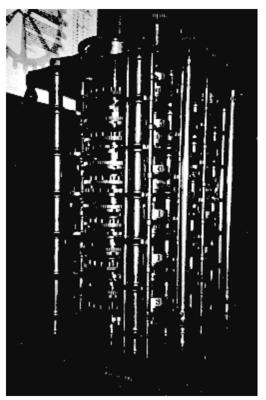
Historical Development

- To fully appreciate the computers of today, it is helpful to understand how things got the way they are
- The evolution of computing machinery has taken place over several centuries
- In modern times computer evolution is usually classified into four generations (the digital ages)

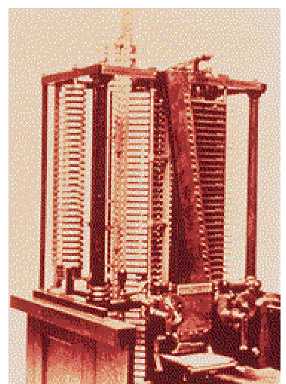
Before the digital age

- Age of the Mechanical Computers
- Charles Babbage invented a viable mechanical computer
- First one was called Difference Engine in early 1800 and was a special purpose calculator



difference engine

Babbage's second computer



analytical engine, 1834

- Analytical engine
  - general-purpose
  - used binary system
  - Arithmetic processing unit
  - punched cards as input
  - branch on result of previous instruction
  - Ada Lovelace (first programmer)
  - never quite completed

Review on Electric Circuit

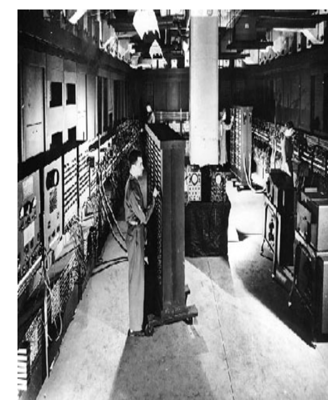
- An **electrical network/circuit** is an interconnection of electrical elements such as resistors, inductors, capacitors, transmission lines, voltage sources, current sources, and switches
- An electric circuit is formed when a conductive path (closed loop) is created to allow movement of charges
- This continuous movement of charges through the conductors of a circuit is called a ***current***, and it is often referred to in terms of "flow," just like the flow of a liquid through a hollow pipe.

Review on Electric Circuit (contd..)

- The force motivating electrons to "flow" in a circuit is called ***voltage***. Voltage is a specific measure of potential energy that is always relative between two points.
  - how much ***potential*** energy exists to move charges from one particular point in that circuit to another particular point
- Free electrons tend to move through conductors with some degree of friction, or opposition to motion. This opposition to motion is more properly called ***resistance***
- The ***amount of current*** in a circuit depends on the amount of voltage available to motivate the electrons, and also the amount of resistance in the circuit to oppose electron flow

First Generation: Vacuum Tube Computers (1945 - 1953)

- ENIAC
  - Electronic Numerical Integrator and Computer
  - The ENIAC was the first *general-purpose electronic* computer
  - Created by John Mauchly and J. Presper Eckert at University of Pennsylvania, 1946
  - Made of vacuum tubes
    - 17468 vacuum tubes, occupied 1800 sq. feet, weighted 30 tons and consumed 174 kilowatts of power



Vacuum Tube

- A device that controls the flow of electrons in evacuated space
- Discovery by Thomas Edison (1883) – birth of vacuum tube
  - Conducted series of electrical experiments in order prevent the light bulb filament from burning away and blackened
  - How the light bulb work?
    - Passing electricity through the filament caused it to heat up, and radiate light (release of photons)

Edison's Discoveries

- By placing filament mounted in a glass bulb which was vacuumized prevented the filament from oxidizing and burning up
- Edison also detected electrons flowing from the lighted filament to a metal plate (positive charged) mounted inside the bulb
- This became known as the ***Edison Effect*** (a.k.a thermionic emission i.e. emission of charged particles called thermions (+/- charges) from a charged metal

Edison Effect

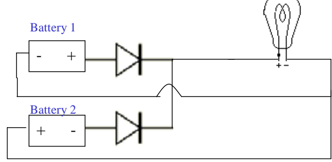
Application of Vacuum Tube

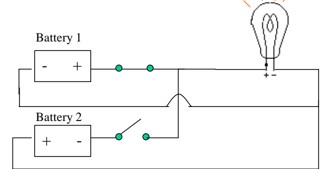
- The current flow was when the plate was connected to the positive end of the filament, but not when the plate was connected to the negative side
- John Fleming studied Edison's effect (around 1905)
  - to detect radio waves and to convert them to electricity
  - Patented this device as ***diode***

- **Diode** is a component that restricts the direction of movement of charge carriers. Essentially, it allows an electric current to flow in one direction, but blocks it in the opposite direction.
  - electrons flow within the tube from the negatively charged (*cathode)* to the positively charged (*anode). Electric current flows in the opposite direction*

Uses of Diodes

- Used as rectifiers, and switches
- Use as a switch
  - Constructing OR function out of diodes

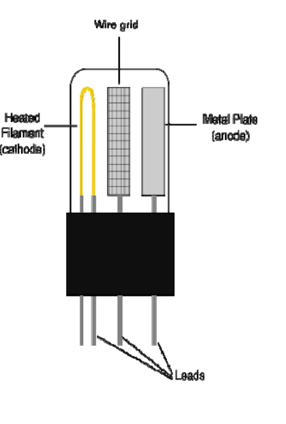




Battery 1 orientation makes the diode conduct (acts as switch closed). Battery 2 orientation does not make the diode conduct (switch is opened). However Battery 1 orientation forms a path in the circuit such that current flows through the circuit and bulb is lit. There are 4 different combinations of battery orientation. Testing these orientation, you will soon realize that this is an OR function i.e. either diodes conducts then the bulb lights up.

Triode

- Same function as Diode
- Additional element added to the diode called control grid
- The control grid can reduce or prevent electron flow from the cathode to the anode of the a diode
- The control grid is charged negatively to repel electrons from flowing from filament to plate
  - The larger the charge on the grid, the smaller the electron flow to plate



Drawback of the Vacuum Tube

- Occupied a lot of space
  - ENIAC occupied 1800 sq feet

- They use enormous amounts of electrical power to heat the cathode
  - The heat dissipation can cause melt down of the device
  - Overcome the problem large cooling systems are employed e.g. ENIAC consumed 174 kilowatts of power

Second Generation:Transistorized Computers (1954 - 1965)

- Invented at Bell Labs in 1948
  - Inventors: John Bardeen, Walter Brattain, and William Shockley (Nobel prize, 1956)
- A **transistor** is a semiconductor device that is used as amplifier or switch
  - Equivalent to Triode of the First generation

Semiconductor

- Semiconductor material are in between conductors (like metals) and insulators (like glass)
  - In pure form, it is non-conducting
  - E.g. silicon crystal - has 4 electrons in its outer orbital, which form perfect covalent bonds with four neighboring atoms, leaving no electrons to conduct electric current
  - You can change the behavior of silicon and turn it into a conductor by **doping** it. In doping, you mix a small amount of an **impurity** into the silicon crystal

Semiconductor Doping

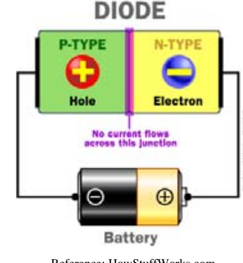
- **N-type**
  - In N-type doping, phosphorus or arsenic is added to the silicon in small quantities
  - Phosphorus and arsenic each have 5 outer electrons
  - The 5th electron has nothing to bond to, so it's free to move around
  - The created free electrons allow an electric current to flow through the silicon
  - Electrons have a negative charge, hence the name N-type

Semiconductor Doping (contd..)

- P-type
  - ➢ In P-type doping, boron or gallium is the dopant
  - ➢ Boron and gallium each have only three outer electrons
  - ➢ When mixed with silicon, they form "holes" in the where a silicon electron has nothing to bond to
  - ➢ The absence of an electron creates the effect of a positive charge, hence the name P-type
  - ➢ Holes can conduct current. A hole happily accepts an electron from a neighbor, moving the hole over a space

Semiconductor Diode

- No current flows across the junction because the holes and the electrons are each moving in the wrong direction
- But If we reverse the direction then….



