Use Authorization

In presenting this dissertation in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my dissertation for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this dissertation for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

Semiconductor Process and Yield Improvements

By

Ikhoon Shin

A Dissertation

submitted in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy in Engineering and Applied Science

in the Department of Electrical Engineering

Idaho State University

Spring 2018

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Ikhoon Shin find it satisfactory and recommended that it be accepted.

_____

Dr. George Imel
Major Advisor

_____

Dr. Steve Chiu
Committee Member

_____

Dr. R. Eugene Stuffle
Committee Member

_____

Dr. Anish Sebastian
Committee Member

_____

Dr. Joshua Pak
Graduate Faculty Representative

DEDICATION

I dedicate my dissertation work to my family and many of my friends and colleagues at ON Semiconductor who have supported me throughout the process. A special gratitude to my loving parents, MyungGoo Shin and InSok Jun and my wife Jinee Park and my son Phillip Shin who had to go to late night coffee shop with me while I was preparing for qualification exam. I dedicate this work and give special thanks to my friend and best supporters Todd Corsetti and Paul Streva for being there for me throughout the entire doctorate program.

ACKNOWLEDGEMENTS

I wish to thank my committee members Dr. Steve Chiu, Dr. R. Eugene Stuffle, Dr. Anish Sebastian, and Dr. Joshua Pak who were more than generous with their time, expertise, understanding and support which made it possible for me to work full time at ON Semiconductor and still able to pursue Doctorate program. I also would like to thank Glenn Florence and Scott Donaldson for proofreading and providing valuable feedback.

I would like to express the deepest appreciation to my Major Advisor, Professor Dr. George Imel. He has been supportive since the first class I took from him and the completion of the undertaking could not have been possible without him accepting me as his student.

Table of Contents

List of Figures

List of Tables

Semiconductor Process and Yield Improvement

Dissertation Abstract – Idaho Stature University (2018)

Since the invention of the electrical transistor in 1947, nowadays it is unimaginable to think that modern society can function without electronic devices. The semiconductor industries enabled the increased use of new electronics in unlikely applications such as automotive. By understanding critical methods and its applications of industry standard semiconductor processes and failure analysis on various stages of the processes, one can use that knowledge to minimize process variation and improve production efficiency leading to final semiconductor product yield increases. Minimization of future failures of new semiconductor devices in the field is also achieved.

The polysilicon hole defects arise from a three-way interaction between the boron implant in polysilicon, subsequent thermal processing, and the BOE chemistry. During the annealing and activation thermal steps after implantation, boron with a high solubility solid limit will be strongly affected by the morphological lattice of polysilicon and will cause the grains to grow in size during redistribution, which rearranges the grain boundaries. Therefore, grain boundaries and lattices of polysilicon are physically impacted by the boron implant and become more susceptible to micro-roughing due to NH_4OH and possible metal contamination in BOE chemistry when adding ammonium fluoride.

Key Words: Defects, Polysilicon Holes, Semiconductor Process, Yield Improvement.

1. What is the semiconductor process?

The semiconductor process utilizes the silicon base starting material in the form of a wafer. This can be manipulated by various complex chemical, physical and thermal process techniques with special equipment to create a desired N type (extra electrons) or P type (extra holes) material. Finally interconnections and dielectric isolations form an electrically functional device. The most common device in the semiconductor process is called the Complementary Metal Oxide Silicon (CMOS) transistor. It is called complementary since N-type and P-type transistors are created to form a logic circuit. The transistor consists of a drain, source, gate and body. The body of the transistor is formed by lightly doped N-type well or P-type well where the entire N or P-type of the transistor is enclosed. For example, in the N-type well, heavily doped P-type drains and sources are formed while in the P-type well, heavily doped N-type drains and sources are formed.

The silicon atom contains four valence electrons at the outer energy band and it shares four other electrons from its nearest neighboring atoms. Its electrical characteristics can be changed by adding or removing electrons by way of implants and impurity diffusion processes. With controlled diffusion using the temperature profile, the conduction property of silicon can be precisely controlled. To form a certain device, the exact region of N-type or P-type area (active area) needs to be defined by way of photolithography techniques. Photolithography is defined as the transfer of a pattern from a photomask layer (reticle) onto a light-sensitive chemical film (photoresist). The light used to transfer the pattern is commonly from an ultraviolet (UV) light source. The reticle is a fused silica base glass that is patterned with chrome metal film to define the area to be transparent or opaque for the UV light source.

Protective layers on the wafer such as a specific photoresist material, oxide or nitride layer can be used for patterning. If oxide or nitride layers are used as the protective layer it is commonly called “hard mask”. For example, a protective layer is patterned by placing a reticle over the wafer and exposing it to a UV light source. Only photoresist areas that are exposed to UV will be removed and this exposed area will be implanted or etched and the rest of the area with photoresist material will be protected. A nitride layer can be deposited by Low Pressure Chemical Vapor Deposition (LPCVD) or Plasma Enhanced Chemical Vapor Deposition (PECVD). An oxide layer can be grown in a high temperature (typically ~1200°C) furnace with supplied oxygen atoms. Nitride and oxide layers are also commonly used as insulating layers in between conducting stacks or used as dielectrics for capacitors. Various etch processes are employed to selectively remove layers that are not needed. Chemical etch or plasma base sputter processes are used to create trenches, define device patterns or used for surface cleaning. Cleaning process steps in between photo, etch, implant and depositions are critical to minimize extrinsic defects associated with contamination or residual particles. As part of the cleaning process, a hydrofluoric acid (HF) chemical solution is used to remove the native oxide layer which is grown naturally in the ambient environment. A metallization process is used to interconnect devices. In general aluminum copper (AlCu) is use for interconnections greater than 0.18 µm geometry. For processing nodes below 0.11 µm, Cu metallization is used to reduce parasitic resistance in metal interconnects. In general Cu metal is electroplated onto the wafer and since Cu cannot be etched with commonly used chemicals or the plasma etch process, additive patterning called the damascene process is used. Detailed information is provided in the Metallization section.

Figure 1.1 shows general process flows of typical semiconductor processes of deposition, patterning, material modification, removal, and planarization steps to form a CMOS transistor.

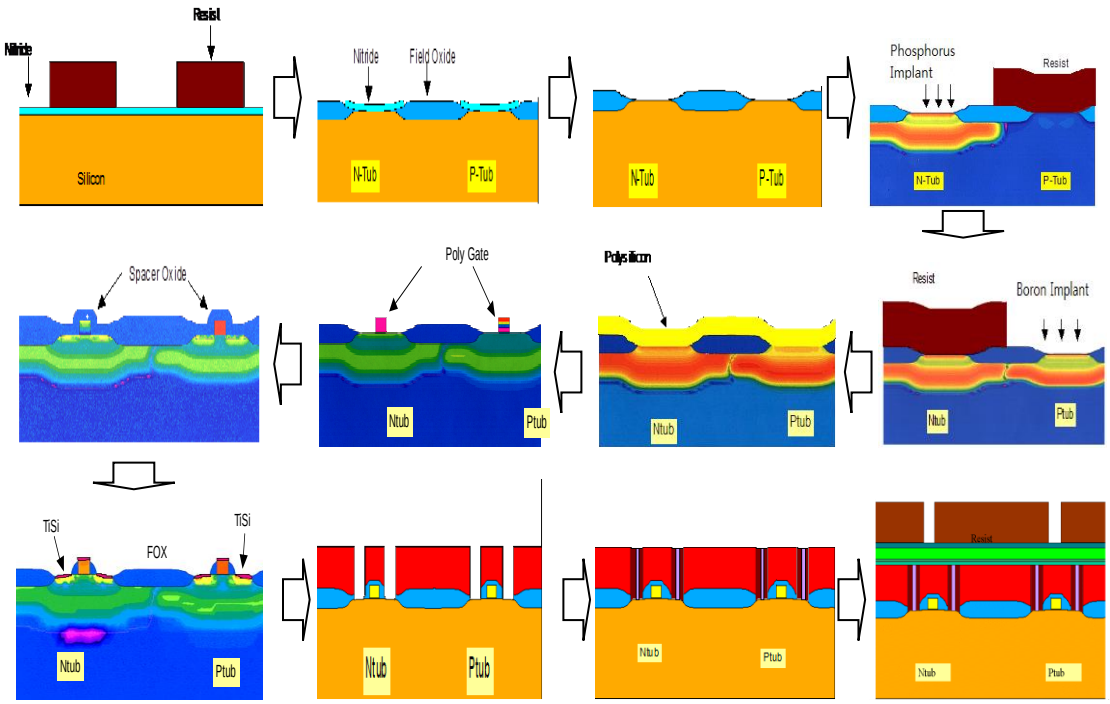


Figure 1.1 Overview of typical semiconductor process flow of transisitor build up. Critical steps show simulated diffusion and carrier concentration levels of of N and P type areas. (Credit: ON Semiconductor internal document archives)

1.1 Deposition of different materials

A multitude of layers of different materials have to be deposited or grown during the semiconductor fabrication process. As mentioned above, oxides and nitrides are used as non-conductive dielectric layers between metals and deposited on specific areas of the device. Also an oxide layer is grown on silicon to form SiO_2 as an isolation structure to separate active areas. With adequate isolation, leakage currents between active areas are minimized and thus ensure the minimum distance allowed (design rule) for specific technology requirements. To prevent oxide growth on an active area, nitride is deposited as a blocking layer. Figure 1.2 shows Local Oxidation of Silicon (LOCOS) isolation cross-section view with active areas.

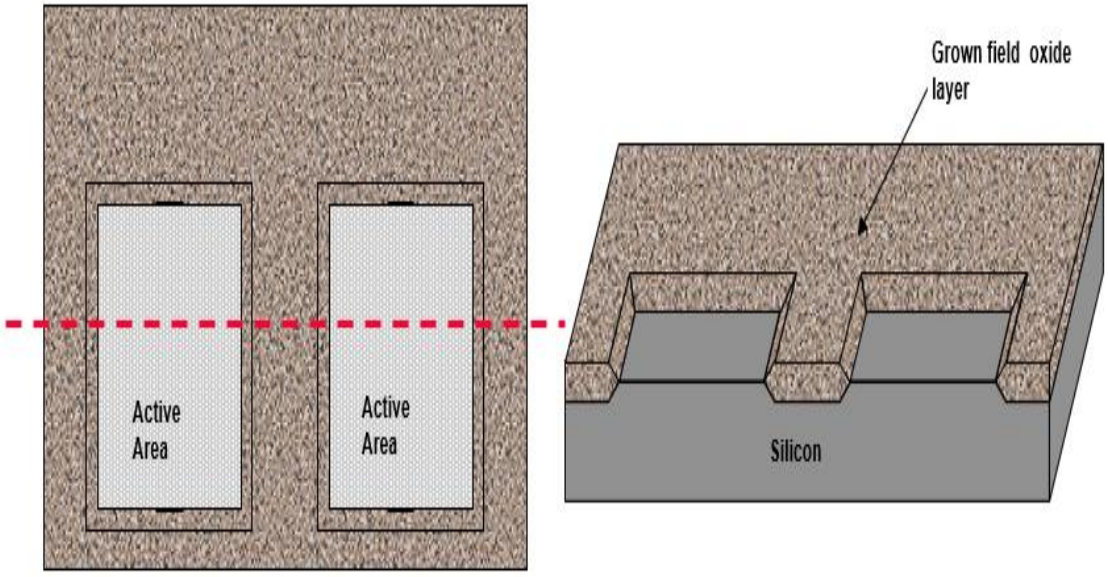


Figure 1.2 Cross-section view of Local Oxidation of Silicon (LOCOS) isolation with active areas. (Credit: ON Semiconductor internal document archives)

A thin layer of grown SiO_2 (around 20 nm-100 nm) thick is used as the gate oxide of CMOS transistors, where the SiO_2 layer is between the polysilicon gate and underlying conductive channel of the device. Externally applied gate voltage will be modulated based on the thickness

and the quality of the gate oxide. Electric fields from the gate voltage will be reduced with thicker gate oxide on channel areas below. To increase the electric field, one can reduce the gate oxide thickness but this may increase the leakage current if the oxide is of poor quality. At higher electric fields, *Fowler-Nordheim tunneling* occurs where electrons tunnel into the oxide layer, thus increasing the leakage current causing unwanted power dissipation of the device, reliability and dielectric integrity issues. At an oxide field of 8 MV/cm, the measured Fowler-Nordheim tunneling current density is about $5x10^{-7}$ A/cm^2. **[1]** With extremely thin gate oxide thickness (below 5 nm), direct tunneling can occur even with higher quality oxide.

Figure 1.3 shows a typical CMOS gate oxide location.

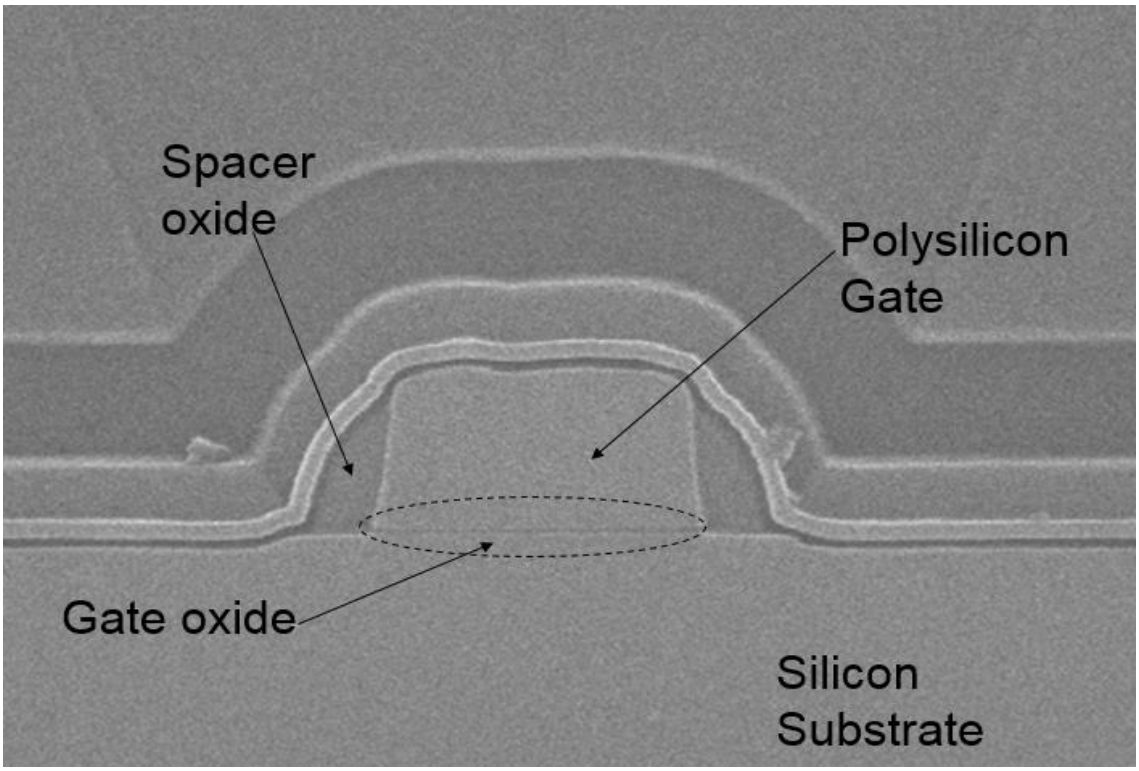


Figure 1.3 General cross-section view of CMOS showing 70Å oxide (inside dotted circle) in between polysilicon gate and silicon substrate. (Credit: ON Semiconductor internal document archives)

