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
5	37 Rb 85.47	38 Sr 87.62	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.94	43 Tc (98)	44 Ru 101.1	45 Rh 102.9	46 Pd 106.4	47 Ag 107.9	48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8	52 Te 127.6	53 I 126.9	54 Xe 131.3
6	55 Cs 132.9	56 Ba 137.3	57 La* 138.9	72 Hf 178.5	73 Ta 180.9	74 W 183.9	75 Re 186.2	76 Os 190.2	77 Ir 190.2	78 Pt 195.1	79 Au 197.0	80 Hg 200.5	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0	84 Po (210)	85 At (210)	86 Rn (222)
7	87 Fr (223)	88 Ra (226)	89 Ac~ (227)	104 Rf (257)	105 Db (260)	106 Sg (263)	107 Bh (262)	108 Hs (265)	109 Mt (266)	110 --- ()	111 --- ()	112 --- ()		114 --- ()		116 --- ()		118 --- ()

Lanthanide Series*	58 Ce 140.1	59 Pr 140.9	60 Nd 144.2	61 Pm (147)	62 Sm 150.4	63 Eu 152.0	64 Gd 157.3	65 Tb 158.9	66 Dy 162.5	67 Ho 164.9	68 Er 167.3	69 Tm 168.9	70 Yb 173.0	71 Lu 175.0
Actinide Series~	90 Th 232.0	91 Pa (231)	92 U (238)	93 Np (237)	94 Pu (242)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (249)	99 Es (254)	100 Fm (253)	101 Md (256)	102 No (254)	103 Lr (257)

APPENDIX B

RCA process recipe

Steps	Chemical	Process Time
1	EDR	0s (Transfer Bath)
2	SC1	600s
3	QDR	480s
4	SC2	600s
5	QDR	480s
6	EDR	0s (Transfer Bath)
7	FR	120s
8	LPD	527s

Chemical Abbreviation:

EDR	Etch Drain Rinse. Cold de-ionized (DI) water rinse. An EDR is in series after each HF process bath.
SC1	1:2:10(by volume, $NH_4OH:H_2O_2:H_2O$) at temperature 50.0±5.0°C
SC2	1:1:20(by volume, $HCl:H_2O_2:H_2O$) at temperature 50.0±5.0°C
QDR	Quick Drain Rinse. Hot/Cold DI water rinse bath. A QDR is in series after each process bath except HF.
FR	Final rinsing bath
LPD	Heating process at 130.0±5.0°C