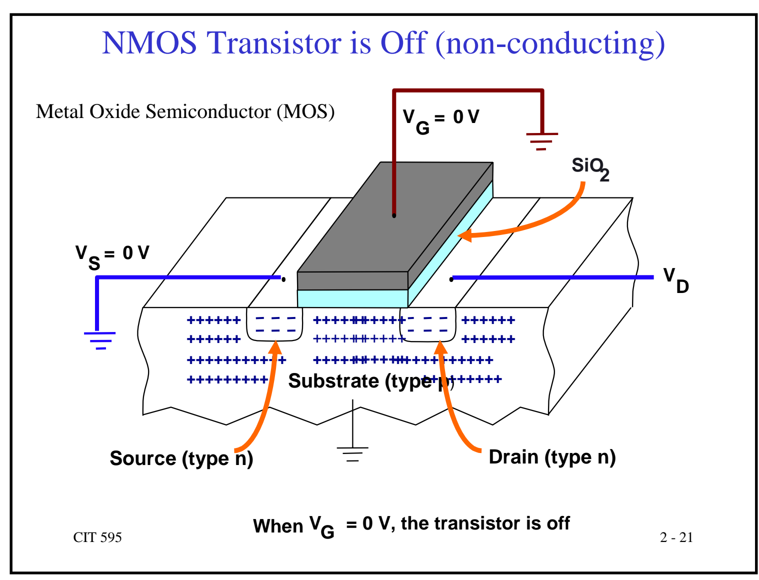
Reference: HowStuffWorks.com

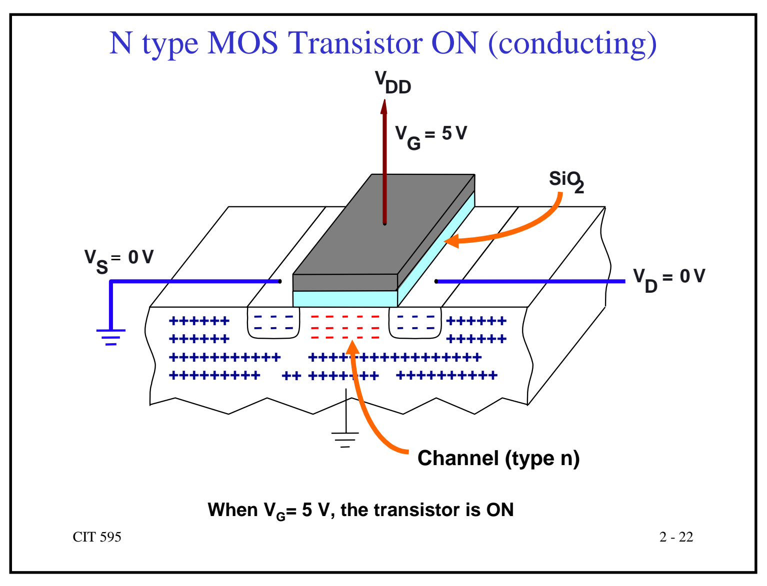
Semiconductor Diode (contd..)

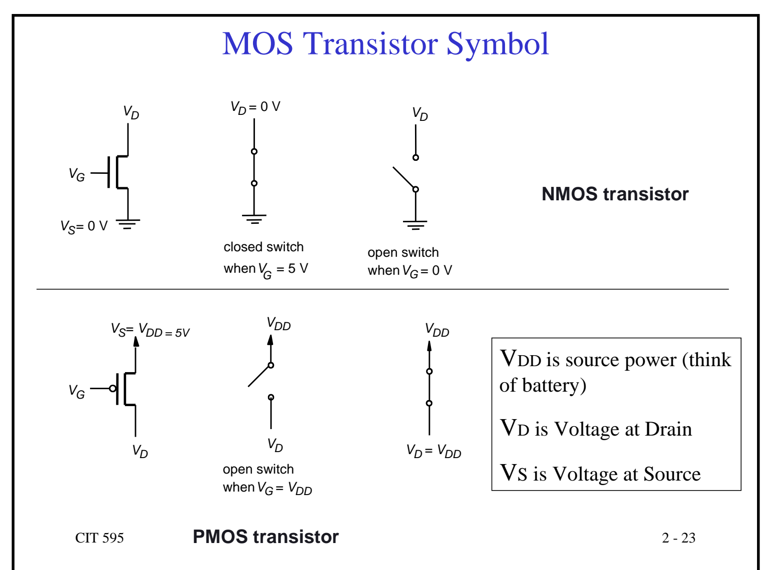
- If you **flip the battery around**, the diode conducts electricity
  - ➢ The free electrons in the N-type silicon are repelled by the negative terminal of the battery. The holes in the P-type silicon are repelled by the positive terminal
  - ➢ At the **junction** between the N-type and P-type silicon, holes and free electrons meet. The electrons fill the holes. Those holes and free electrons cease to exist, and new holes and electrons spring up to take their place. The effect is that **current flows** through the junction
    - ▪ Basically enough pressure (potential built up) so that the charges move/current flows

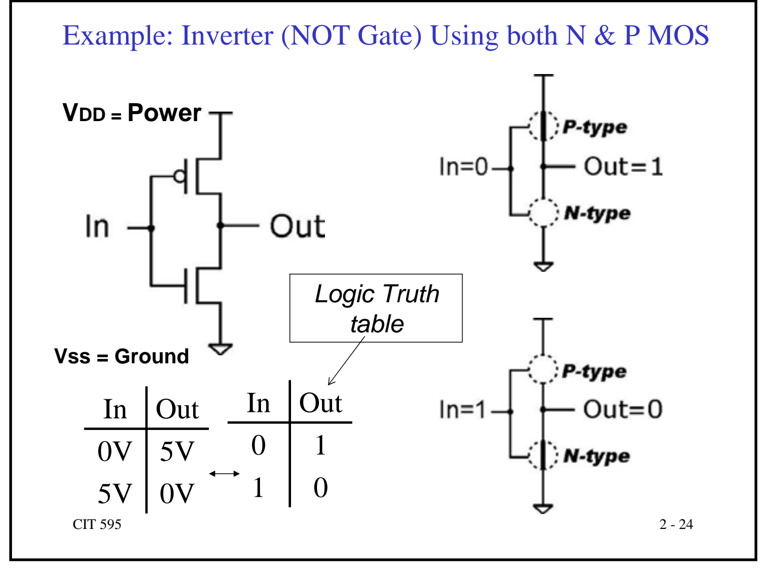
Semiconductor Transistor

- A **transistor** is created by using **three layers** instead of two layers like in the diode
  - ➢ You can create either an NPN or a PNP sandwich.
- A transistor looks like two diodes back-to-back. You'd imagine that no current could flow through a transistor because back-to-back diodes would block current both ways
- However, when you apply a small current to the **center layer** of the sandwich, a much larger current can flow through the sandwich as a whole. This gives a transistor its **switching** behavior.
  - ➢ A small current can turn a larger current on and off


NMOS Transistor is Off (non-conducting)
Metal Oxide Semiconductor (MOS)
VG = 0 V
SiO2
VS = 0 V
VD
Substrate (type p)
Source (type n)
Drain (type n)
When VG = 0 V, the transistor is off
CIT 595
2 - 21


N type MOS Transistor ON (conducting)
VDD
VG = 5 V
SiO2
VS = 0 V
VD = 0 V
Channel (type n)
When VG= 5 V, the transistor is ON
CIT 595
2 - 22


MOS Transistor Symbol
VD
VG
VS= 0 V
VD = 0 V
closed switch
when VG = 5 V
VD
open switch
when VG = 0 V
NMOS transistor
VS= VDD = 5V
VG
VD
VDD
VD
open switch
when VG = VDD
VDD
VD = VDD
VDD is source power (think of battery)
VD is Voltage at Drain
VS is Voltage at Source
CIT 595
PMOS transistor
2 - 23


Example: Inverter (NOT Gate) Using both N & P MOS
VDD = Power
In
Out
Vss = Ground
P-type
In=0
Out=1
N-type
Logic Truth table
In | Out
0V | 5V
5V | 0V
In | Out
0 | 1
1 | 0
P-type
In=1
Out=0
N-type
CIT 595
2 - 24

Previous Example: CMOS Inverter

- Complementary MOS (CMOS)
- Uses both n-type and p-type MOS transistors
  - p-type
    - Attached to POWER (high voltage)
    - Pulls output voltage UP when input is zero
  - n-type
    - Attached to GROUND (low voltage)
    - Pulls output voltage DOWN when input is one

Basic Functions

- In the last example we saw how inverter function is implemented using transistors
- Similarly other basic functions such as AND, OR, NAND, NOR etc can be built
- In Chapter 3 we will study that from basic functions we can make more complex functions such as XOR, MUX, Adders, Memory

Advantage of Transistor over Vacuum Tube

- Smaller size
- Highly automated manufacturing
- Lower cost (in volume production)
- No warm-up period (most vacuum tubes need 10 to 60 seconds to function correctly)
- Lower power dissipation
- Higher reliability and greater physical ruggedness

Third Generation: Integrated Circuit Computer (1965 - 1980)

- Integrated Circuits or chips contain dozens transistors on single silicon chip
- Made computers faster, smaller, cheaper
- Computer manufacturers of this era were characterized as IBM and the BUNCH (Burroughs, Unisys, NCR, Control Data, and Honeywell)

Fourth Generation: VLSI Computers (1980 – Current)

- Very Large Scale Integrated circuits (VLSI) have more than 10,000 components per chip
- Enabled the creation of microprocessors
  - The first was the 4-bit Intel 4004
  - Later versions, such as the 8080, 8086, and 8088 spawned the idea of "personal computing."
  - Intel 80 series had around 29,000 transistors

Integrated Circuits (IC) Examples

7404 Chip

(6 Inverters on one chip)

V_{DD}

Gnd

Millions of Transistors on silicon Wafer (VLSI)