1.1.1 Oxidation

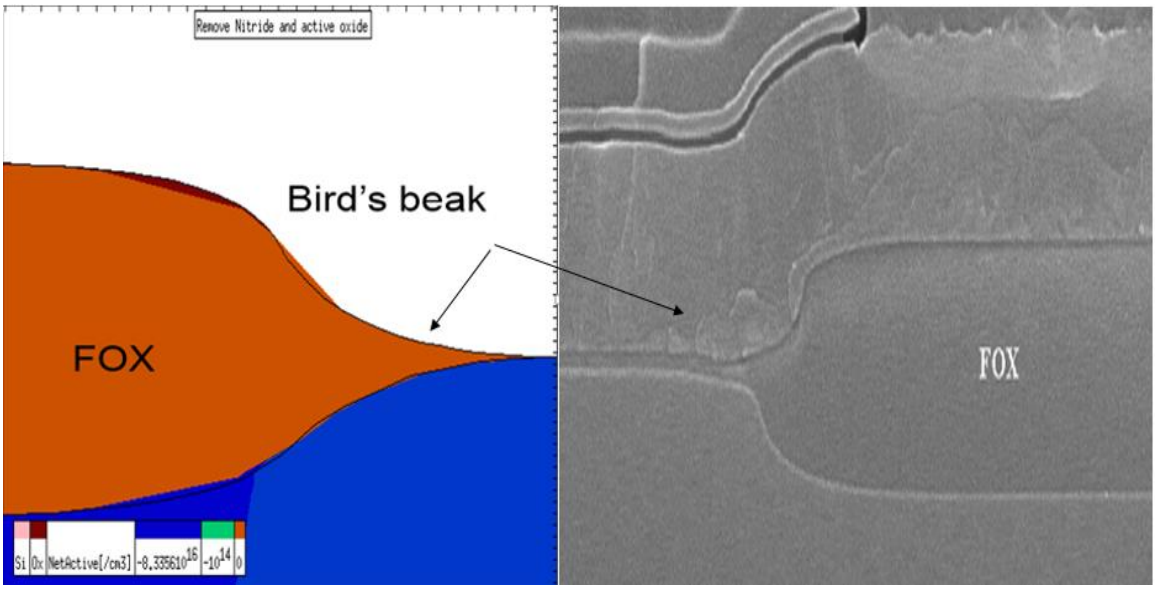
The oxidation of the silicon surface occurs from the transport of oxygen atoms through the growing oxide and forms an oxide–silicon interface at elevated temperatures. **[2]** There are two different types of oxidation in the semiconductor process. One is using H_2O as input (wet oxidation) with oxidation temperature around 900 °C - 1000 °C and the other is using O_2 gas (dry oxidation) at oxidation temperatures around 1200 °C. Due to the high temperature of the process, oxidation steps can be used for diffusion of implanted species; details will be discussed in the diffusion section. The wet oxidation can achieve faster oxide growth but may have large process variation of thickness. Dry oxidation results in slower oxide growth but can achieve better thickness control into the angstrom range and achieve higher density oxide, thus providing a higher quality oxide layer. In general, wet oxidation is used for isolation structures such as LOCOS field oxide and dry oxidation is used for critical layers such as a dielectric layer between a poly-crystal gate and a silicon channel area of a transistor device. Figure 1.4 shows the typical LOCOS field oxide formation.

Dry oxidation reaction on silicon:

$$Si + O_2 = SiO_2$$

Wet oxidation reaction on silicon:

$$Si + 2H_2O = SiO_2 + 2H_2$$



(a) (b)

Figure 1.4 Technology Computer Aided Design (TCAD) simulation result (the –x and y axis are in µm) (a) and actual cross-section view (b) of LOCOS Field Oxide (FOX) growth - forms bird's beak after removal of nitride blocking layer. (Credit: ON Semiconductor internal document archives)

The oxidation process consumes around 40% of the silicon since the SiO_2 interface has approximately twice the volume of silicon and this SiO_2 layer grows in both directions almost equally. **[3]** For example, for 1000Å of oxide growth, 400Å of silicon is consumed. Since oxide will grow in both directions even with an oxide blocking nitride layer, some limited amount of oxide will grow laterally at the edge of the nitride layer. After oxidation, the nitride layer is removed and the thin layer of oxide region is referred to as a bird's beak. It is an important factor to consider for overall dimension and height control of the device because overall field oxide height is only 60% of total growth. The general relationship for the oxidation of silicon surface is given in **[4]**

$$X_0^2 + AX_0 = B\,(t + T)$$

where

$$A = 2D\left(\frac{1}{k} + \frac{1}{h}\right)$$

$$B = \frac{2DC}{Ni}$$

where

t = Time (s)

T = Correction factor for initial native oxide layer that is generally present on silicon surface in the range of 10 - 50 angstrom (10^{-10} m)

B = Parabolic rate constant (cm^2/s)

C = Maximum solubility of the oxidizing species in the oxide (gram/100 gram)

Ni = Number of oxidant molecules in the oxide (molecules in gram/unit volume)

D = Diffusion coefficient of the oxidizing species in the oxide (cm^2/s)

k = Chemical reaction rate constant for oxidation at the silicon surface (cm/s)

h = Mass transfer coefficient of the oxidizing species from the gas phase to the oxide surface (cm/s)

Quality control of oxide thickness is one of the most critical areas of CMOS transistor devices. For example, the turn-on voltage of a transistor can be changed by different thicknesses of gate oxide due to electric field impact as voltage is applied to the gate node. The oxide quality directly impacts the reliability of the device because poor gate oxide quality may cause oxide breakdown at the applied operating voltage and cause product failures. Temperature zone control and its consistency across the furnace are important to ensure the uniform oxide growth across the wafer. Oxide thickness differences between the center and edge of the

wafer can cause device performance variation and add process variation to subsequent process steps that may cause the wafer to be scrapped. The furnace temperature is kept constant at 650 °C or 800 °C and then temperature is ramped up to the operating temperature of 900 °C - 1200 °C for oxidation on the silicon surface. The wafer boat must be slowly pushed into the furnace to minimize temperature coefficient variations on the wafer that can cause stress to the wafers potentially cracking and /or warps. Warped wafers can cause focus issues at the photo steps and in extreme cases, wafers will need to be scrapped. Furnace loading size can be from 25 to 150 wafers depending on the equipment size configuration. The temperature profile must be carefully chosen based on empirical results of given equipment that minimize the difference in the oxide thickness variation from front to back in the wafer boat. It is essential to avoid using nitrogen gas as the purging agent during the loading because silicon in nitrogen gas above 900 °C will form silicon nitride by direct reaction and this silicon nitride layer will prevent oxide growth on the affected area.

1.1.2 Chemical vapor deposition

Various materials such as oxide, nitride, polysilicon, and metal can be deposited using a chemical process called Chemical Vapor Deposition (CVD). In the CVD process, the silicon wafer is exposed to one or more volatile precursors, which react and/or decompose on the substrate surface to produce the desired deposit of conformal thin films such as oxide layer. The two most important deposition methods of depositing films in semiconductor processes are the Low Pressure Chemical Vapor Deposition (LPCVD) and the Plasma Enhanced Chemical Vapor Deposition (PECVD).

1.1.2.1 LPCVD

The LPCVD reactor operates at low pressure (around 0.1 -1 Torr). In the LPCVD reactor, the deposition rate of thin film depends on gas flow rate, pressure and temperature conditions and with proper control, good film thickness uniformity across the wafer can be achieved. It's especially sensitive to the temperature of the chamber and it is controlled in a resistance heated furnace to reach 300 °C - 1075 °C depending on desired material that needs to be deposited.

Polycrystalline silicon which is used as part of the gate structure of a CMOS transistor is deposited from trichlorosilane ($SiHCl_3$) or silane (SiH_4), using the following reactions. **[5]**

$$SiH_3Cl \rightarrow Si + H_2 + HCl$$

$$SiH_4 \rightarrow Si + 2\ H_2$$

A silicon nitride layer which is used commonly for a dielectric passivation layer can be deposited using dichlorosilance ($SiCl_2H_2$) and ammonia (NH_3) as source gas:

$$3\ SiCl_2H_2 + 4\ NH_3 \rightarrow Si_3N_4 + 6\ HCl + 6\ H_2 \quad (\text{at } 800°C)$$

The byproduct gases and unused gases must be properly exhausted out of the reactor system so that high quality of films can be achieved. A proper plumbing system and regular maintenance will prevent system malfunction and minimize environmental issues.

1.1.2.2 PECVD

Plasma enhanced chemical vapor deposition utilizes radio frequency (RF) plasma to enhance chemical reaction rates of the precursors similar to LPCVD. For example, silicon dioxide can also be deposited from a tetraethoxysilane (TEOS) silicon precursor in an oxygen or oxygen-argon plasma.

$Si(OC_2H_5)_4 + 2H_2O \rightarrow SiO_2 + 4\ C_2H_5OH$: TEOS conversion to SiO_2 with water

$Si(OC_2H_5)_4 \rightarrow SiO_2 + 2\ (C_2H_5)_2O$: TEOS conversion to SiO_2 with heat (>600°C)

With added energy from the plasma, PECVD processing allows deposition at lower temperatures than LPCVD, which is often critical in the manufacture of semiconductors that minimize unwanted thermal budget of the process. Figure 1.5 shows the general cross-section view of a PECVD reaction chamber.

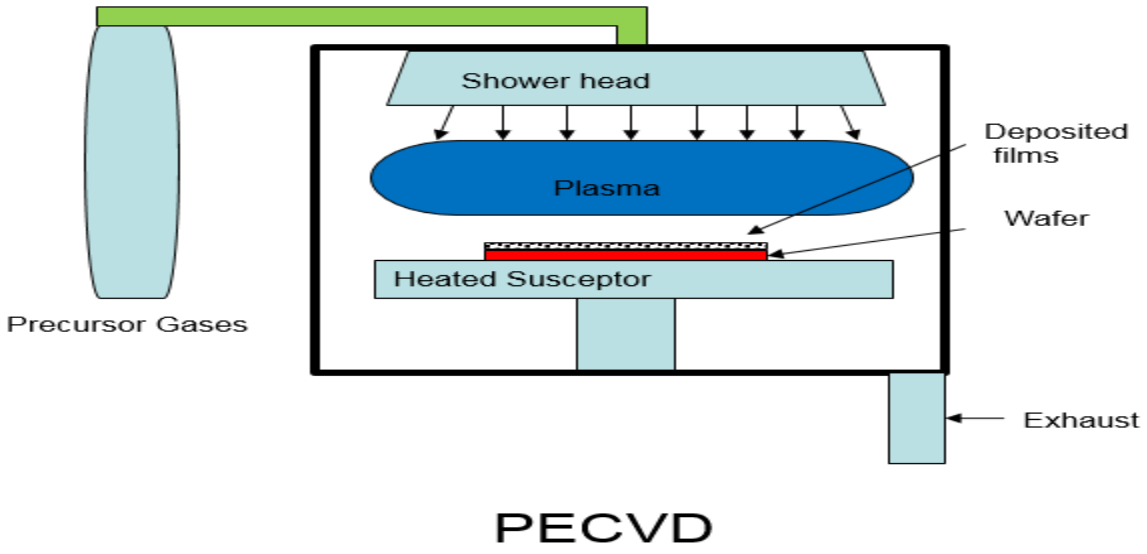


Figure 1.5 General cross-section view of PECVD reaction chamber.

The lower temperatures also allow for the deposition of organic base coating layers, such as a polymer layer.**[6]** Also, wafer surfaces exposed to the plasma receive energetic ion

bombardment. This bombardment can lead to increases in density of the film, help remove contaminants, thus improving the film's electrical and mechanical properties. The higher potential across the sheath surrounding an electrically-isolated object can be achieved by adjustments in reactor geometry and conditions such as gas flow rates, temperature, RF plasma power and gas density.

However, with increased configuration set up requirements, it becomes more difficult to control the process once the Process Of Record (POR) parameters are set. In most manufacturing settings, elaborate Statistical Process Control (SPC) are in place to monitor processes. For example, by sampling oxide thicknesses and extracting statistical data such as mean, median, standard deviation, and range and with known upper and lower thickness limits of oxide, the SPC system can alert operations staff if the thickness of oxide is drifting or out of control. Then, process engineers can intervene and bring the system back into controls thus providing conforming products that meets the specification.

1.1.3 Metallization

Metallization is the process which the components of the Integrated Circuit (IC) are interconnected by metallic conductor layers. This process produces a thin-film metal layer that will serve as the required conductor pattern for the interconnection of the various components on the device. Another use of metallization is to produce metalized areas called bonding pads around the periphery of the chip to enable the bonding of wire leads from the package to the chip. The bonding wires are typically 25 μm diameter gold or copper wires, and the opening of bonding pads are usually made to be around 50x50 μm^2. Typically 1-2 μm thick AlCu (99.5wt%

Al and 0.5wt% Cu) is commonly used on technology nodes greater than 180 nanometer (nm). For technologies below 180 nm node, copper metal line is commonly used since the resistance of Cu metal is much lower than Al metal. With lower metal resistance, the thickness of metal can be reduced for the same resistance output; thus dielectric layer thickness can be increased to reduce the parasitic capacitance while maintaining the overall stack thickness. The resistance/capacitance (RC) time delay can be improved, for example on logic circuit performance using copper metal. **[7]** The resistance and capacitance of the basic structure of conductive material shown in Figure 1.6 and their relationship is described as follows.