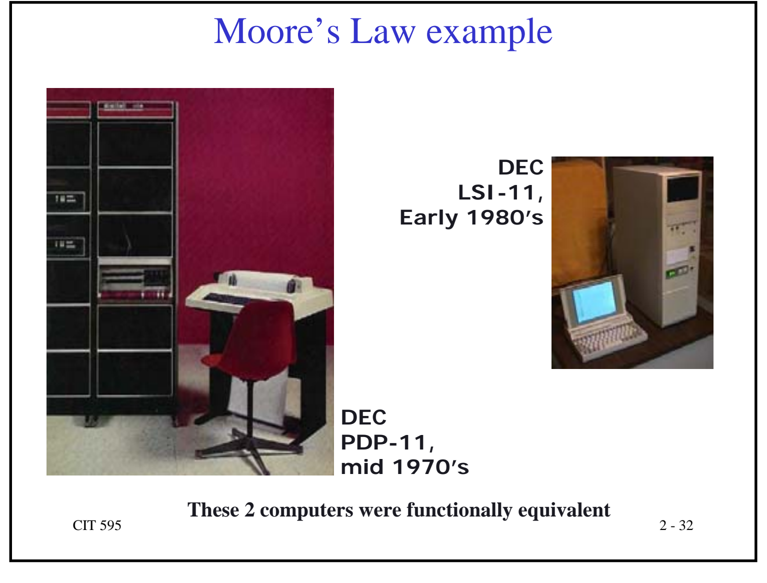
intel pentium 4

Moore's Law (Rule of Thumb)

- Gordon Moore, Intel founder said

  "The density of transistors in an integrated circuit will double every year."

- Contemporary version: "The density of silicon chips doubles every 18 months."
  - also, prices decline
  - first described in 1965
  - experts predict this trend might continue until ~2020
  - limited when size reaches molecular level



Transistor Density Increase

- **Intel Pentium II processor(1997)**
  - ➢7 million transistors
- **Intel Pentium III processor(1999)**
  - ➢28 million transistors
- **Intel Pentium 4 processor(2000)**
  - ➢42 million transistors

Corollary to Moore's Law

- Rock's Law (Intel financier)
  - ➢"The cost of capital equipment to build semiconductors will double every four years."
  - ➢In 1968, a new chip plant cost about $12,000.
  - ➢In late 1990s the cost had come to $12 million

Need alternative means to build chips**…may be quantum computing???**

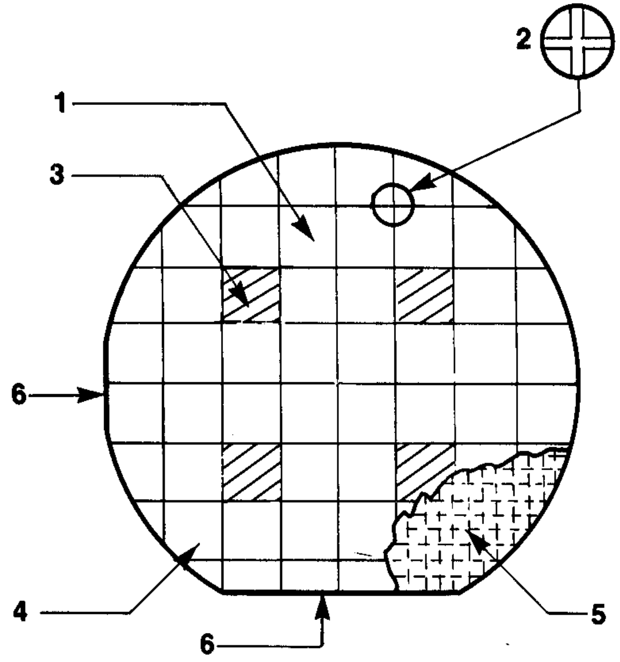
CMOS Technology

CMOS Technology

- Basic Fabrication Operations
- Steps for Fabricating a NMOS Transistor
- LOCOS Process
- n-Well CMOS Technology
- Layout Design Rules
- CMOS Inverter Layout Design
- Circuit Extraction, Electrical Process Parameters
- Layout Tool Demonstration
- Appendix: MOSIS, EUROPRACTICE

Wafer Terminology

1. Chip = Die = Microchip = Bar
2. Scribe Lines
3. Engineering Test Die
4. Edge Die
5. Crystal Planes
6. Wafer Flats



Basic Wafer Fabrication Operations

The number of steps in IC fabrication flow depends upon the technology process and the complexity of the circuit

Example:

CMOS n-Well process - 30 major steps, and each major step may involve up to 15 substeps

Only **three basic operations** are performed on the wafer:

- Layering
- Patterning
- Doping

Layering

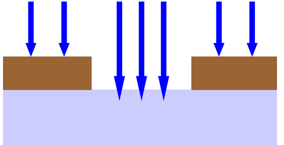
Grow or deposit thin layers of different materials on the wafer surface

Layers	Technique			
	Thermal oxidation	**ChemicalVapor Deposition (CVD)**	**Evaporation**	**Sputtering**
Insulators	Silicon Dioxide (SiO2)	Silicon Dioxide (SiO2) Silicon Nitrides (Si3N4)		Silicon Dioxide (SiO2) Silicon Monoxide (SiO)
Semiconductors		Epitaxial Silicon Poly Silicon		
Conductors		Dopedpolysilicon Metals Al/Si Alloys Silicides	Metals Alloys	Metals Alloys

Layering - Thermal Oxidation

SiO_2 functions:

Surface passivation


Diffusion barrier

Field oxide

MOS Gate oxide

Natural oxide: silicon will readily grow an oxide (5-10nm) if exposed to oxygen in the air!

The range for useful oxide thickness: 25nm (MOS gates) - 1500nm (field oxide)

Dry oxidation

$Si + O_2 \rightarrow SiO_2$ (900-1200°C)

700nm oxide: 10hours (1200°C)

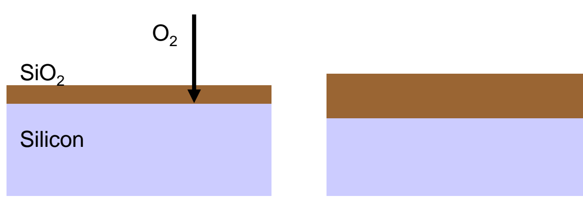
Good oxide quality: gate oxide

Wet oxidation (water vapor or steam)

$Si + H_2O \rightarrow SiO_2 + 2H_2$ (900-1200°C)

700nm oxide: 0.65hours (1200°C)

Poor oxide quality: field oxide



Layering - Chemical Vapor Deposition (CVD)

Deposited materials:

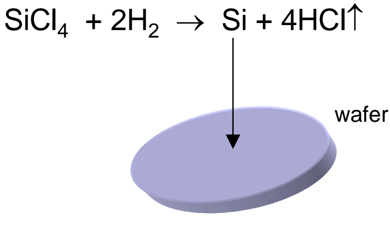
- **Insulators & Dielectrics**: SiO_2, Si_3N_4, Phosphorus Silicate Glass (PSG), Doped Oxide
- **Semiconductors**: Si
- **Conductors**: Al, Cu, Ni, Au, Pt, Ti, W, Mo, Cr, Silicides (WSi_2, $MoSi_2$), doped polysilicon

Basic CVD processing:

- a gas containing an atom(s) of the material to be deposited reacts with another gas liberating the desired material
- the freed material (atom or molecular form) "deposits" on the substrate
- the unwanted products of the chemical reaction leave the reaction chamber

Example: CVD of silicon from silicon tetrachloride

$$SiCl_4 + 2H_2 \rightarrow Si + 4HCl\uparrow$$

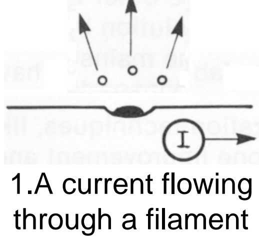


Layering - Evaporation

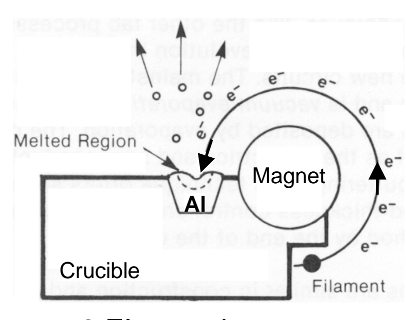
Used to deposit **conductive layers** (**metallization**): Al, Al/Si, Al/Cu, Au, Mo, Pt

When temperature is raised high enough, atoms of solid material (Al) will melt and "evaporate" into the atmosphere and deposit on to the wafer

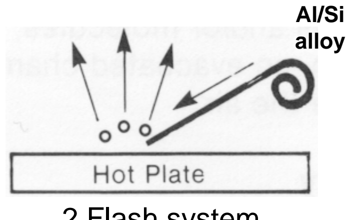
External energy needed to evaporate the metal are provided by:



1.A current flowing through a filament

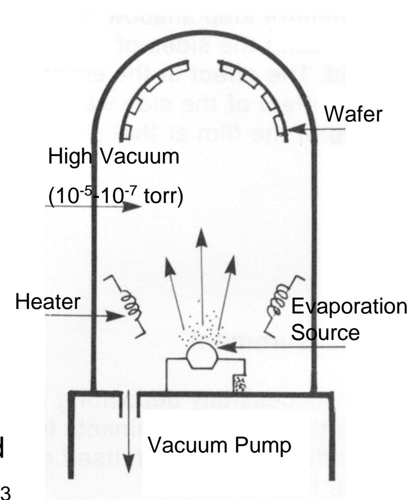


3.Electron beam



2.Flash system

The evaporation take place into an evacuated chamber; otherwise Al would combine with oxygen in air to form Al_2O_3