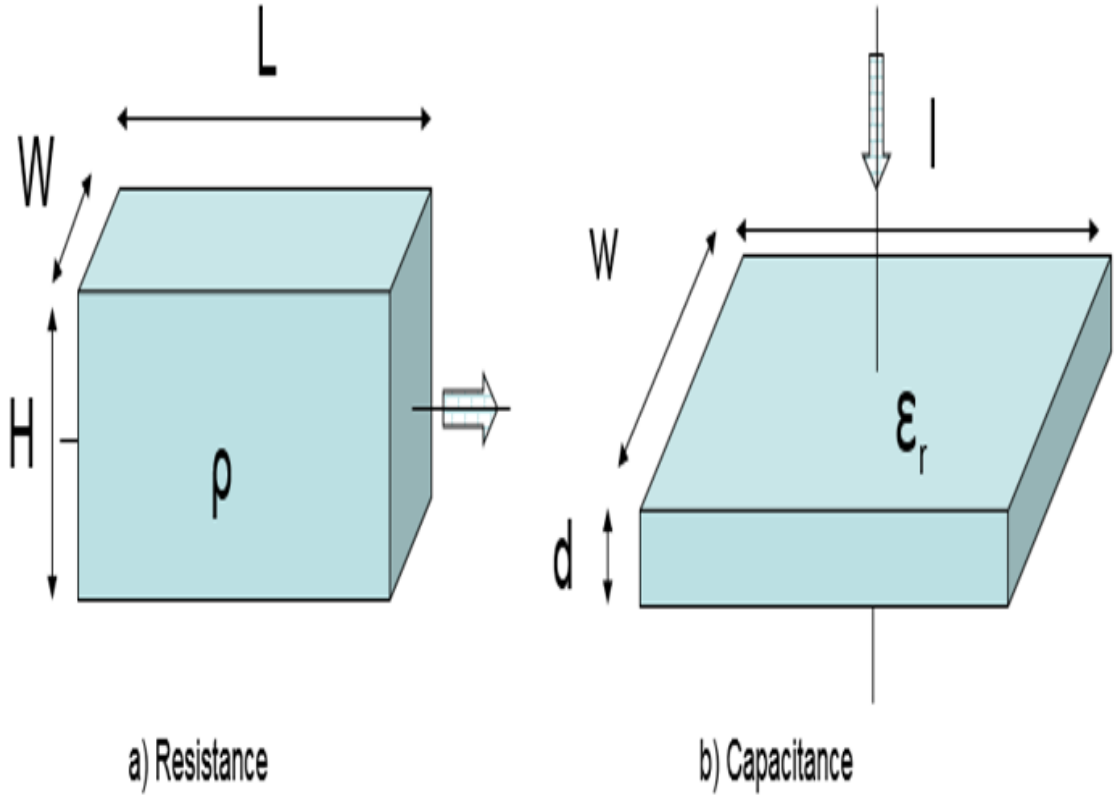


Figure. 1.6 shows basic structures for resistance (a) and capacitance (b)

$$R = \frac{\rho L}{WH} \qquad C = \frac{lw\varepsilon_r\varepsilon_o}{d}$$

where

R = Resistance (Ohm (Ω))

ρ = Resistivity (Ω·m)

L = Length (m)

W = Width (m)

H = Thickness (m)

C = Capacitance (farad (F))

l = Length (m)

w = Width (m)

ε_r = Relative permittivity of the dielectric

ε_o = Permittivity of free space

($\approx 8.854\times10^{-12}$ F·m^{-1})

d = Distance between plates (m)

The main reason that copper metal was not used widely in earlier stages of development was due to cross contamination of copper particles, and the fact that copper cannot be easily etched with commonly used dry plasma etch processes for anisotropic pattern results. Group 11 metals such as copper form few volatile compounds at temperatures below 150°C, which limits the approaches that can be used to perform etching that precludes low temperature dry etch processes. **[8]** A new copper process called damascene was developed by utilizing electroplated copper and additive patterning to form copper interconnect lines. Figure 1.7 shows general overview of metal process of Al and Cu.

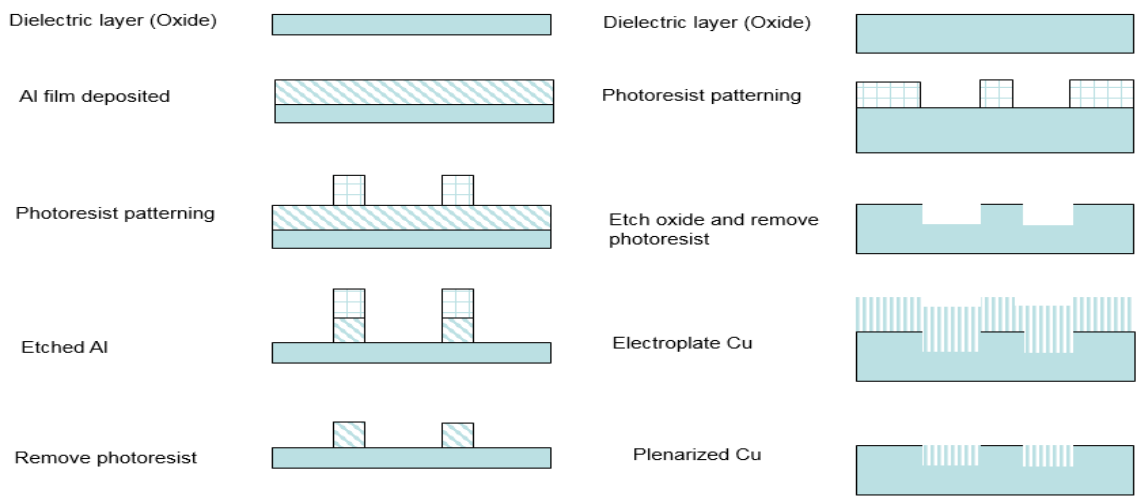


Figure 1.7 General metal process overview. (a) Al process with conventional deposition and etch process and (b) Cu is process using damascene process.

In order for aluminum interconnects to conduct electrical current optimally, a stack of three types of material is applied to create metal interconnects instead of single metal such as AlCu. First a barrier metal such as a titanium nitride (TiN) layer is deposited to prevent direct contact of the conductive metal layer applied to the underlying silicon. If aluminum is directly contacted with silicon, some of the silicon would migrate into the aluminum and aluminum atoms would replace the migrated silicon atoms and this results in electrical shorts. The barrier metal will help reduce the effect of the electromigration problem. Electromigration is generally considered to be the result of momentum transfer from the electrons, which move in the direction of applied electric field to the ions which make up the lattice of the interconnect material. When electrons are conducted through a metal and interact with ions in lattice, electron scattering occurs. Electron scattering increases with imperfections in the lattice and thermal vibration of atoms. Once the electron scattering reaches a certain strength, atoms become separated from the grain boundaries and are transported in the direction of the

current.**[9]** As electron scattering increases it also increases thermal vibration of atom and the effect causes thermal runaway that eventually breaks the metal line.

The second step of the metallization stack is to deposit the body of the metal layer. The most commonly used metal layer for technology larger than 0.18 µm geometry is aluminum (99.5%) copper (0.5%) AlCu. Addition of copper enhances the conductivity of aluminum and reduces the electromigration problem. When depositing aluminum, it does not perfectly follow the topography of the wafers (due to the non-conformal deposition property of aluminum). The barrier metal layer can assist in conducting currents in the contacted area. In general, a 1 µm – 2 µm thick metal layer is deposited for the interconnect.

The third step is to deposit an Anti-Reflective Coating layer (ARC) such as TiN to prevent problems at subsequent photolithography and etch process steps. Even though wafers have been planarized, the surface may still contain some curvature after the metal has been deposited. Without an ARC layer, the UV lights used to expose the photoresist development would be reflected onto parts of the photoresist that are supposed be free of UV light exposure. This phenomenon is called “reflective notching” and unintended removal of photoresist may cause defects in the metal line since subsequent etch steps will remove metal areas that are supposed to be protected by photoresist. Figure 1.8 shows an example of reflective notching on an unflattened metal surface without ARC layer and Figure 1.9 shows a Scanning Electron Microscope (SEM) image of AlCu and associated barrier metal layer.

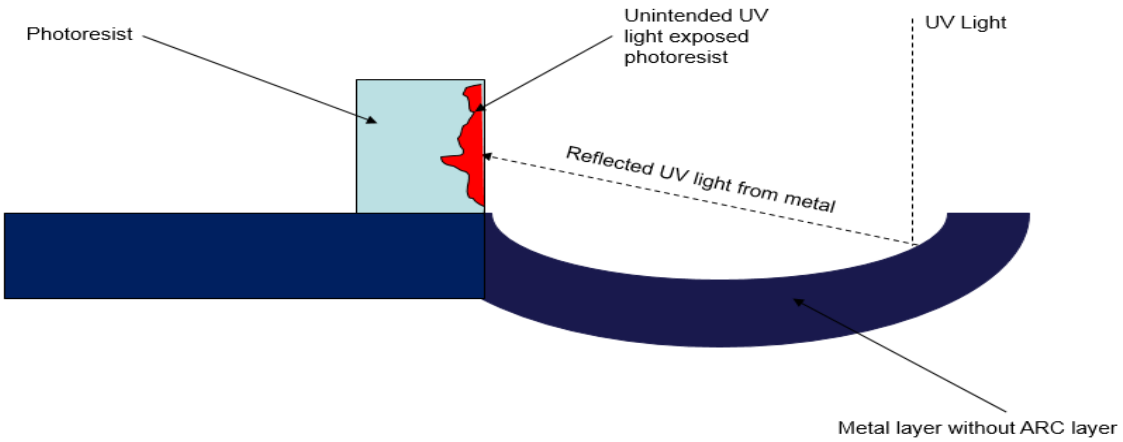


Figure 1.8 Without an Anti-Reflective Coating Layer (ARC), UV light would be reflected onto parts of the photoresist that need to be free of UV light exposure during a photolithography step.

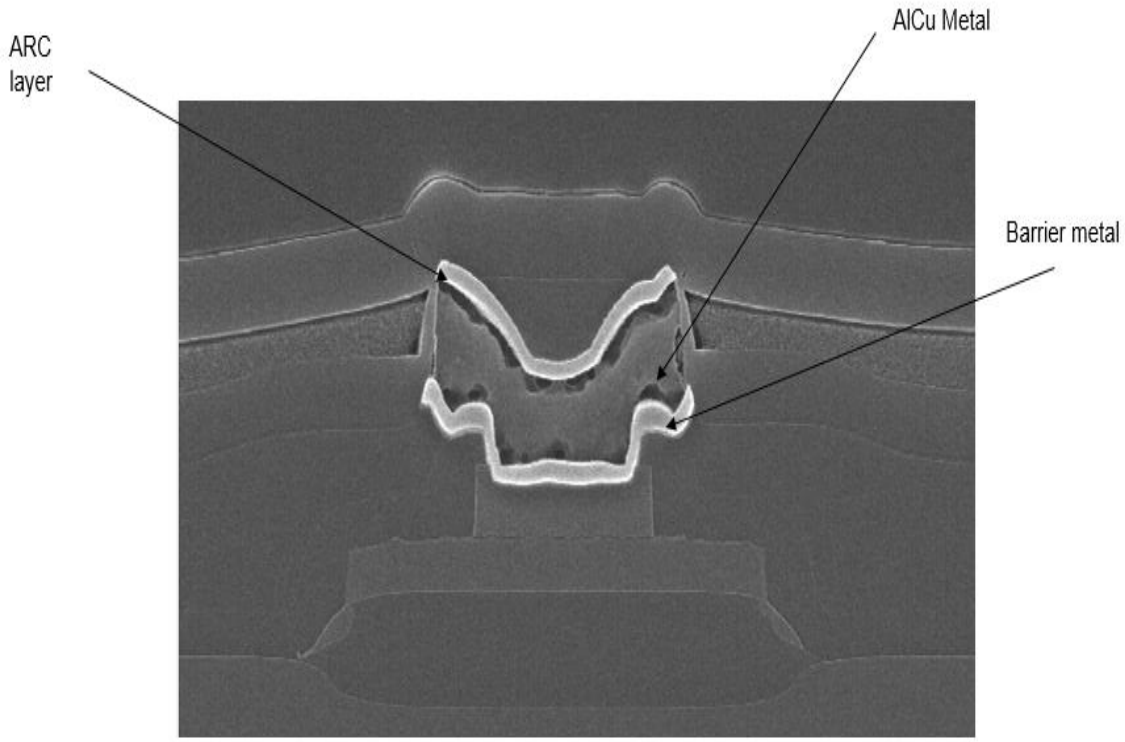


Figure 1.9 SEM image of AlCu metal, TiN barrier metal and ARC metal layer (Picture credit: ON Semiconductor internal document archives)

1.2 Patterning

There are several different layers that require the photo masking step to either expose or cover from subsequent process steps such as implant or etching. The wafer will be coated with photoresist and followed up with a temperature cure process to harden the liquid base photoresist. An ARC layer is used to reduce standing wave interference by minimizing UV light reflection. Reflected UV light means the photoresist only received partial UV light or, UV light may expose an unwanted area of the photoresist and it will cause the pattern to not form properly.

The photomask (reticle) set is used to define specific areas of the device such as the active area, trench and metal lines and typically transparent fused silica blanks covered with a pattern defined with a chrome metal-absorbing film. Each reticle commonly contains only a single layer pattern and to complete an entire CMOS transistor around 25-30 photomask steps are used. A given pattern on the photomask is projected and reduced by several times onto the wafer surface. For example a reticle with 4X size pattern will be shrunk four times on to the wafer. To achieve complete wafer coverage, the wafer is repeatedly "stepped" from position to position under the optical column until full exposure is achieved.

Modern semiconductor technologies utilize sub-micron dimensions, around 350 nm down to 14 nm dimension. The semiconductor device minimum geometry limit is often set by the limit of photo-lithography equipment and its ability to accurately align (overlay) between layers. The alignment between levels is complicated by variations in reticles, temperature caused expansion of layers, wafer flatness and contamination. Wavelength of light vs the dimension of

the device is the critical item that limits the final device geometry. The UV light used to expose photoresist layers is usually from a high-power mercury arc lamp with wafer length of:

0.436 µm (G-line)
0.404 µm (H-line)
0.365 µm (I-line)
0.255 µm and smaller (DUV – deep UV)

Historically, photolithography has used ultraviolet light from gas-discharge lamps using a mercury source. These lamps produce light across a broad spectrum with several strong peaks in the ultraviolet range. However, a mercury light source has its limitation beyond 0.245 µm geometry and new light sources such as ArF (0.193 µm) and F_2 (0.157 µm) are used in excimer laser lithography machines developed in 1982; this is commonly called a scanner photo tool. The term scanner is used since laser beam is scanned across the wafer surface. **[10]**

The minimum line width that can be resolved is limited by the wavelength of the exposure light, the quality of the optics, and the planarity of the wafer. In older contact and proximity printing technology, where the reticle is clamped in contact with the wafer for exposure, the minimum line width Lm, is approximated by:

$$Lm = \sqrt{\lambda g}$$

where

g = Gap between focused image and the wafer surface (µm)
λ = Wavelength of the light (µm)

For a typical example using the G-line photolithography, with wavelength of 0.436 µm, assume a gap of 5 µm; this gives the minimum line width of 1 µm.

The minimum feature size also known as Critical Dimension (CD) for a newer projection system, where the reticle is separated from the wafer thus avoiding any contact, is given approximately by:

$$CD = k_1 \frac{\lambda}{NA}$$

where

k_1 = Coefficient of process factors (dimensionless)

λ = Wavelength of the light (µm)

NA = Numerical aperture of the lens of the photo tool (dimensionless)

where

$NA = n \sin\left(\frac{AA}{2}\right)$

$AA = \text{Angular aperture (degree (°))}$

$n = \text{Lowest refractive index (dimensionless)}$

Using a projection system with I-line photolithography, with a wavelength of 0.365 µm, the minimum CD can be as low as to 0.35 µm.

A typical photolithography process consists of wafers placed on track/cluster equipment, pre-treatment (apply photoresist adhesion promoter), spin photoresist coating, prebake of photoresist, reticle/photomask exposure on stepper or scanners, post-exposure bake, development of photoresist to form a pattern, rinse, and inspection for quality verification. **[11]**

1.2.1 Positive photoresist

With positive photoresist in areas that are exposed to UV light, chemical structure changes and it becomes more soluble in the photoresist developer. These exposed areas are then removed with the photoresist developer solvent. A positive photoresist is composed of a light sensitive ortho-quinone diazide (inhibitor), and alkali insoluble novolak resin. After developing the photoresist with an alkaline solution, the areas that are free of photoresist can be etched, as shown in Figure 1.10. Since only inhibitor molecules exposed to UV light change chemically, it provides excellent edge definition.

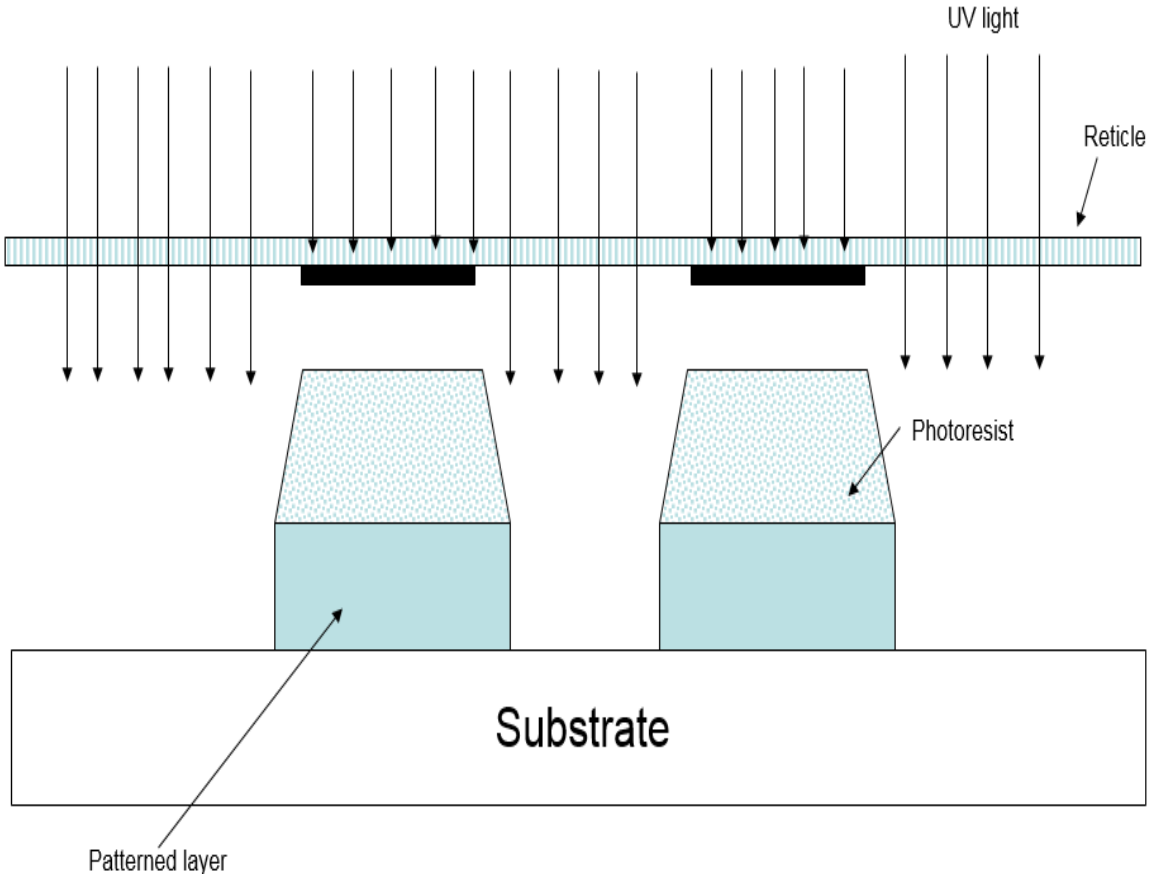


Figure 1.10 Exposed, developed and etched positive photoresist pattern.

1.2.2 Negative photoresist

The negative photoresist is just opposite of positive. In the negative, when the UV light hits the photoresist it polymerizes and becomes difficult to dissolve. Thus, exposed photoresist area remains on the surface and the rest of the photoresist gets removed. Negative photoresist are polymers containing double bonds, usually polyvinyl cinnamate derivatives or cyclized rubber derivatives. **[12]** When negative photoresist is exposed with UV light, with assist from a sensitizer, a chemical reaction occurs and resin forms a three-dimensional, cross-linked polymer. One disadvantage of negative photoresist is that solvents used to develop it can be absorbed into the exposed resist region and cause swelling which reduces the line definition. However, where resolution is less critical such as P-type or N-type well (Tub) regions where CMOS transistors are formed, there are distinct cost advantages in the use of negative photoresist. Figure 1.11 shows the negative photoresist overview.

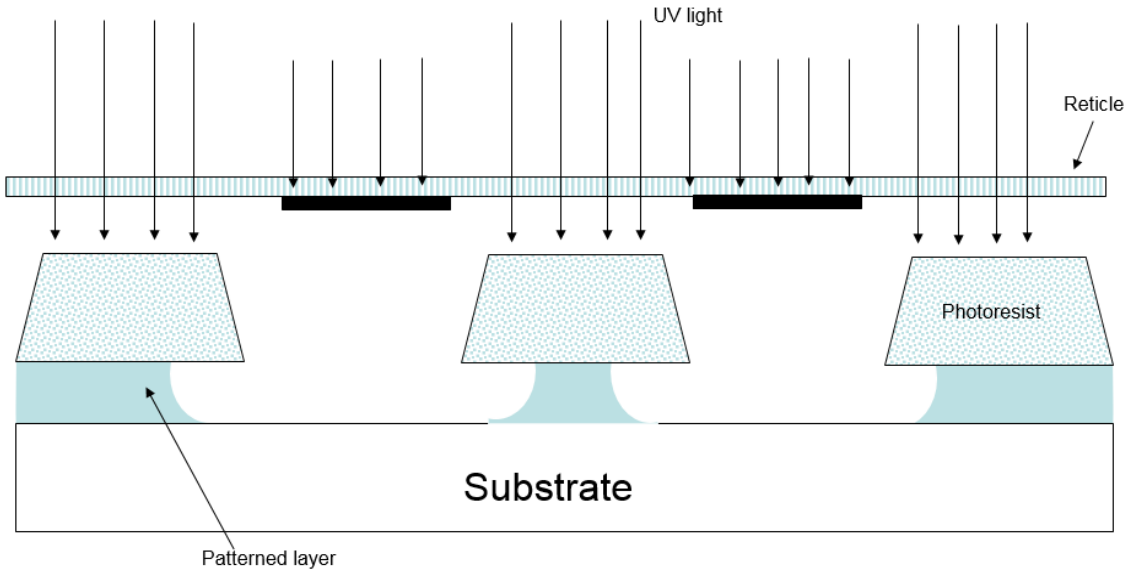


Figure 1.11 Exposed, developed and etched negative photoresist pattern. Patterned layer with isotropic etch profile.

1.3 Modification of material property

1.3.1 Implantation

The implantation step is a key process to convert the neutral silicon area into various N-type or P-type areas. The most commonly used dopants in the semiconductor process are boron for P-type and phosphorous for N-type. The boron atom has 3 electrons in its outer band and phosphorus has 5 electrons. Implanted species either give up electrons or take extra electrons from the silicon lattice and provide extra electrons or holes and thus change the silicon to become N-type or P-type respectively. Figure 1.12 shows the silicon crystalline lattice structures with boron and phosphorus atoms to forming N-type and P-type silicon.

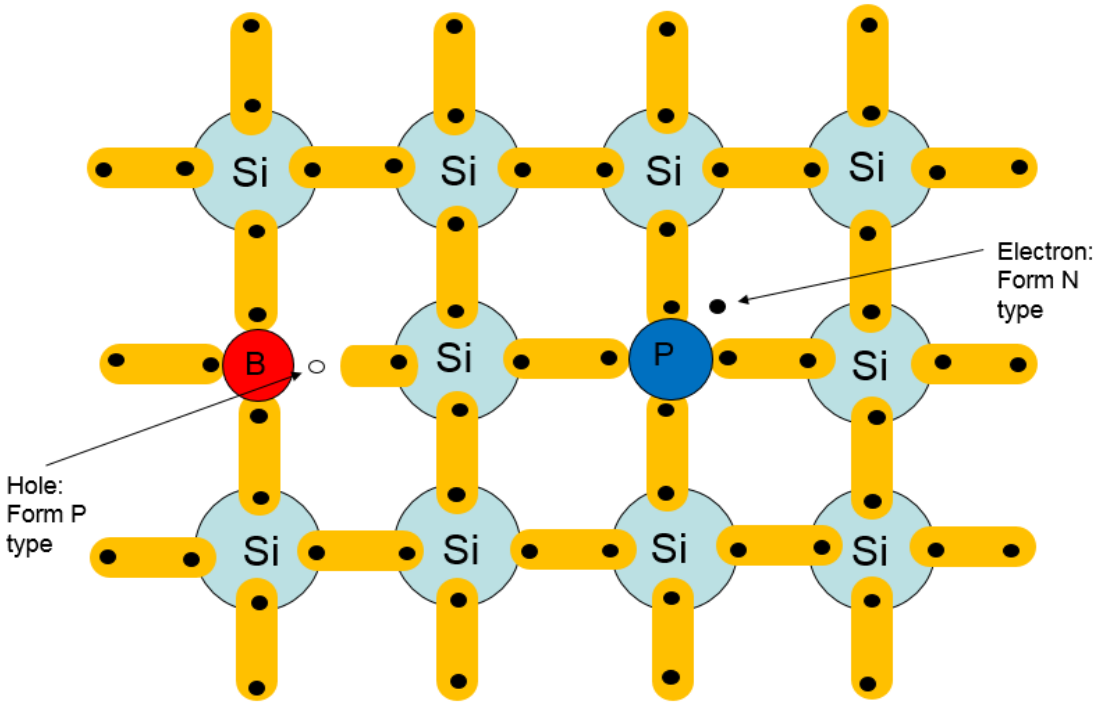
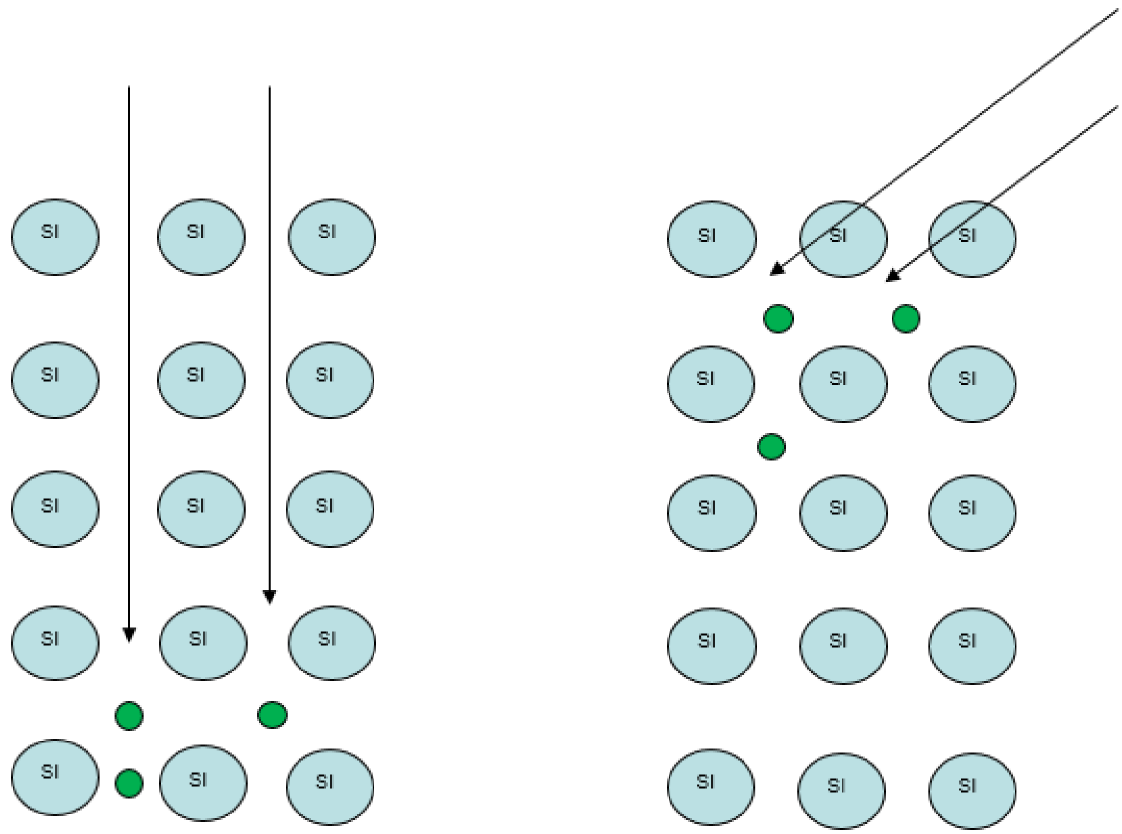


Figure 1.12 Silicon crystalline lattice with B and P atoms are introduced to form a N-type and P-type silicon.

Implanter input process parameters such as the concentration dose level of species, tilt angle, twist angle, and rotation determine effectiveness of implantation. Implanted ions come to rest

in the silicon target by a combination of electronic and nuclear stopping force. The electronic forces are inelastic in nature in which the ions incrementally slow down without large trajectory deflections. On the other hand, nuclear stopping causes elastic collisions, which significantly slow down the ions**. [13]** There is a known problem unique to implantation which is called "channeling", a condition where dopant atoms travel between the silicon atoms lattice, deep into the silicon. When dopant ions go too deep or too shallow, the intended junction depth cannot be formed and it will cause failure in device performance. For example, since the boron atom is relatively small and light compared to the silicon atom, the ion stopping force in the silicon atom lattice is reduced. In order to prevent the channeling effect, fluorine is added to form BF_2 which increases the increase probability of nuclear stopping through hard collisions. Another method is to implant at an angle to increase collisions with silicon atoms, thus preventing them from travelling too far through the lattice. Figure 1.13 shows the straight down implant vs angled implant.



(a) (b)

Figure 1.13 Channeling occurs when dopant atoms are implanted straight down (a) while angled implant (b) will minimize the channeling effect by increased probability of collisions with silicon atoms.

A thin layer of oxide (around 100Å - 150Å) is deposited prior to the implant step to minimize silicon surface damage and contamination during ion implantation.

1.3.2 Diffusion

The key feature of semiconductor technology is the control of the type and concentration of impurities at specific areas of the device. Diffusion is the movement of impurity atoms in a semiconductor material at high temperatures. The objective is to control the concentration, and depth of elements to be doped in silicon crystal. The diffusion coefficient D is given by the relationship below. **[14]**

$$D = D_1 \exp\left(\frac{Q}{kT}\right)$$

where

Q = Activation energy of diffusion (joules (J))

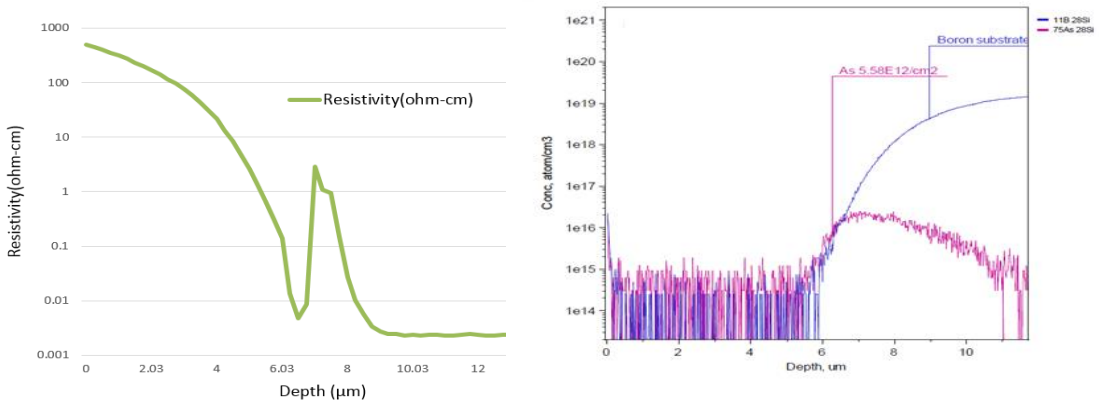
D_1 = Pre-exponential factor (empirical relationship between temperature and rate coefficient from Arrhenius equation)

k = Boltzmann's constant (1.380 $J \cdot K^{-1}$)

T = Absolute temperature (kelvin (K))

Depending on the process flow, specific diffusion process steps can be inserted. However, during the field oxidation process steps, due to their high temperature process conditions, a portion of the oxidation process can be used as a diffusion drive step on previously implanted regions. There are two most commonly used methods to verify diffusion results: Series Resistance Probe (SRP) and Secondary Ion Mass Spectrometry (SIMS). For the SRP measurement, a sample wafer is prepared by angled polishing. Using a four-point probe the resistance of the material is measured from top to bottom along the angled surface. The change of resistance according to depth can be used to determine junction depth and concentration level. With known dopant types and resistivity values it can be converted to a concentration level. On other hand SIMS can detect very low concentrations of dopants and impurities. The technique provides elemental depth profiles ranging from a few angstroms (Å) to tens of micrometers (µm). The sample surface is sputtered/etched with a beam of primary ions (usually O_2+ or Cs+) while secondary ions formed during the sputtering process are

extracted and analyzed using a mass spectrometer (magnetic sector or time of flight.). The secondary ions can range in concentration from matrix levels down to sub-ppm trace levels. Figure 1.14 shows the sample SRP and SIMS results.