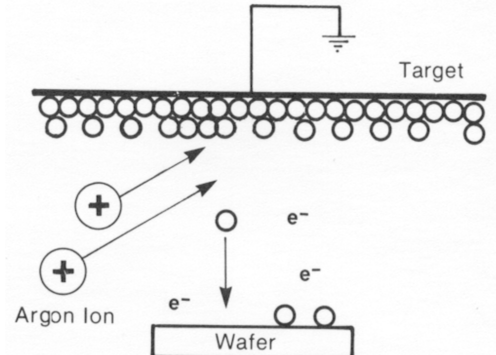
Layering - Sputtering

Used to deposit **thin metal/alloys** films and **insulators**: Al, Ti, Mo, Al/Si, Al/Cu, SiO_2

Sputtering process:

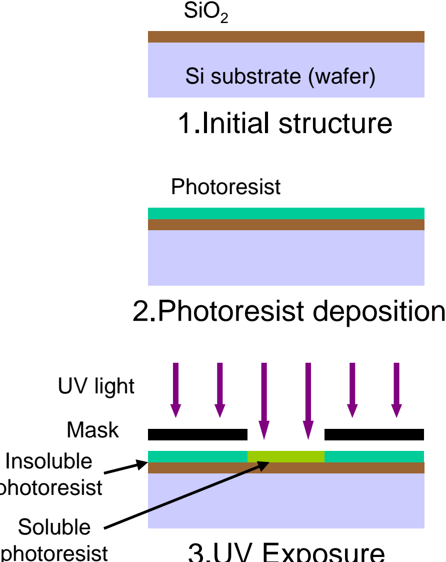
- ionized argon atoms (+) are introduced into an evacuated chamber
- the target (Al) is maintained at negative potential
- the argon ions accelerated towards the negative charge
- following the impact some of the target material atoms tear off
- the liberated material settles on everything in the chamber, including the wafers

The material to be sputtered does not have to be heated



Patterning

- Patterning = Lithography = Masking
- Selective removal of the top layer(s) on the wafers
- Ex.: Process steps required for patterning SiO_2



4.Soluble photoresist etching

Chemical/Dry etch

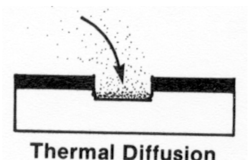
5.SiO_2 etching

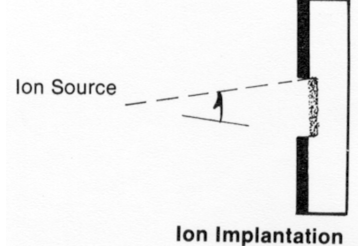
5.SiO_2 etching (end)

6.Photoresist etching

Doping

- Change conductivity type and resistivity on selected regions of wafer
- Doping takes place to the wafer through the holes patterned in the surface layer
- **Two techniques** are used:
    - Thermal diffusion
    - Ion implantation

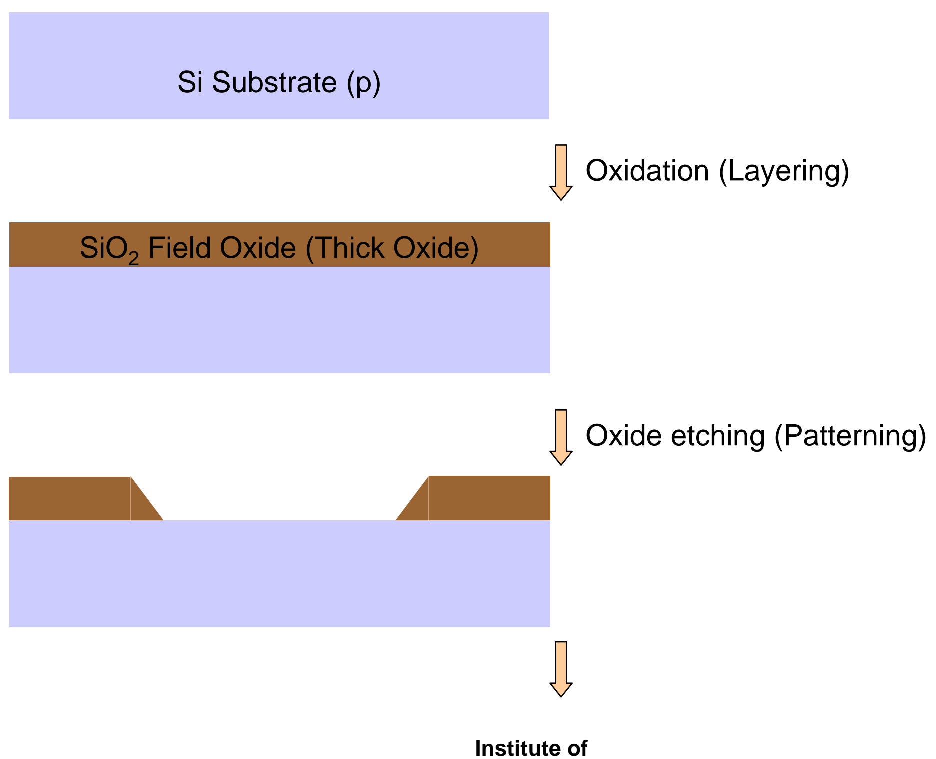




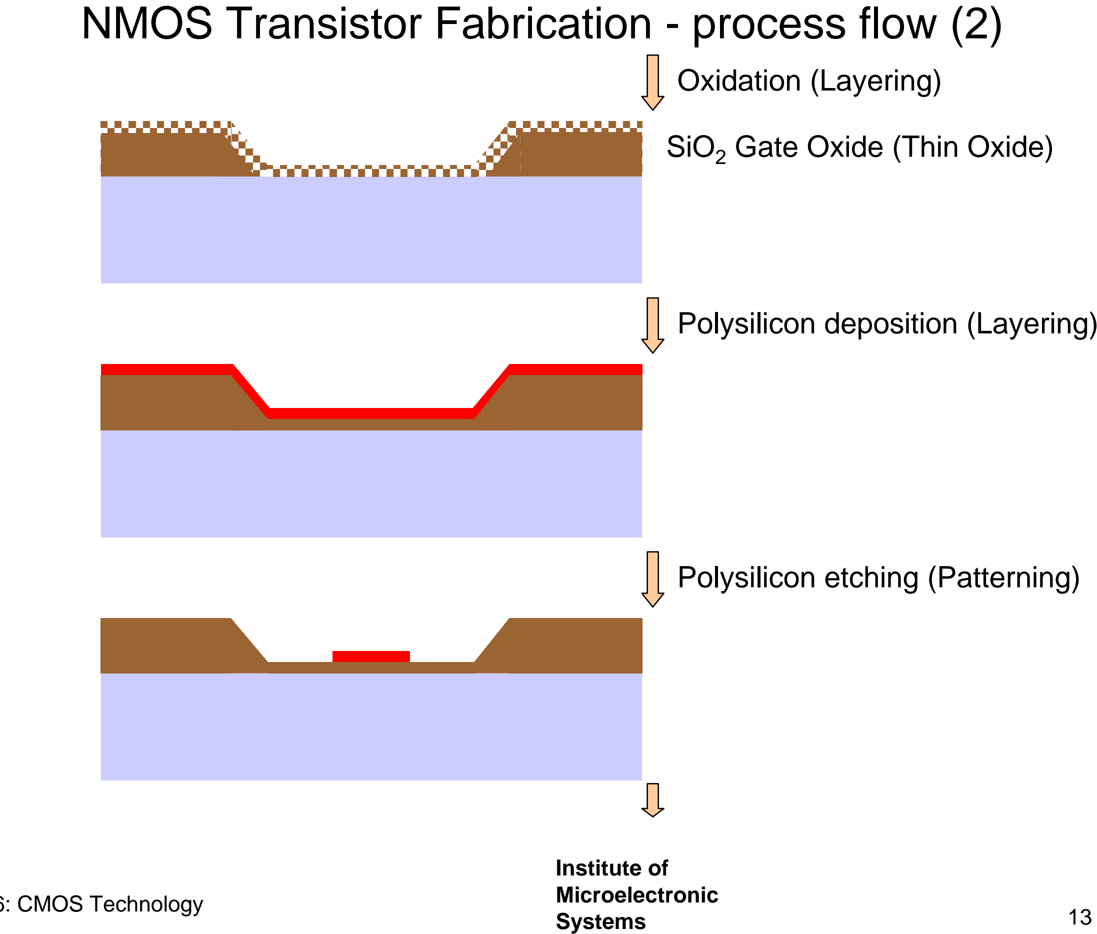
Thermal diffusion:
- heat the wafer to the vicinity of 1000°C
- expose the wafer to vapors containing the desired dopant
- the dopant atoms diffuse into the wafer surface creating a p/n region

Ion implantation:
- room temperature
- dopant atoms are accelerated to a high speed and “shot” into the wafer surface
- an annealing (heating) step is necessary to reorder the crystal structure damaged by implant