(a) (b)

Figure 1.14 Show SRP (a) and SIMS (b) results. SRP results are shown as measured resistivity (ohm-cm) and SIMS result are shown as concentration level (atom/cm^3) . SRP shows a resistivity spike and resistivity drop near the junction between the Si substate and the deposited Si layer.

The thermal budget is a critical factor for device performance and any process steps that require 650 °C or higher process temperature must be counted in the thermal budget to avoid unwanted diffusion to occur that may affect the overall junction depth. Boron tends to be "absorbed or up diffused" by each thermal step and as a result extra amounts of boron may be required to meet specific doping levels of a P-type well after the diffusion step. Another benefit of a high temperature diffusion step is that it anneals the silicon crystal lattice structure damaged by the high energy implant step. Figure 1.15 shows a simulation result of doping density of various wells formed after a diffusion step.

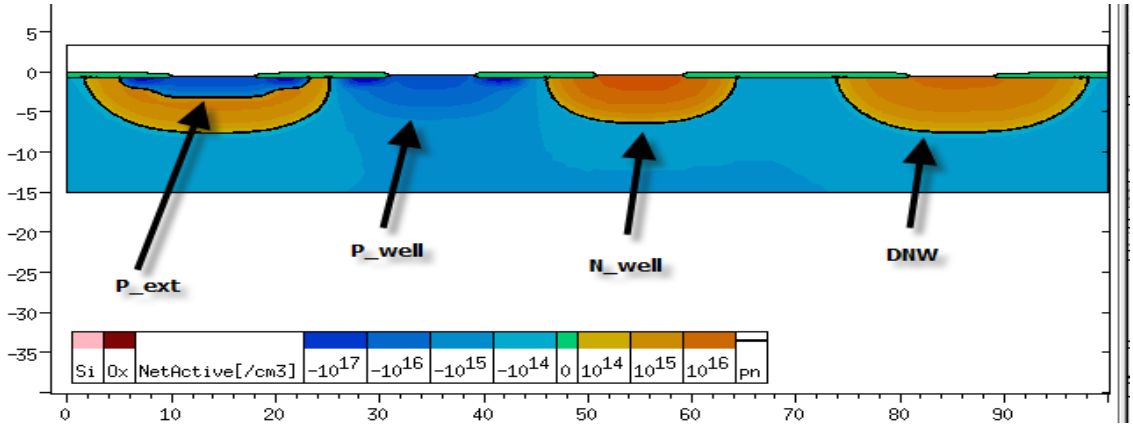


Figure 1.15 TCAD simulation shows various type of body wells formed on silicon after implantation and diffusion. The different color denotes the concentration level and Y- axis shows corresponding depth.

1.3.3 Rapid thermal annealing

Rapid Thermal Annealing (RTA) after implant or metal deposition is required to ensure that implanted species are activated and bonded with silicon atoms by release or acquisition of outer most electrons. Also required is repair of lattice damage during the ion implant and state changes (phase shift) of grown films such as Ni silicide which controls sheet resistance. The salicidation of specific materials such as Ni, Ti, and Co can reduce the sheet resistance by one to two orders of magnitude. This salicidation process is a critical part for submicron CMOS technology as it provides solutions for gate level interconnection resistance reduction. **[15]** Figure 1.16 shows sheet resistance and sheet resistance non-uniformity for Ni silicide that correspond to temperature and different phases of Ni_xSi_x.

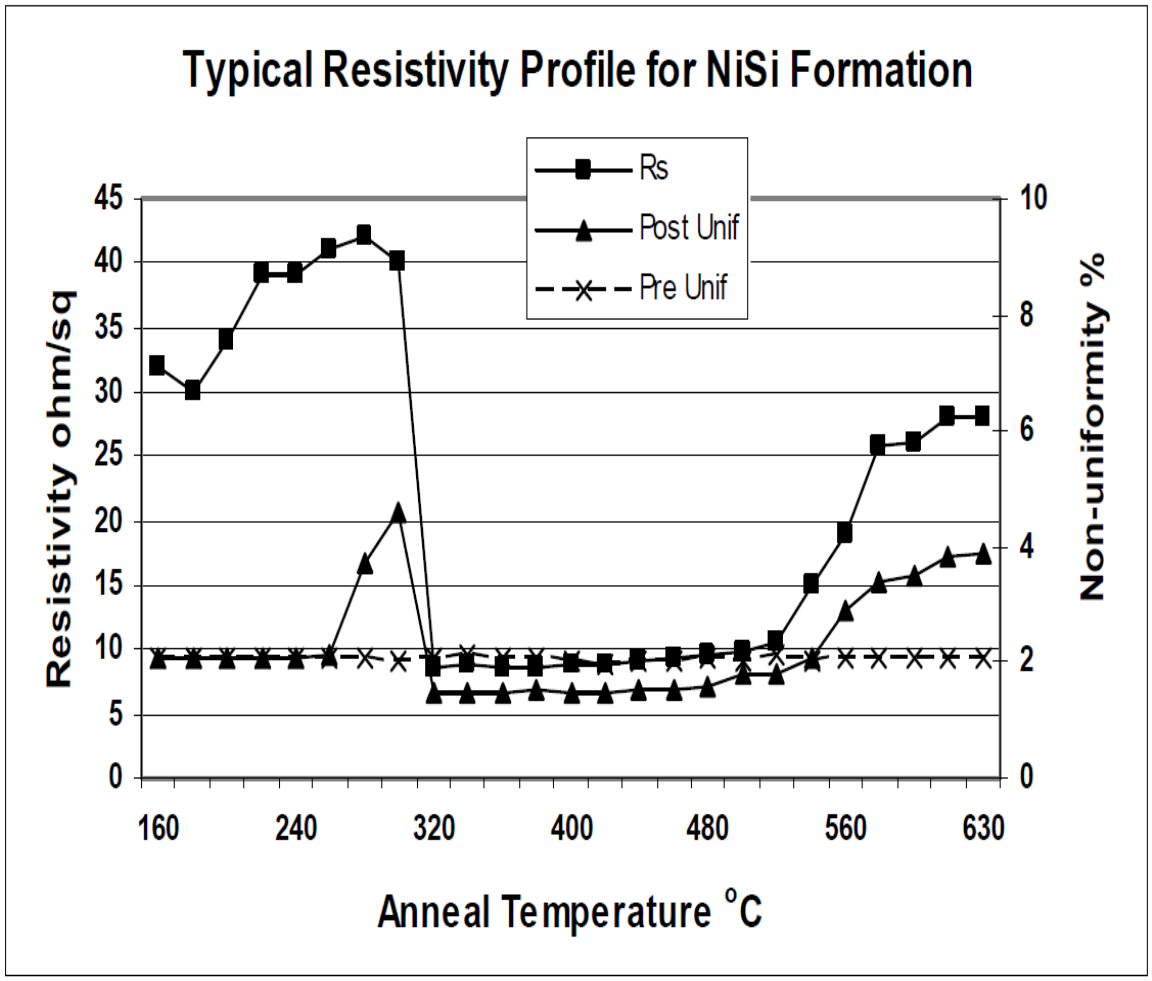


Figure1.16 At temperatures below 300 °C Ni_2Si is formed, and at 400 °C phase shifted and lower resistivity NiSi is formed but at temperature above 500 °C phase shifts occur again to form NiSi2 with higher resistivity film**. [16]**

The Rapid Thermal Process requires high temperatures (400 °C - 900 °C) for a short period of time (around 10-50 seconds). Using short thermal exposure times for the RTA step minimizes redistribution of dopant where additional diffusion will be limited. In certain cases, multiple RTA steps are used to control specific sheet resistance. Control of temperature uniformity across the heating source is critical for the RTA process since it is very difficult to accurately measure the absolute temperature range. To verify the temperature of the RTA tool, periodic temperature measurements are done on test wafers using a pyrometer. In general different

random locations on the test wafer are measured and compared with reference data collected previously. **[17]**

1.4 Removal

Wafer etching and cleaning in semiconductor processing requires Ultra-High Purity (UHP) chemicals for expected device performance and yield. A major source of yield losses can be attributed to particles and contamination from inadequate cleaning processes. Particulate impurities such as metallic contamination can lead to leakage increase or shorting between signal lines of the device**. [18]** Also, if particles generated during the etch process are not removed properly, this can also cause device failures. Particle inspections on critical steps after etching and cleaning steps are part of the normal process flow where with early detection of particles, wafers can be submitted for rework or scrapped inline in order to replace with recovery wafers. The cost of fully processed wafer scrap at testing is much higher than early termination.

1.4.1 Etching

There are several different types of etching required during semiconductor processes based on materials and process modules to properly pattern a given device. There are two major types of etch processes in semiconductor process. One is plasma based dry etch and the other is chemical based wet etch. Plasma based dry etch is an anisotropic process that removes materials in an applied single direction. On the other hand, wet etches are generally isotropic which remove material equally in all directions.

1.4.1.1 Wet etch

Wet etching has been used from the beginning of semiconductor process technology and it is still widely used for basic device manufacturing. A key advantage of wet etches is the high selectivity and ability to remove undesirable ions and contamination from the surface as it etches. High selectivity means for example, while diluted HF is used to remove silicon dioxide, the surrounding silicon area will not be etched. One disadvantage of wet etches is the isotropic nature of the etch profile which limits the ability to form a specific structure such as a metal line. Figure 1.17 shows the wet etch isotropic profile.

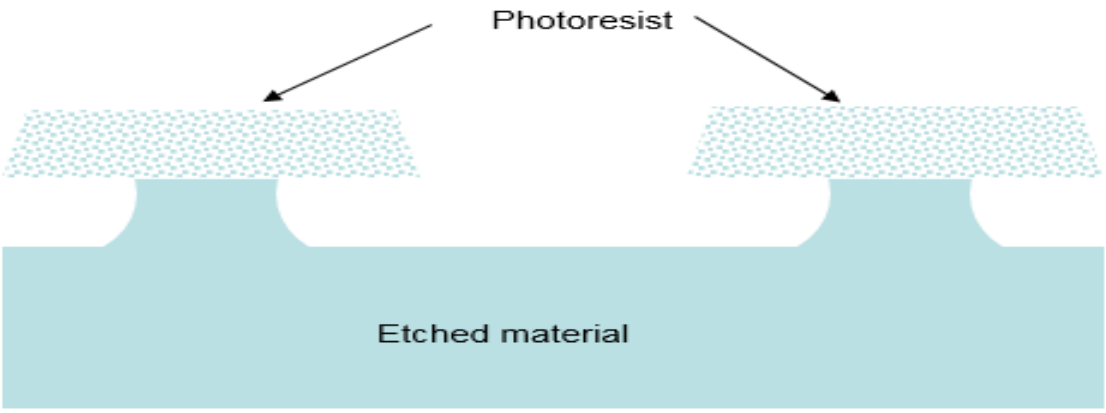


Figure 1.17 Wet etch produces an isotropic etch profile of material.

In most cases, hard reticle layers such as oxide and nitride layers greater than 4500Å will receive the wet chemical etch since it is blanket removal of these layers where etch profile does not matter. Also, thick metal above 3 µm may use wet etch to remove a portion of metal first, and the rest of etch is done using plasma based dry etch. The reason for this two-step etch is to reduce plasma etch tool maintenance. Prolonged use of Radio Frequency (RF) plasma etching will require more frequent of preventive maintenance of the tool. Table 1.1 shows most commonly used wet etchants.

Table 1.1 Commonly used wet etchants in semiconductor process

Etchant	Purpose	Note
Buffered Oxide Etch (BOE) HF + NH_4F	Selective etch for silicon dioxide	Ratio of 40% NH_4F and 49% HF for BOE
Diluted HF	Oxide removal	1 part by volume 49%HF : 100 parts H_2O
H_3P0_4 with HNO_3 and acetic acid	Selective etch for Aluminum with 1% silicon	Etch rate – 200 nm/min at 25°C
H_3PO_4	Removal of silicon nitride	180 °C temperature required
H_2O_2	Remove Ti:W	
H_2O_2:NH_4OH:H_2O	Remove Ti:N	

1.4.1.2 Dry etch

Various materials such as polysilicon, silicon dioxide, oxide, nitride and metal layers can be dry etched in a RF generated plasma. Polymer buildup during a plasma based etch step will coat side walls which will improve the anisotropic etch profile by preventing side wall etching. Over-passivation (more sidewall polymer formation) occurs on metal sidewalls that face large open spaces because they are more accessible to the passivating gases. Polymer formation during metal etch is a normal characteristic of metal etch processing. This sidewall polymer formation helps protect the AlCu sidewalls so they remain smooth, void-free and not pitted by the Cl_2 gas that is used to remove the AlCu from the areas being cleared. Any time metallic-organic molecules (i.e., metal etch polymers) are generated, they are more difficult to remove in general by the wet etch and thus the plasma based etch is needed. Dry etch rate can be

affected by density of the pattern that needs to be etched on the wafer. Due to design complexity, some areas may have isolated patterns and on the same chip there could be a high density area that needs to be etched at the same time. This is known as loading effect and the etch effectiveness can be best stated by this relation:

$$R = \frac{1}{Ae}$$

where

R = Etch effectiveness (cm^{-2})

Ae = Exposed surface area (cm^2)

The loading effects from large exposed surface areas will generate more quantity of sidewall polymer, thus incorporating also more metallic ions (Al, Ti) into the polymer. This makes the polymer that much harder to be removed by the wet clean. In general, discrete technology with high field density typically produces significantly more polymer than other CMOS resist types during dry metal etch steps, especially with 3 µm thick metals that impacts the surface area. In order to minimize pattern density differences, a dummy pattern can be placed on less dense areas which improves etch rate uniformity on the wafer. Figure 1.18 shows the anisotropic dry etch profile.