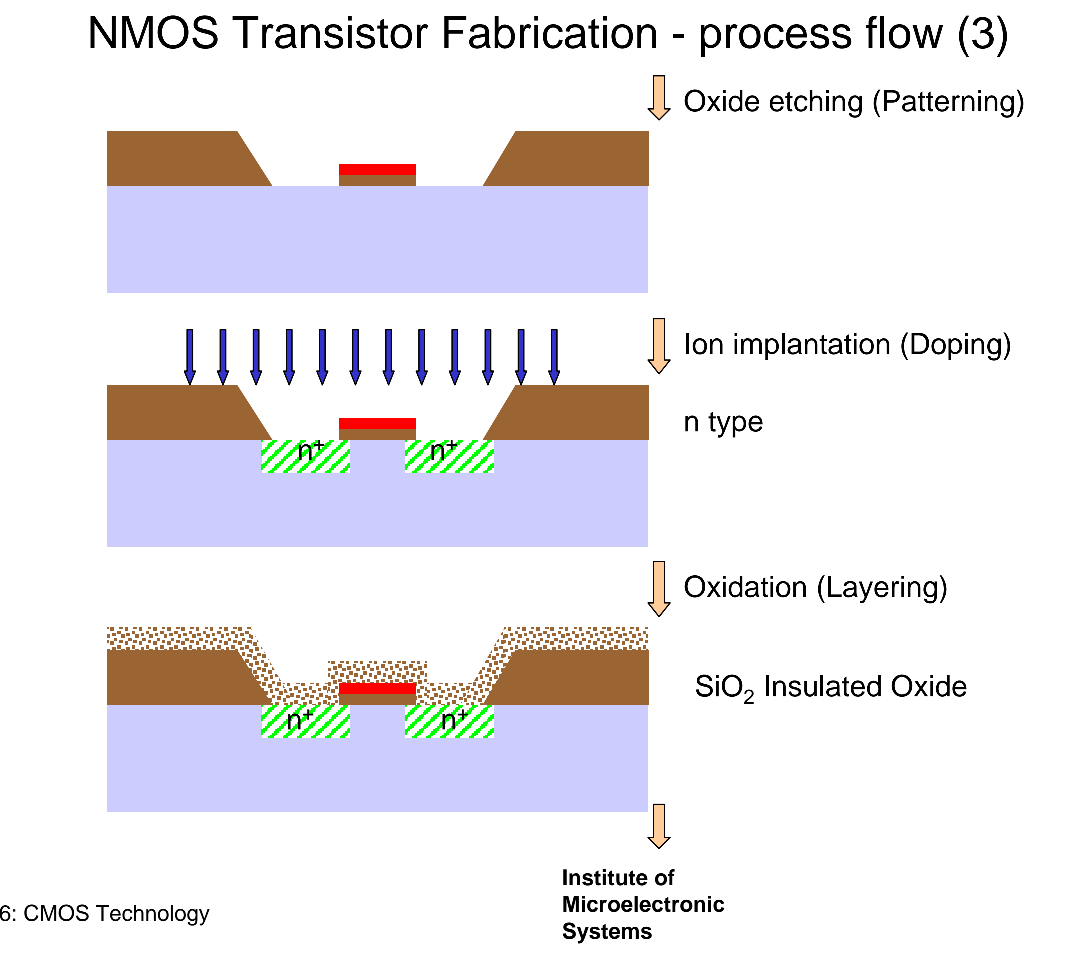
NMOS Transistor Fabrication - process flow (1)



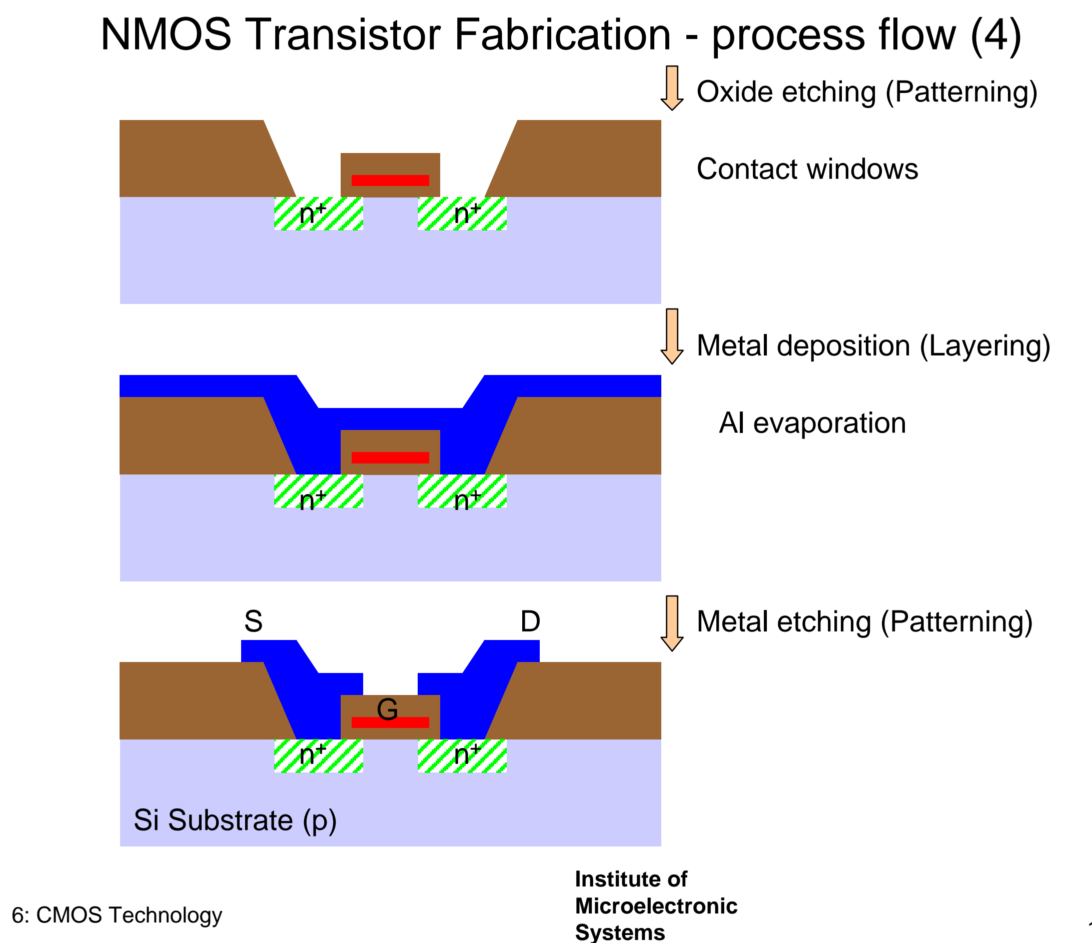
NMOS Transistor Fabrication - process flow (2)



NMOS Transistor Fabrication - process flow (3)



NMOS Transistor Fabrication - process flow (4)



Device Isolation Techniques

MOS transistors must be **electrically isolated** from each other in order to:

- prevent unwanted conduction paths between devices
- avoid creation of inversion layers outside the channel regions
- reduce the leakage currents

Each device is created in dedicated regions - **active areas**

Each active area is surrounded by a field oxide barrier using few techniques:

A) **Etched field-oxide isolation**

1) grow a field oxide over the entire surface of the chip

2) pattern the oxide and define active areas

Drawbacks: -large oxide steps at the boundaries between active areas and field regions!

-cracking of polysilicon/metal subsequent deposited layers!

Not used!

B) **Local Oxidation of Silicon (LOCOS)**

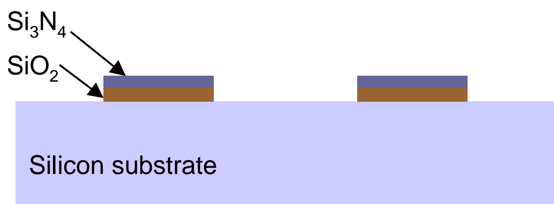
Local Oxidation of Silicon (LOCOS) (1)

More planar surface topology

Selectively growing the field oxide in certain regions - process flow:

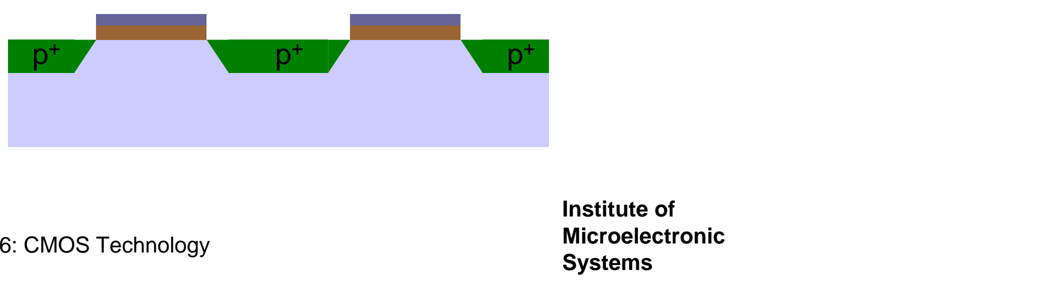
1) grow a thin pad oxide (SiO_2) on the silicon surface

2) define active area : deposition and patterning a silicon nitride (Si_3N_4) layer



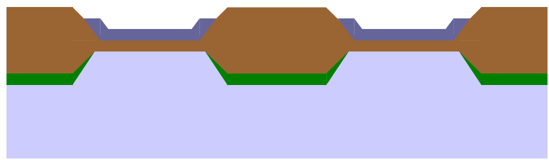
The thin pad oxide - protect the silicon surface from stress caused by nitride

3) channel stop implant: p-type regions that surround the transistors



Local Oxidation of Silicon (LOCOS) (2)

4) Grow a thick field oxide

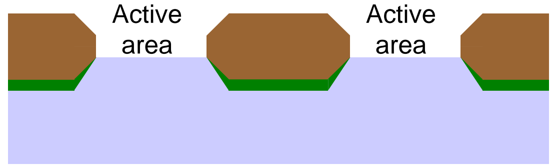


Field oxide is partially recessed into the surface (oxidation consume some of the silicon)

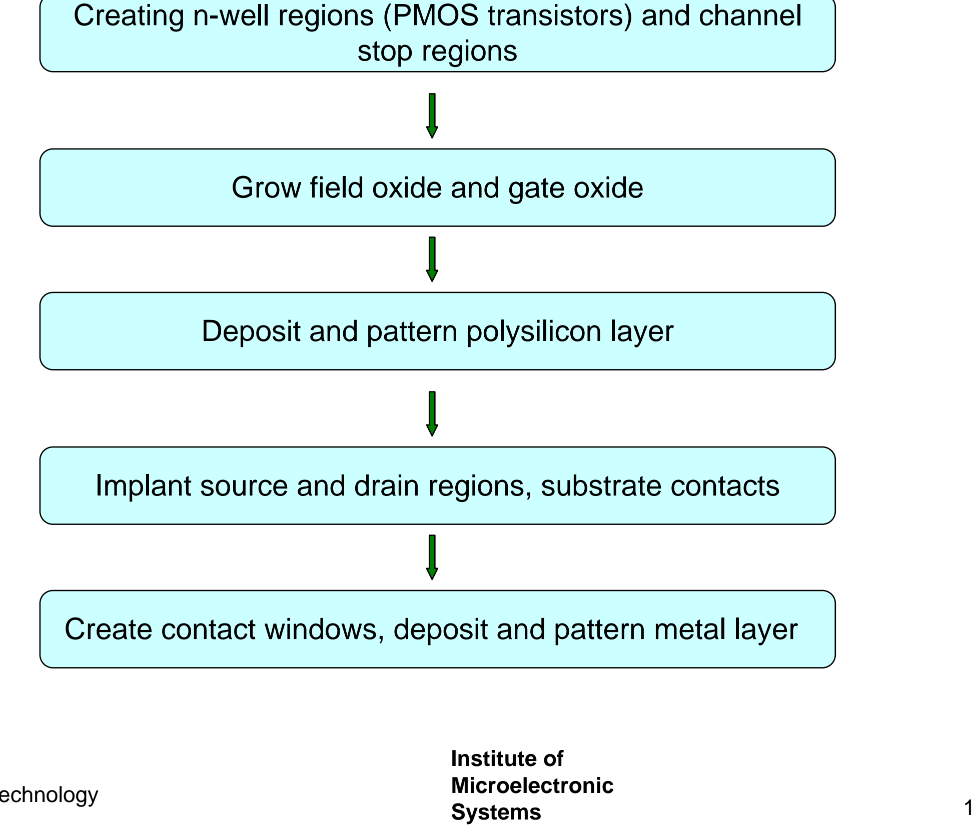
Field oxides forms a **lateral extension** under the nitride layer - **bird`s beak region**

Bird's beak region **limits device scaling** and **device density** in VLSI circuits!

5) Etch the nitride layer and the thin oxide pad layer



n-Well CMOS Technology - simplified process sequence

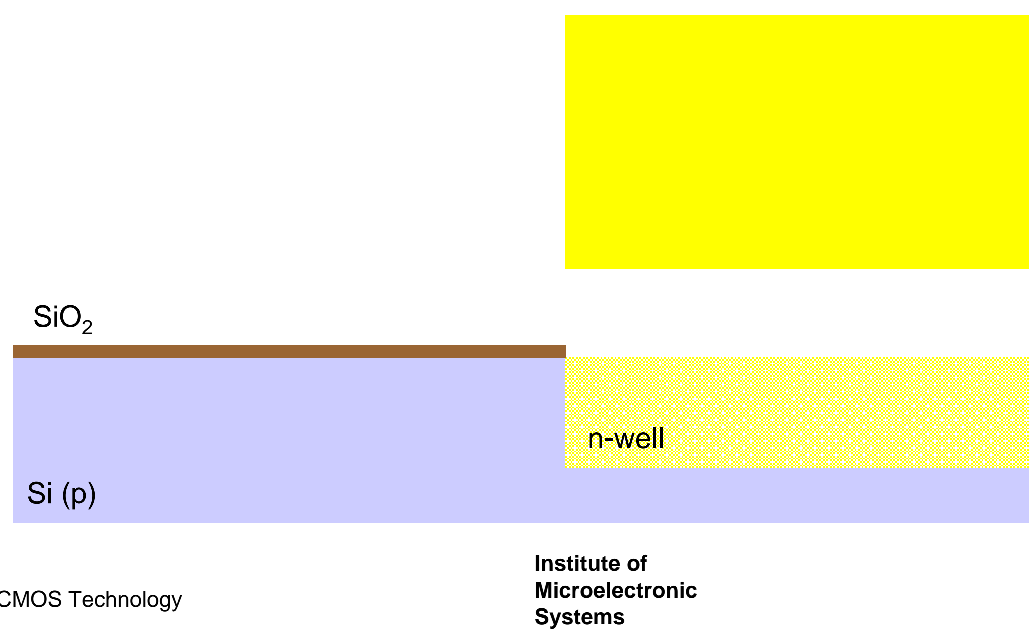


n-Well CMOS Technology - Inverter Example

- Process starts with a moderately doped (10^{15} cm^{-3}) p-type substrate (wafer)
- An initial oxide layer is grown on the entire surface (barrier oxide)

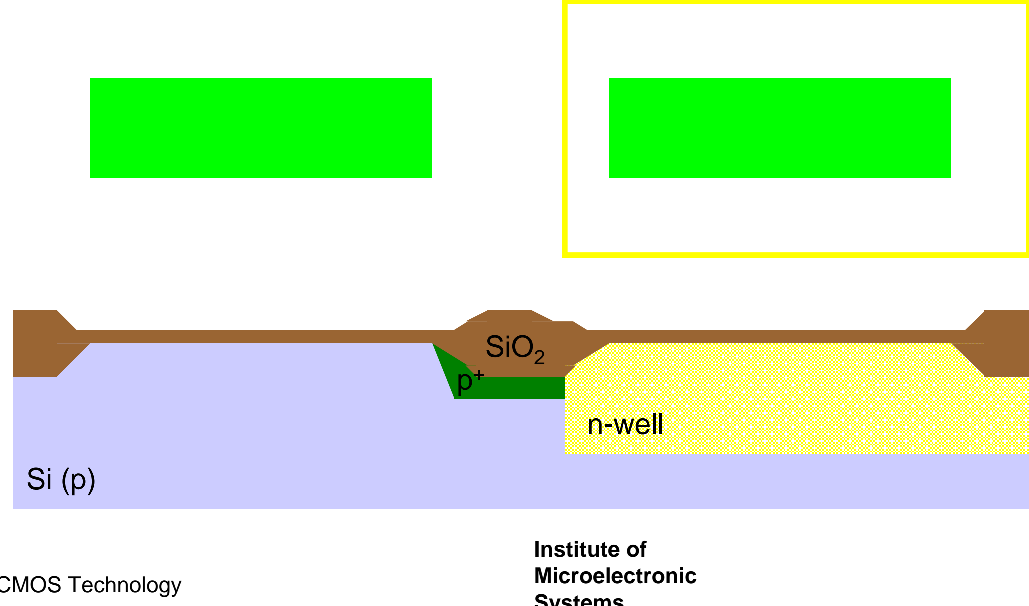
1. **n-Well mask** - defines the n-Well regions

- Pattern the oxide
- Implant n-type impurity atoms (phosphorus) - 10^{16}cm^{-3}
- Drive-in the impurities (vertical but also lateral redistribution - **limits the density**)