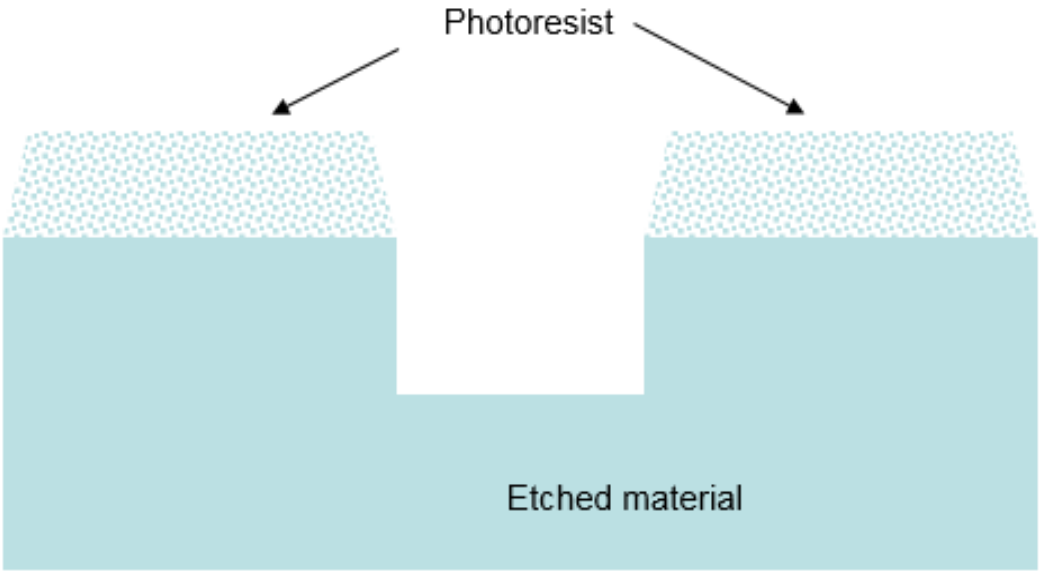


Figure 1.18 Anisotropic dry etch profile of material where vertical side walls are formed.

There are two different types of dry etches: pure chemical process etch and physical material removal process etch.

1.4.1.2.1 Chemical dry etch

Energetic ions from the plasma alter the substrate so that neutral free radicals can chemically combine with the material, forming volatile products. At very low pressures of the plasma chamber, ion flux can bombard the surface. Since the surface of sidewall is perpendicular to the electric field, the sidewalls receive minimal ion flux which assists an anisotropic etch profile. The main advantage of ion enhanced chemical dry etching is due to anisotropic profile, so higher resolution can be achieved, for example smaller metal line width.

1.4.1.2.2 Physical dry etch

Highly energetic argon ions physically bombard the surface of material and literally knock off atoms from surface of etch material. This process is often called sputtering or an ion milling etch process. Inert gas such as argon is commonly used to reduce any unwanted chemical reaction with etching material. This type of etching is generally used for pre-clean of surfaces or to remove residues such as photoresist.

1.4.2 Cleaning

Each major process requires clean steps. The objective of cleaning is to remove particles, contamination, and metal ions from the surface of the wafer that can potentially cause reliability issues or device failures. For example, photoresist after its use needs to be removed, and plasma based sputter etch is used to remove the bulk of photoresist. Heavy inert gas such as argon is used as the primary material for sputter etch. It is followed up by a wet chemical (sulfuric acid with ammonium hydroxide) clean to remove the rest of the photoresist. In the ambient environment, 10-20Å of oxide grows on top of the silicon surface and this native oxide needs to be removed to proceed with next process step. Diluted 100:1 HF is commonly used to remove this native oxide layer.

There is a cleaning step that is in the standard set of wafer cleaning steps called “RCA clean”. Werner Kern developed the basic procedure in 1965 while working for RCA, the Radio Corporation of America and thus the term RCA is associated with this cleaning method **[19].** There are four basic steps required for RCA clean. Around 10 Å of native oxide which forms in the ambient environment can be removed with diluted hydrofluoric acid that can be added before SC-2 step as listed below.

- Step 1. (SC1): organic clean + particle clean:

$NH_4OH : H_2O_2 : H_2O$ in the ratio of 0.05 : 1 : 5 at 75 °C

- Step 2. Deionized (DI) water wafers rinse:

25 °C DI wafer rinse – to avoid micro roughening by NH_4OH

- Step 3. (SC-2): metallic ionic clean:

$HCl : H_2O_2 : H_2O$ in ratio of 1: 1: 5 at 75°C

- Step 4. DI rinsing and drying :75 °C DI wafer rinse

The piranha solution, also known as piranha clean, is a mixture of sulfuric acid (H_2SO_4) and hydrogen peroxide (H_2O_2), used to clean organic residues and metal atoms. During the cleaning process, it passivates the silicon surface with an oxide layer due to an oxidizing agent in the solution. Piranha solution must be prepared with great care because the solution is highly corrosive and a powerful oxidizer. Surfaces must be clean and completely free of organic solvents from previous wash steps before coming into contact with the piranha solution. The piranha solution cleans by dissolving organic contaminants, and a large amount of contaminant will cause violent bubbling and a release of hydrogen-rich gas that is harmful to health, and could cause an explosion. Therefore, the container of the piranha solution must have sensors that detect excessive bubbling and also be able to exhaust the outlet. The byproduct gas also needs to go through a scrubbing chamber which neutralizes the gases before release to the environment.

1.5 Planarization

To form multiple layers of metal stacks, planarization of previous layers becomes critical. The accurate alignment of contact and via structures to submicron active areas and metal interconnects can be severely affected by the planarization of the stacks and it can directly impact the overall yield of product. There are two major areas of concern with non-flat surface layers during backend semiconductor processing - photolithography and etching. If the incoming layer is not flat, photoresist will follow the contour of topography where the thickness of the photoresist will be different on concave areas compared to convex areas. This thickness difference across the wafer will cause development issues since thicker photoresist areas may not be adequately developed or removed. Also, the stepper will be unable to focus accurately on high points and lower points at the same time. Therefore, any marginal quality of incoming layer will cause overlaying problems on subsequent steps. The second issue is at the etching step. Too much variation in the topography can result in non-uniform etch across the wafer. For example, etch rate at a concave area is slower than at a convex area for plasma-based sputter etch since a concave area will be less exposed to incoming ions. Under-etched metal can cause failure of the device due to shorting of two metal lines. Over etch to compensate the etch rate difference will eventually start etching sideways in the areas that are already cleared. These problems get worse with each subsequent level, because the topography becomes more extreme since it is additive and magnifies the concave and convex areas.

There are two planarization techniques that are used depending on whether the incoming wafers have metal layers. An aluminum based interconnect metal layer is deposited at around 400 °C or less and if planarization process techniques requires temperatures significantly higher

than 400 °C the metal will melt. One planarization process technique that can be used before metal is deposited on the wafer is applying Boro-Phospho-Silicate-Glass (BPSG). This is a form of oxide that has been lightly doped with boron and phosphorus to help it "reflow" under a temperature range from 750 °C - 850°C. By adding dopant, it decreases the melting point of the silicate glass. Both N and P type dopants were used to minimize the conductivity of the layer since BPSG is used as part of insulation layer. The one percent increase of boron concentration with phosphorous concentration causes about a 100 °C drop of the glass flow temperature and approximately 50% decrease of the etch rate in 1:6 buffered HF solution. **[20]** The BPSG step involves first depositing a thin layer of undoped oxide layer that acts as a barrier layer and then a layer of BPSG. This thin layer of undoped material layer will prevent some of the boron and phosphorous in the BPSG to diffuse into the silicon over time. After the deposition of BPSG, wafers are heated in a furnace, causing the BPSG to fill in the concave and convex areas. Another planarization process technique that is used for wafers already having a metal layer utilizes a Tetraethylorthosilicate ($Si(C_2H_5O)_4$, TEOS) oxide layer which can be deposited using a PECVD tool at temperatures around 400°C. For technology greater than 0.35 µm geometry, around 5000Å TEOS is deposited. To further enhance the flatness of the wafer on both BPSG and TEOS planarization, a Chemical Mechanical Polishing (CMP) step is used where the excess layer on top is polished to provide additional flatness of the layer. Figure 1.19 shows a two metal layer cross-section view after a planarization process step.

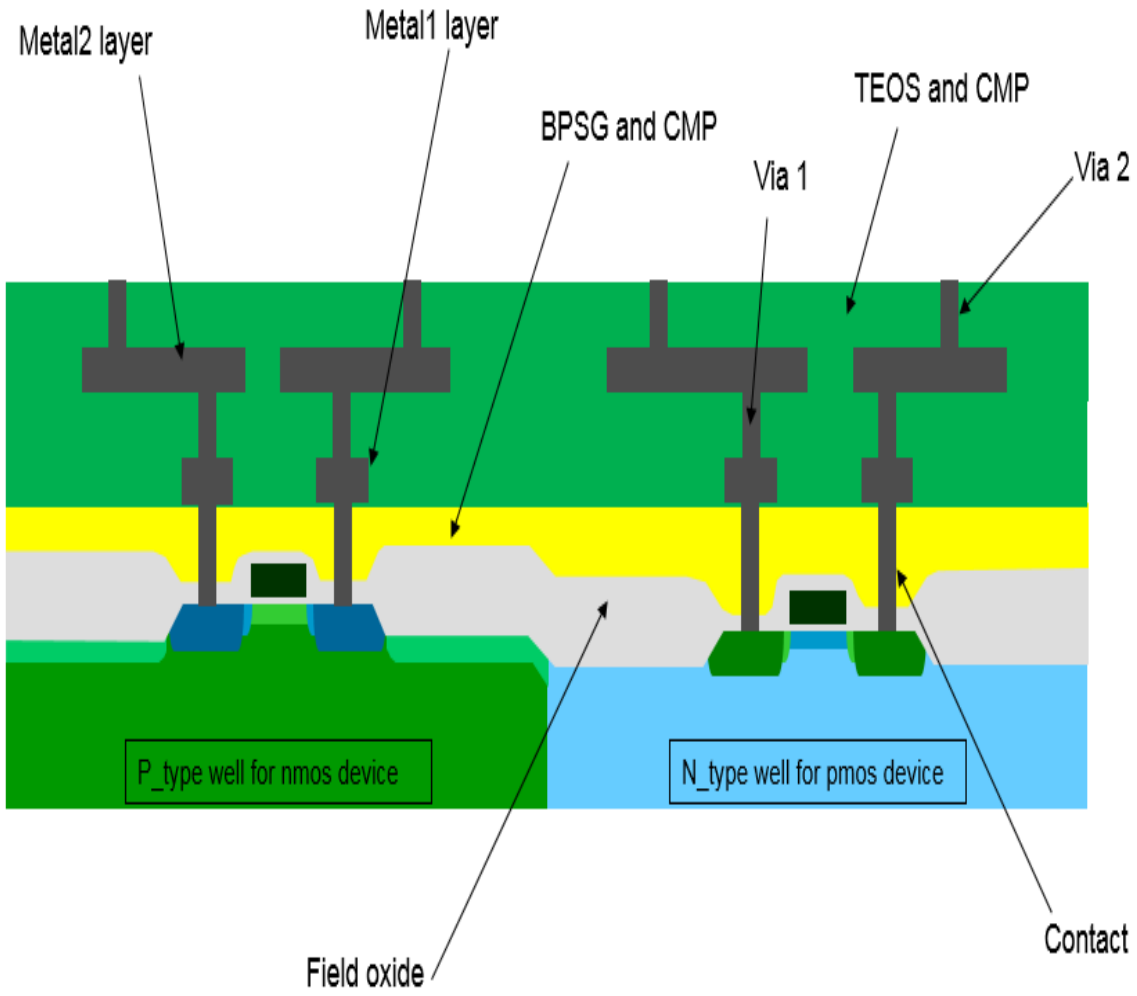


Figure 1.19 Insulation layer with Chemical Mechanical Planarization (CMP).

2. Yield improvements

In semiconductor manufacturing, once the technology and process are established, yield improvement is one of the essential areas for continued success of the process. In general the yield is represented by the functionality and reliability of the integrated circuits or discrete device produced on the wafer surface. During the manufacturing of semiconductor products, yield loss is caused by intrinsic or extrinsic defects, process variations and sensitivity of integrated circuit designs. **[21]**

In many cases, investigating failures and developing a yield improvement solution is not a straight forward activity and may require many different experiments to eliminate factors that do not contribute to the yield loss and to verify the significant factor that is the root cause of the problem. Since the semiconductor process requires up to hundreds of process steps to create a functioning device, interactions between different process steps can impact the overall process yield. Therefore, it is critical to review pre- and post- process steps of known areas of concern. Static Random Access Memory (SRAM) failures due to high leakage currents is an example of functional failure of the device. This kind of yield loss can be attributed to process variation or design sensitivity. Gate oxide failures are one of the major contributors to the unreliability and yield loss of semiconductor devices. While most gate oxide issues are created within the gate oxide module, post-gate processing can also contribute to gate oxide defects. Specifically, defects can be formed during the salicide block module on polysilicon; a module that serves as a hard reticle for selective salicide formation.

2.1 Eliminating a polysilicon hole defect created during oxide removal

As part of a high voltage 0.35 µm technology qualification, Gate Oxide Integrity (GOI) tests including Time Dependent Dielectric Breakdown (TDDB) and Voltage Breakdown of Dielectric (V_{BD}) were performed. TDDB and V_{BD} tests showed that PMOS GOI capacitors exhibited high failure rates on large square active area (PCSQ1), poly edge (PCPE1) and bird's Beak (PCBB1) structures, while all complementary NMOS structures were relatively defect free. Data analysis indicated that the failures were immediate breakdown which indicates extrinsic defects in nature. Excluding these failures otherwise indicated that the intrinsic lifetime for the gate oxide met the 10-year requirements. Table 2.1 outlines the test structure types, its dimensions, and its failure rate.

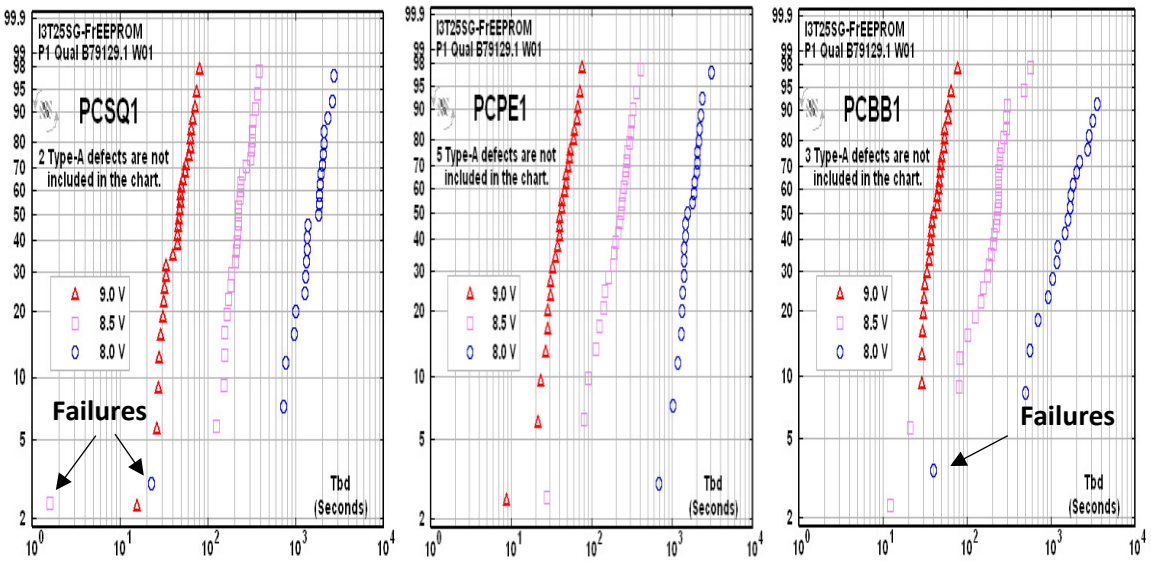
Table 2.1 Oxide qualification test structure module

Structure name and size	**PCSQ1 (Active Area)**	**PCPE1 (Poly edge)**	**PCBB1 (Bird's beak)**
Active Area (μm^2)	243,000	243,000	243,000
Edge length (µm)	PE: 0 FE: 2,070	PE: 194,400 FE: 720	PE: 0 FE: 195,120
# of defects (out of 116 capacitors)	15	28	41

To perform TDDB tests, one first determines the applicable voltage level that will not breakdown the given oxide layer immediately and apply this voltage at a range of temperatures and observe the time to breakdown. In general an oxide layer can handle 8MV/cm. A V_{BD} test is applied with incrementally increased voltage to the oxide and recording at what voltage level

the oxide failed. In general, the failure criteria of the oxide is 1uA current detected when voltage is applied.

Figure 2.1 shows the cumulative probability of TDDB result of test structures where some samples failed almost immediately when external voltage was applied which indicates the oxide quality issues.



(a) (b) (c)

Figure 2.1 TDDB result of test structure modules. The Y axis is cumulative probability in percent and the X axis is time to breakdown in seconds. (a) large square active area (PCSQ1), (b) poly edge (PCPE1) and (c) bird's Beak (PCBB1) structures. (ONSEMI reliability report).

Figure 2.2 shows the V_{BD} result where some of the samples had near zero volt breakdown voltage which indicates that the oxide may have been shorted or severely damaged at the onset.

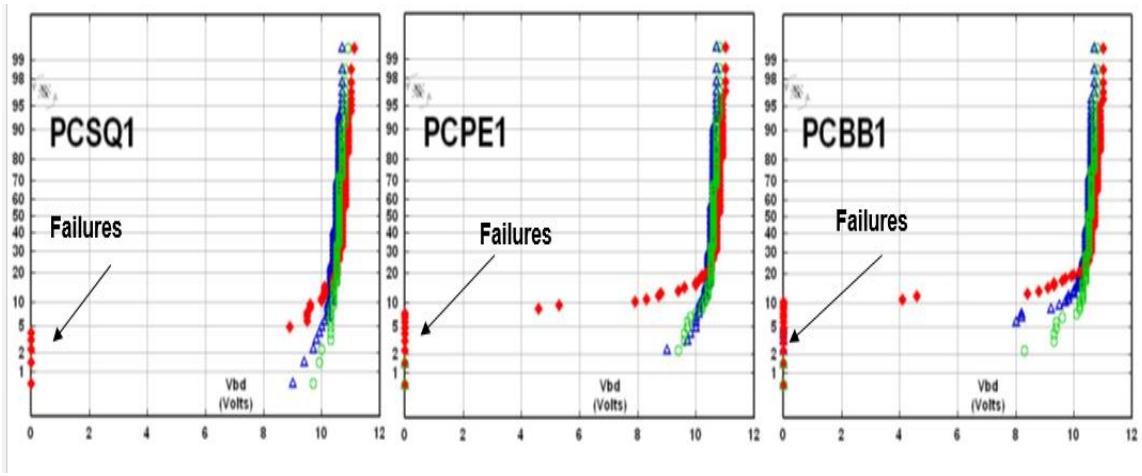


Figure 2.2 V_{BD} test result of test structures. . The Y axis is cumulative probability and the X axis is breakdown voltage in volts (ONSEMI reliability report)

These failed samples were submitted for Failure Analysis (FA) by photo emission and SEM review. The SEM cross-sections showed a round hole in the polysilicon top plate. These polysilicon holes were completely devoid of polysilicon and typically showed titanium silicide which is conductive material on the sidewalls and hole bottom. Figure 2.3 and 2.4 shows a typical appearance of these defects. This failure analysis result directly confirms the early failures observed during the TDDB and V_{BD} tests.

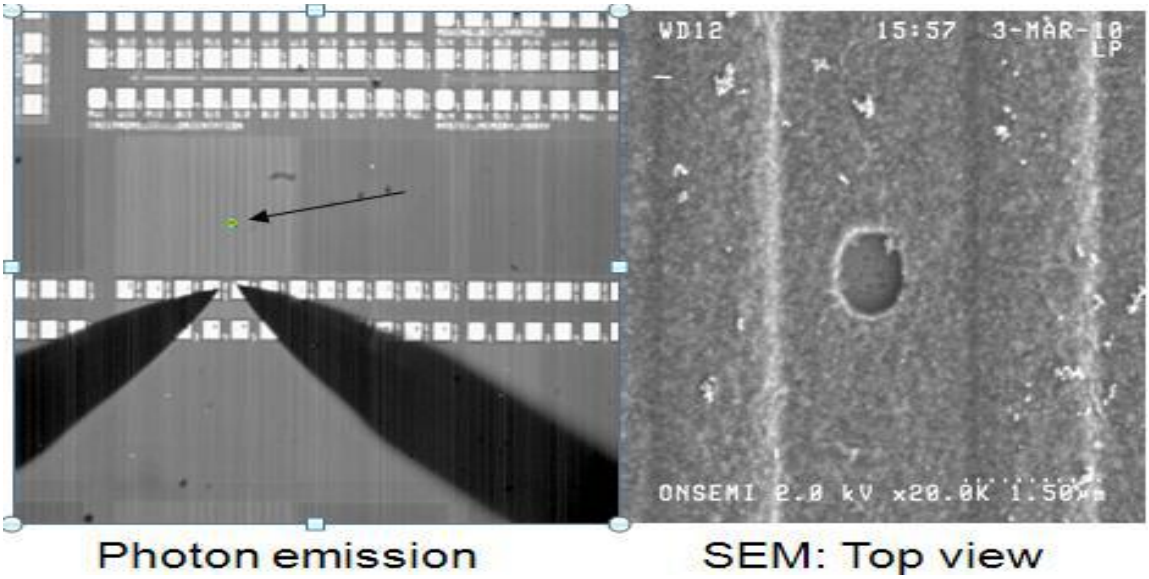
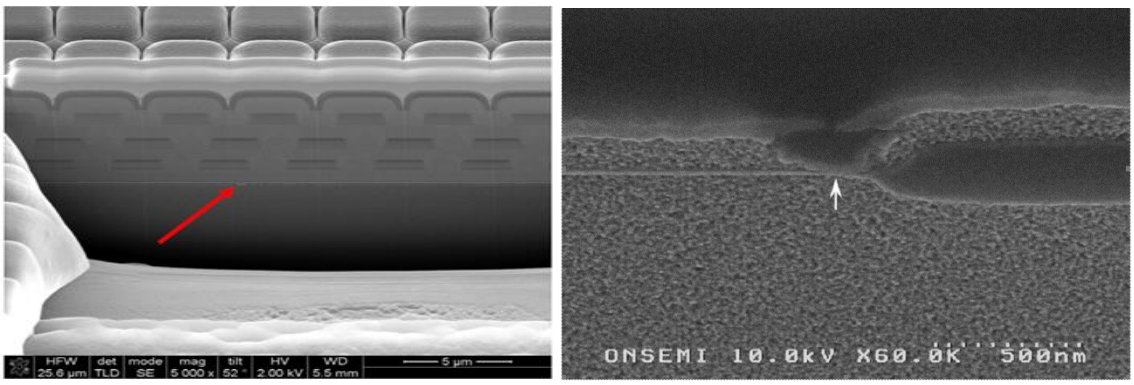


Figure 2.3 Photon emission spot and SEM image of defect site. Failure analysis work by ON SEMI



(a) (b)

Figure 2.4 The picture (a) show Focused Ion Beam (FIB) cross-section cut view and (b) polysilicon hole with portion of Ti-silicide layer at the bottom of the hole which will cause short to oxide layer. (Failure analysis work by ON SEMI)

Based upon the failure analysis results, in-line product scans were intensified in the manufacturing line using a KLA2132 bright-field wafer scan tool. The focus of these scans was in the salicide block module, as listed in Table 2.2.

Table 2.2 Salicide block processing module

Process Steps	**Comments**
SiO_2	Chemical vapor deposition TEOS
Si_2N_4 deposition	Chemical vapor deposition
Patterning / etch / ash / clean	
Pre-metal clean	Buffed oxide etch wet clean
Ti deposition	Endura metal deposition
TiSi formation	Rapid thermal activation
Ti Strip	SC1 strip

Stepwise scan recipes were developed to determine where in the salicide block processing module that the defect type was first formed. For these inline scans, a relative visual defect density (D_0) was determined through review of the defects found with KLA scans. This is defined as number of defects per square centimeter of area scanned. Through the in-line scans, the polysilicon defect was determined to be first detected after the pre-metal clean (see Table 2.2). This clean is used to remove any native oxide prior to Ti deposition and can be considered as a primary process step in the creation of polysilicon hole defects. With this discovery, additional experiments were performed to identify contributing factors to the defect formation and to understand the mechanism of polysilicon hole formation. Figure 2.5 shows a typical KLA map with the defect location on a given wafer of 8 inches in diameter.

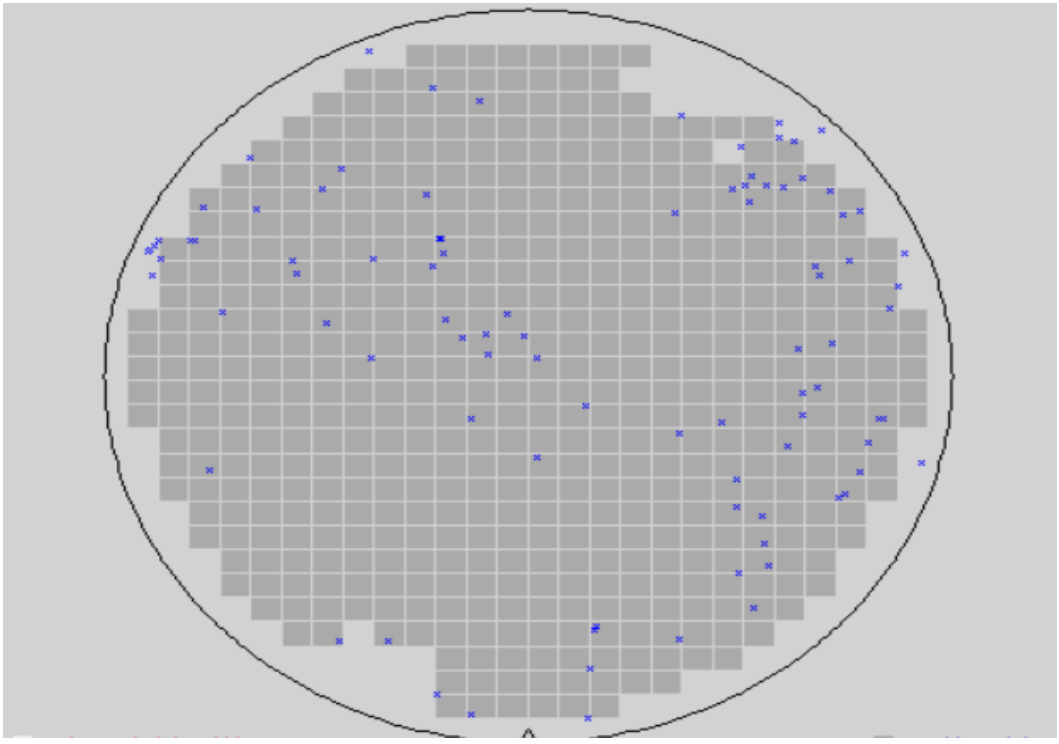


Figure 2.5 Typical KLA wafermap that shows defect location on 8in wafers.

2.1.1 Contributing factors experiments

In order to reduce processing time, a short-loop flow was created based on the salicide block module utilizing patterned implanted wafers. With the short-loop flow wafers, many experiments were designed to find the root cause for the defect of interest. Below are summaries of the important experiments that led to determining a root cause for the polysilicon hole defects. All experiment as results are based on visual defect counts.

2.1.1.1 Implant damage

One theory of the defect formation was that the phosphorus ion (P+) implant is physically damaging the polysilicon and allowing the pre-metal clean to remove the damaged polysilicon. An experiment designed to test this theory used B, BF_2, and argon (Ar) as implant species with same dose and energy condition as the standard P+ source/drain implant used by the technology. The inert argon gas was chosen to focus on study the physical damage effect of an implant on the polysilicon (removing chemical interactions). Table 2.3 summarizes the visual defect density for the implant damage theory experiment.

Table 2.3 Effect of implant on polysilicon holes

	BF_2 (POR)	**B**	**Ar only**	**B and Ar**
Overall KLA visual defect counts	88	271	2	561
Polysilicon visual defect counts	70	221	0	505

The experimental results showed that any boron base implant produced polysilicon holes, while argon did not produce any holes. This result indicates the polysilicon hole was not the result of physical bombardment during the implant process.

2.1.1.2 Plasma charge damage

The second possible cause investigated was whether the P+ source/drain implant conditions were the initiator of the problem. Specifically, it was hypothesized that the implant recipe arc current was too high and thus caused some type of electrostatic discharge damage to the polysilicon which was later enlarged via the pre-metal clean. For this experiment, the arc current was varied to produce poor electron neutralization conditions with results shown in Figure 2.6.

Scans showed that a polysilicon defect was created, but differed in characteristics from the defect under study. The defect not only formed in the polysilicon but also in the Si substrate in the case of high arc current. Also the created defect had the appearance of a typical electrical discharge, and thus was not circular in nature. From these observations, it was thus concluded that the polysilicon hole defects were not due to wafer charging during ion implantation.

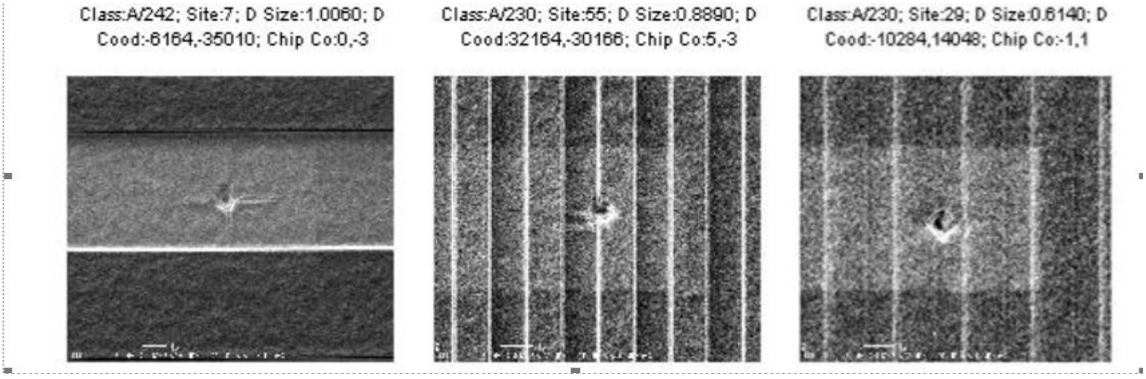


Figure 2.6 SEM image of high arc current implant charging damage defects on wafer. (Failure analysis work by ON SEMI)

2.1.1.3 Pre-metal clean chemical

As KLA scans indicated that the polysilicon hole defects first appeared following the pre-metal clean, an experiment was centered around the cleaning conditions. The process of record pre-metal clean used a buffer oxide etch (BOE: HF + NH_4F) into HF based clean in order to remove native oxide prior to titanium deposition. An alternative of a 100:1 HF solution was proposed. A split was performed between these two chemicals and the resultant defect density is shown in table 2.4. A significant reduction of polysilicon hole production with HF was observed.