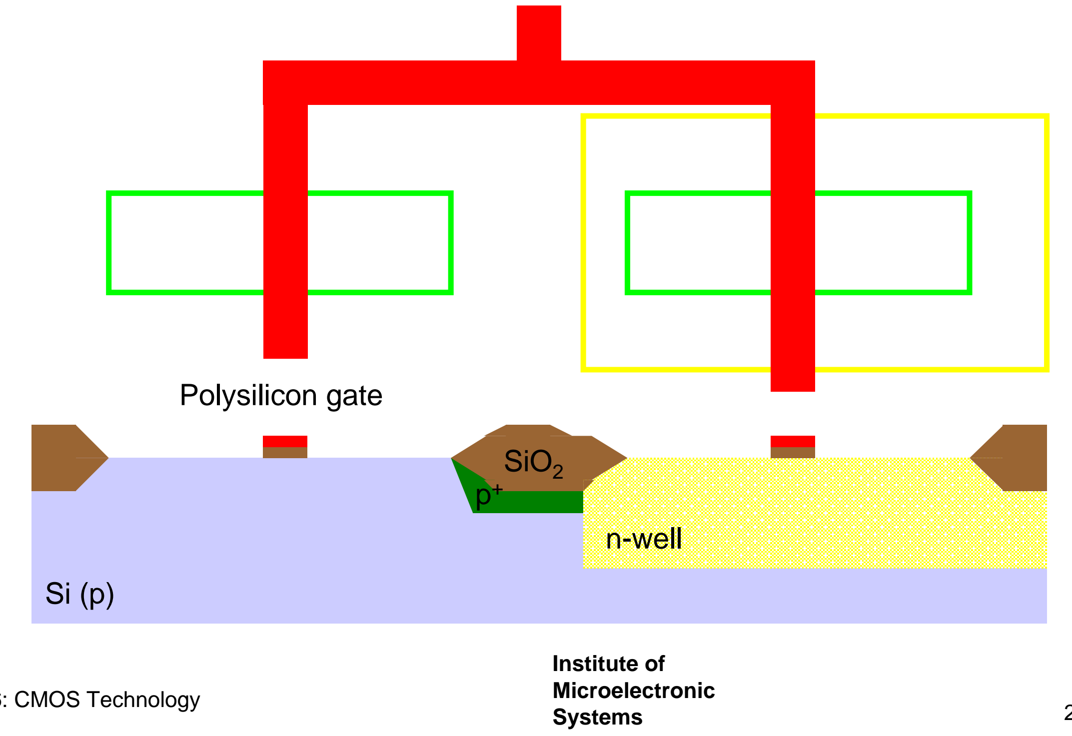
2. **Active area mask** - define the regions in which MOS devices will be created

- LOCOS process to isolate NMOS and PMOS transistors
  - lateral penetration of bird's beak region ~ oxide thickness
  - channel stop p^+ implants (boron)
- Grow gate oxide (dry oxidation) - only in the open area of active region



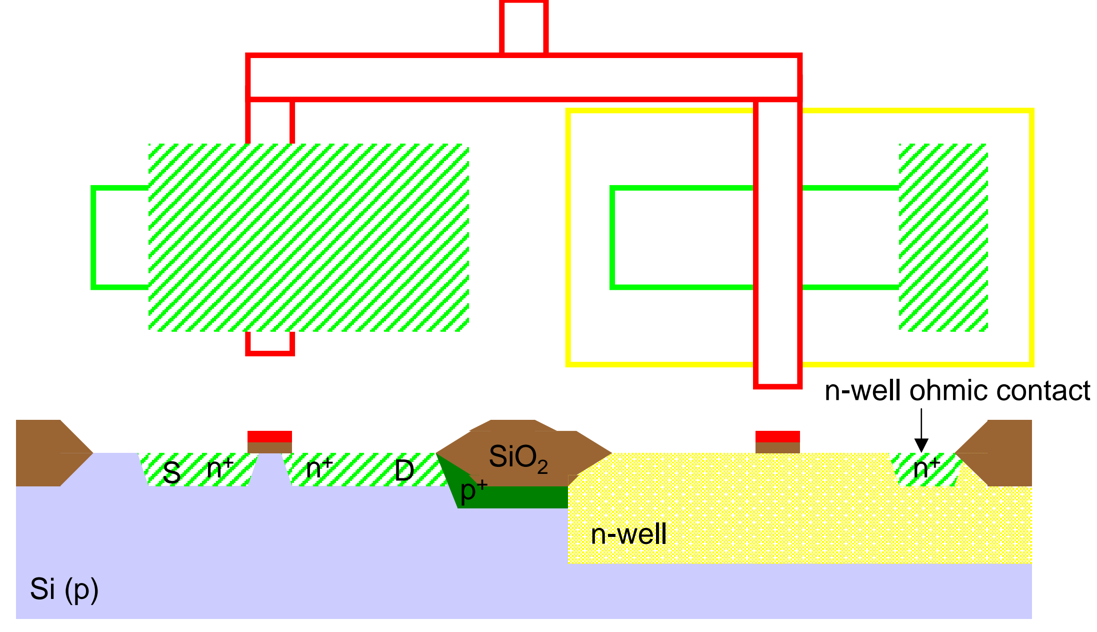
3. **Polysilicon mask** - define the gates of the MOS transistors

- Polysilicon is deposited over the entire wafer (CVD process) and doped (typically n-type)
- Pattern the polysilicon in the dry (plasma) etching process
- Etch the gate oxide



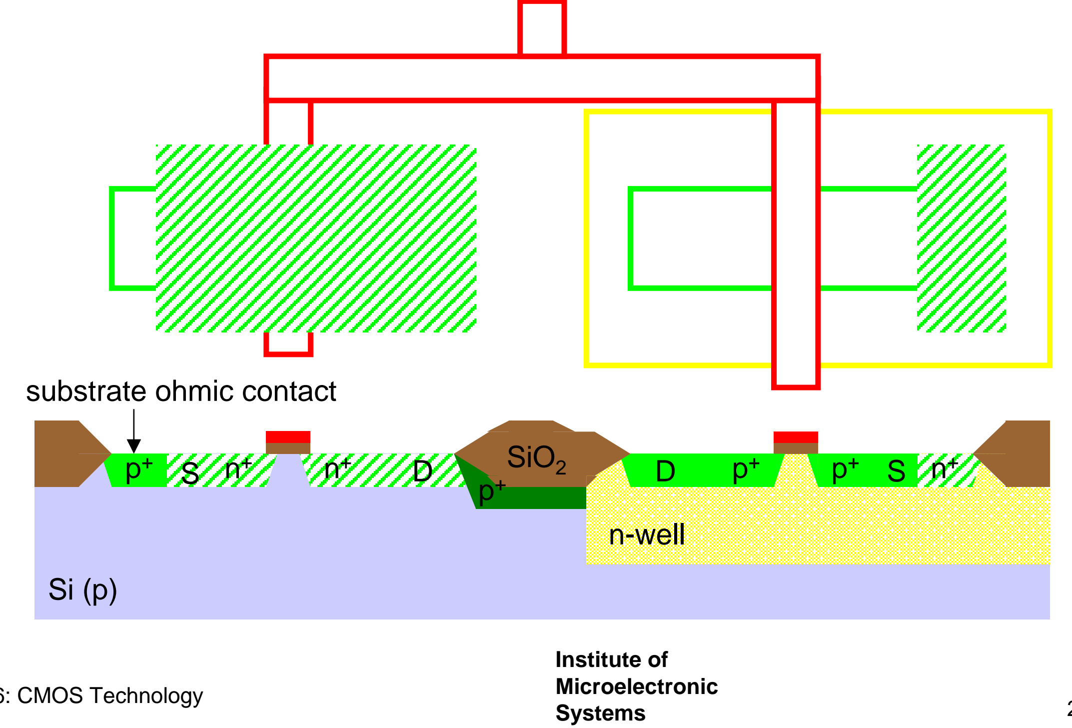
4. **n-Select mask** - define the n^+ source/drain regions of NMOS transistors

- Define an ohmic contact to the n-well
- Implant n-type impurity atoms (arsenic)
- Polisilicon layer protects transistor channel regions from the arsenic dopant

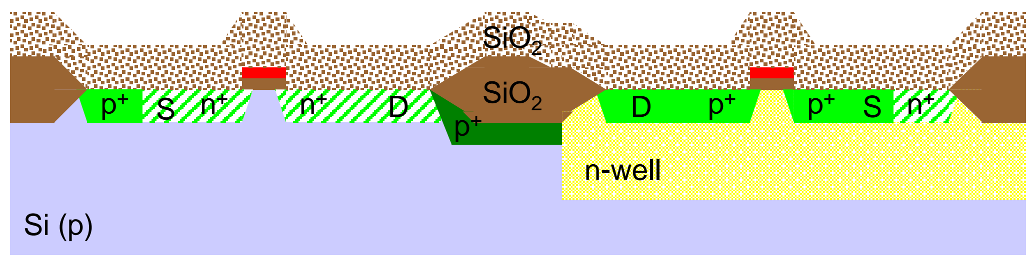


5. **Complement of the n-select mask** - define the p^+ source/drain regions of PMOS transistors

- Define the ohmic contacts to the substrate
- Implant p-type impurity atoms (boron)
- Polisilicon layer protects transistor channel regions from the boron dopant