Table 2.4 Pre-metal cleans

	HF	**BOE**
Overall KLA visual defect counts	6.5	105
Polysilicon visual defect counts	0	105

2.1.1.4 LPCVD vs PECVD vs HDP films

Silicon rich oxide with a RI (refractive index) of 1.56 when used as hard photomask for selective salicidation, also prevents polysilicon hole generation.**[22]** To determine if this was the case, a short-loop experiment was designed to test the SiO_2 deposition method. In the current process, both SiO_2 (TEOS generated) and Si_3N_4 are used to form the salicide blocking photomask. For this test, both LPCVD, PECVD and High Density Plasma (HDP) oxide films and LPCVD, and PECVD nitride films were studied. The PECVD films employed standard operation condition the same as other intra-metal dielectric films used in the fabrication line, while the HDP wafer received a SiO_2 film having a refractive index of 1.56. Table 2.5 shows defect density for this Design of

Experiment (DOE). From the results it can be concluded that the HDP silicon rich film is capable of controlling the generation of the polysilicon hole defects but not to the same level as seen by changing the pre-metal clean chemistry as seen from Table 2.4.

Table 2.5 Different oxide material split experiment result

	Process of record	HDP (Si Rich-RI of 1.56)	HDP + RTA	PECVD TEOS	PECVD TEOS + RTA	PECVD SiN	SiN removed via hot Phos
KLA visual defect	88	3	3	184	23	26	139
Polysilicon visual defect	70	3	2	184	23	26	125

2.1.1.5 Different salicide block module

The fact that there are salicide block module integration differences between 0.35 μm processes in our fabrication line was noted. In-line and electrical testing showed that one integration scheme did not have the polysilicon hole defects while the other did. The major difference was that salicide block module included a Si_3N_4 layer on the process that had the polysilicon hole defects. A short-loop experiment was again used to determine if the factor was Si_3N_4 or other previous processing differences between the two flows. Splits were done around the salicide block module steps. Table 2.6 shows no polysilicon defect creation when the Si_3N_4 layer is omitted from the flow.

Table 2.6 Salicide block module integration difference

	Process of record	Thicker TEOS	Extended TEOS	No SiN layer	No SiN with BOE	STD with HF
KLA visual defect	146	73	143	14	24	7
Polysilicon visual defect	146	43	143	0	0	0

2.1.2 Discussion

Several hypothesis tests were developed based upon known failure mechanisms in an effort to resolve and understand the polysilicon hole defect issue. Based on experiment results, the polysilicon hole defects arise from a three-way interaction between the boron implant in polysilicon, subsequent thermal processing, and the BOE chemistry. During the salicide block depositions, the boron agglomerated in the implanted polysilicon due to the thermal energy in the LPCVD deposition steps. This aggregated boron-rich-silicon is then etched by the BOE etch chemistry with NH_4F with metallic ion contamination in the HF which can lead to a polysilicon hole defect.

The purity of process chemicals plays a vital role on device reliability. Anionic contaminants such as chloride, sulfate, phosphate and nitrate in HF acid have been detected using Ion Chromatography (IC) techniques. **[23]** It is known that enhanced pitting and irreparable damage

to the silicon surface can occur from cleaning solutions where are contaminated by Fe or Ag in HF solutions (HF/$Fe(NO_3)$ or HF/$AgNO_3$) or SC1 solutions (part of RCA cleaning). **[24] [25]**

$Fe^{3+} + e^-_{VB} \rightarrow Fe^{2+}$; $Ag^+ + e^-_{VB} \rightarrow Ag^0(s)$: Cathode reaction

$Si(s) + 2H_20 \rightarrow SiO_2 + 4H+ + 4e^-_{VB}$: Anode reaction

$SiO_2(s) + 6HF \rightarrow H_2SiF_6 + 2H_2O$

Also, higher concentration of NH_4OH must be controlled to avoid micro-roughening the silicon surface. This occurs because in basic solution, silicon is dissolved and in hot water small amounts of NH_4OH will enhance the micro-roughening. **[26]** Additionally, by changing the salicide block integration and film stoichiometry, the defects were modulated. This last result can be explained in a variety of ways: (by a polysilicon grain boundary stuffing) **[22]**, changes to film stress, or changes to the wafer thermal budget. can all explain the phenomenon.

With the results listed in table 2.6, it can be concluded that the BOE solution contaminated with metallic ions allows for polysilicon to pit and eventually form a circular hole.

Following implementation of this new pre-metal clean with 100:1 diluted HF instead of BOE, no polysilicon hole defects have been detected with in-line inspections. Additionally, V_{BD} testing confirmed the defect density results; HF based pre-metal clean shows improved V_{BD} performance (Figure 2.7). **[27]**

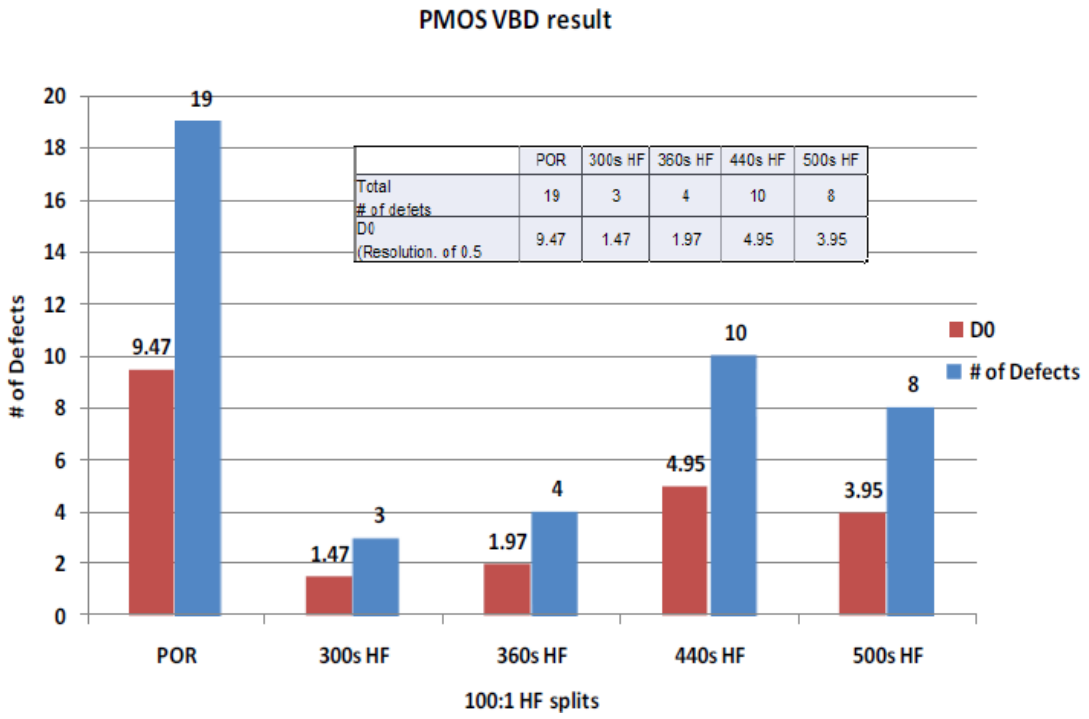


Figure 2.7 V_{BD} result on 100:1 HF

3. Conclusion

Even well-established process steps and methods can cause yield loss on some products with different layout design, components used and pattern densities. The polysilicon hole defects arise from a three-way interaction between the boron implant in polysilicon, subsequent thermal processing, and the BOE chemistry. During the annealing and activation thermal steps after implantation, boron with a high solubility solid limit will be strongly affected by the morphological lattice of polysilicon and will cause the grains to grow in size during redistribution, which rearranges the grain boundaries.**[28]** It is known that electrically inactive boron and energetic ions can create a substantial number of defects.**[29]** Therefore, grain boundaries and lattices of polysilicon are physically impacted by the boron implant and become more susceptible to micro-roughing due to NH_4OH and possible metal contamination in BOE chemistry when adding ammonium fluoride.

In this failure case, questioning the well-accepted BOE etch chemistry would not be the top priority of most investigations because all the process parameter measurements were well within process control. Additionally several product lines also use the BOE without any failures reported. This type of yield improvement can be challenging where precise understanding of electrical failure is required and experiments to disprove or prove the failure mechanism must be done. After several experiments were performed the simple solution was to replace the BOE with diluted HF since it was the most cost effective. By understanding critical methods and their applications to semiconductor processes as well as failure analysis on various stages of the process steps significant leverage for yield improvement is obtained. With concrete experiment

results, process modification was implemented so that product yield was improved and reliability qualification of the product could be achieved.

4. References

1. Taur, Yuan, Ning, Tak H, *Fundamental of Modern VLSI Devices*, Cambridge University Press, p.96. (1998)
2. M. M. Atalia, *In Properties of Elemental and Compound Semiconductors*, Vol. 5, ed. H. Gatos, Interscience, New York, p. 163, (1960)
3. Neil.H.E. Weste, K. Eshraghian, *Principles of CMOS VLSI Design*, Addison-Wesley Publishing Company, Massachusetts, p. 111. (1993)
4. B.E Deal, A.S Grove, "General Relationship for the Thermal Oxidation of Silicon", Journal of Applied Physics, Vol.36 p.12. (1965)
5. Simmler, W. (2005), *"Silicon Compounds, Inorganic",* Ullmann's Encyclopedia of Industrial Chemistry, Weinheim: Wiley-VCH, doi:10.1002/14356007.a24_001
6. Tavares, Jason; Swanson, E.J.; Coulombe, S. *"Plasma Synthesis of Coated Metal Nanoparticles with Surface Properties Tailored for Dispersion".* Plasma Processes and Polymers. 5 (8): 759, (2008)
7. Annabelle Pratt, *"Overview of the Use of Copper Interconnects in the Semiconductor Industry",* Advanced Energy Industries, Whitepaper, (2004)
8. Tae-Seop Choi, Dennis W. Hess, "*Chemical Etching and Patterning of Copper, Silver, and Gold Films at Low Temperatures",* ECS Journal of Solid State Science and Technology, volume 4, issue 1, (2015)
9. Jens Lienig, "Electromigration and Its Impact on Physical Design in Future Technologies", ACM international Symposium on Physical Design (ISPD), p. 33-40, (2013)
10. Jain, K. *"Excimer Laser Lithography"*, SPIE Press, Bellingham, WA, (1990)

11. B. Grosman, S Lachman-Shalem, R Swissa and D.R. Lewin, *"Yield Enhancement in Photolithography through Model-based Process Control: Average Mode Control"* IEEE Transactions of Semiconductor Manufacturing, February (2004)

12. O D Trapp, L J Lopp, R Blanchard, *Semiconductor Technology Handbook*, Portola Valley. Calif. Technology Associates, p. 7, 10, (1993)

13. J.F Ziegler, *Handbook of Ion Implantation Technology*, IBM-Research Yorktown Heights, New York U.S.A, p. 122, (1992)

14. K.V. Ravi, *Imperfections and Impurities in Semiconductor Silicon*, A Wiley-Interscience Publication, p.12, (1981)

15. Chih-Yuan Lu, "*Process Limitation and Device Design Tradeoffs of Self-Aligned $TiSi_2$ Junction Formation in Submicrometer CMOS Devices",* IEEE Transactions on Electron Devices. Vol 38. NO.2, (1991)

16. John Foggiato, Woo Sik Yoo, *"Optimizing the Formation of Nickel Silicide"*, WaferMasters, Inc, Whitepaper.

17. Keith Orss, Phillip Sherbenou, "*Ni_5Pt and Ni_5Pt+Ti R_s as a Function of SCRTP03X Set-point and tc-Wafer Temperature"*, ON Semiconductor whitepaper, (2014)

18. Marc Heyns, Marc Meuris, Paul Mertens, Ultra Clean Processing of Silicon Surfaces*,* Acco Leuven/Amersfoort, p.13, (1996)

19. Kern, W. *"The Evolution of Silicon Wafer Cleaning Technology".* Journal of the Electrochemical Society. 137 (6): p. 1887–1892, (1990)

20. P. Grabiec, S. Pietruszko. *AN INFLUENCE OF DOPING ON THE KINETICS OF BPSG*

DEPOSITION PROCESS AND THE PROPERTIES OF THE DEPOSITED LAYER.

Journal de Physique Colloques, pp.C5-585-C5-593, (1989)

21. The international Technology Roadmap for semiconductor, p. 2, (2011)

22. Summer F.C. Tseng, Wei-Ting kary Chien, Bing-Chu Cai, "*Improvement of polysilicon hole induced gate oxide failure by silicon rich oxidation,"* Microelectronics Reliability 43, p. 713–724, (2003)

23. Jian-Ge Chen, ManLi Wu, *"Determination of PPB Anions in 49% Ultrapure Hydrofluric Acid"*, Ultra Clean Processing of Silicon Surface Symposium, p.37-39, (1996)

24. Kuiqing Peng, Juejun Hu, Yunjie Yan, *"Fabrication of Single-Crystalline Silicon Nanowires by Screatching a Silicon Surface with Catalytic Metal Particles"* Advanced Funcational Materials, p. 387-394. (2006)

25. D. Martin Knotter, Laurent Mouche, "*Silicon Surface Roughening in an Iron Contaminated SC1 Bath"*, Ultra Clean Processing of Silicon Surface Symposium, p.103-106, (1996)

26. O.D Trapp, R.A Blanchard, W. H Shepherd, *Semiconductor Technology Handbook*, Portola Valley, Calif, Technology Associates, p.55. (1980)

27. Ikhoon Shin, Jason Doub, Keith Mortensen, Raymond Lappan, *"Eliminating a Polysilicon Hole Defect Created During Oxide Removal"*, ASMC, (2011)

28. S. Abadli, F. Mansour*,"Grains-Growth and clustering Effects on Boron Diffusion in Polysilicon During Thermal Post-Implantation Annealing",* 16th International Conference on Microelectronics proceedings, IEEE, (2004)

29. Masashi Uematsu, *"simulation of high-concentration boron dissusion in silicon during post-implantation annealing*", Jpn. J. Appl. Phys, Vol. 38, p.3433. (1999)

5. Appendix: Acronyms

A		
	ARC	Anti-reflective Coating
B		
	BOE	Buffer Oxide Etch
	BPSG	Boro-Phospho-Silicate-Glass
C		
	CD	Critical Dimension
	CMOS	Complementary Metal Oxide Silicon
	CMP	Chemical Mechanical Polishing
	CVD	Chemical Vapor Deposition
D		
	D_0	Defect Density
	DOE	Design of Experiment
E		
	ESD	Electrostatic Discharge
F		
	FA	Failure Analysis
	FIB	Focused Ion Beam
	FOX	Field Oxide
G		
	GOI	Gate Oxide Integrity
H		
	HDP	High Density Plasma
I		
	IC	Integrated Circuit
L		
	LOCOS	Local Oxidation of Silicon
	LPCVD	Chemical Vapor Deposition
P		
	PECVD	Plasma Enhanced Chemical Vapor Deposition
R		

	RF	Radio Frequency
	RTA	Rapid Thermal Annealing
S		
	SEM	Scanning Electron Measurement
	SIMS	Secondary Ion Mass Spectrometry
	SPC	Statistical Process Control
	SRAM	Static Random Access Memory
	SRP	Series Resistance Probe
T		
	TDDB	Time Dependent Dielectric Breakdown
	TEOS	Tetraethoxysilane
U		
	UV	Ultraviolet
V		
	VBD	Voltage Breakdown of Dielectric