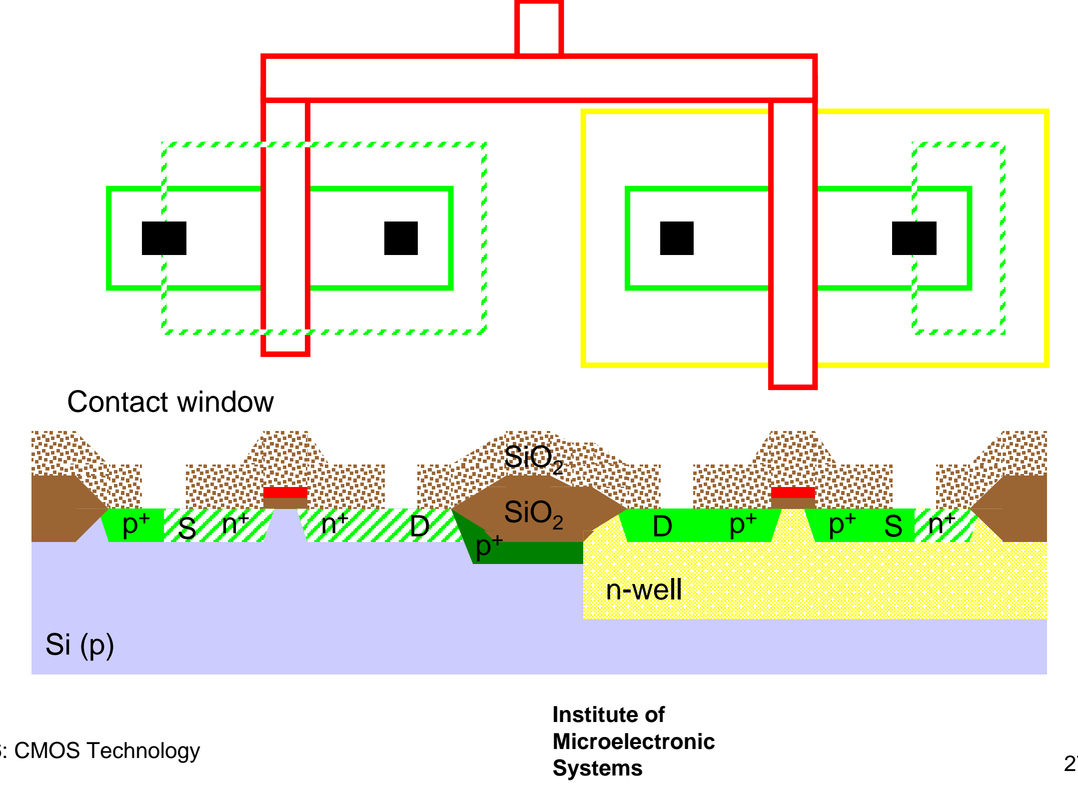


- In the n-well two p^+ and one n^+ regions are created
- After source/drain implantation a short thermal process is performed (annealing):
  - moderate temperature
  - drive the impurities deeper into the substrate
  - repair some of the crystal structure damage
  - lateral diffusion under the gate: overlap capacitances
- Next the SiO_2 **insulated layer** is deposited over the entire wafer area using a CVD technique
- The surface becomes nonplanar: impact on the metal deposition step



6. **Contact mask** - define the contact cuts in the insulating layer

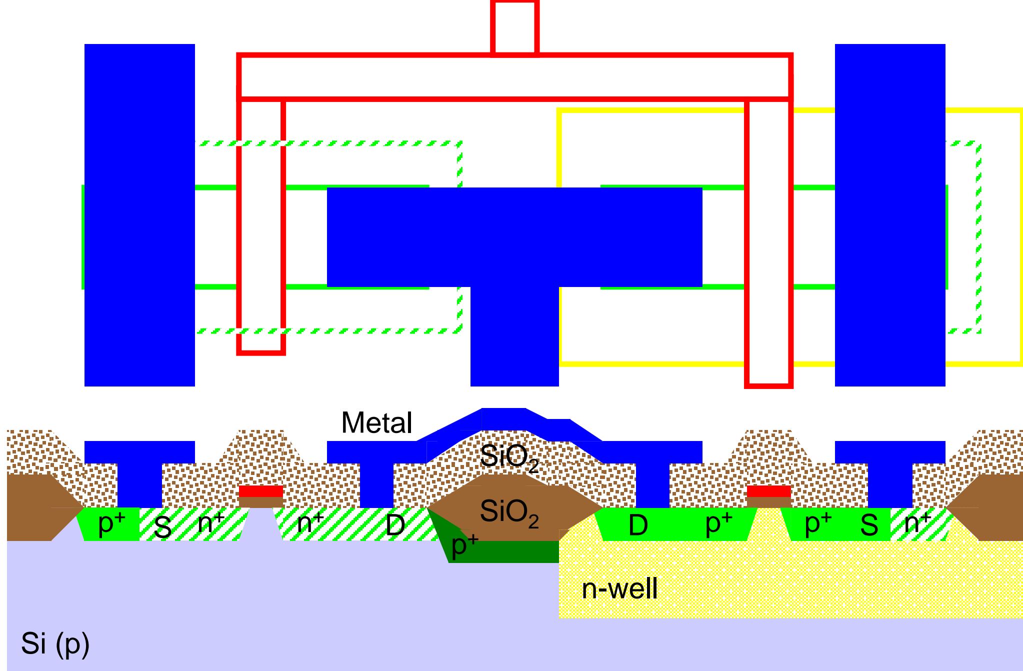
• Contacts to polysilicon must be made outside the gate region (avoid metal spikes through the poly and the thin gate oxide)



7. **Metallization mask** - define the interconnection pattern

• Aluminum is deposited over the entire wafer (evaporation) and selectively etched

• The step coverage in this process is most critical (nonplanarity of the wafer surface)



- The final step: the entire surface is passivated (overglass layer)
- Protect the surface from contaminants and scratches
- Than opening are etched to the bond pads to allow for wire bonding

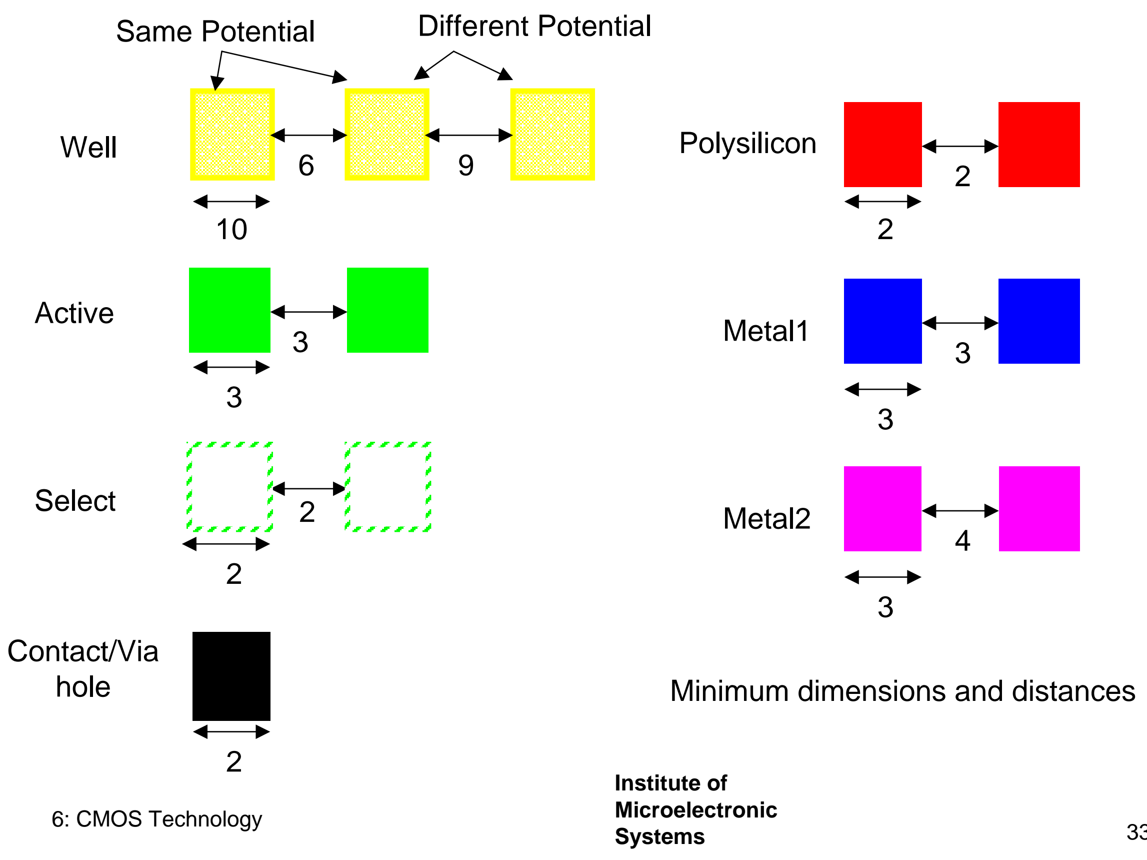
Design Rules

- Interface between designer and process engineer
- Guidelines for constructing process masks
- Unit dimension: minimum line width
- **Scalable design rules** - lambda (λ) parameter:
  - define all rules as a function of a single parameter λ
  - scaling of the minimum dimension: change the value of λ - **linear scaling**!
  - linear scaling is only possible over a limited range of dimensions (1-3μm)
  - are conservative: they have to represent the worst case rules for the whole set
  - for small projects are a flexible and versatile design methodology
- **Micron rules** - absolute dimensions:
  - can exploit the features of a given process to a maximum degree
  - scaling and porting designs between technologies is more demanding: manually or using advanced CAD tools!
- Ex.: Scalable CMOS design rules

CMOS Process Layers

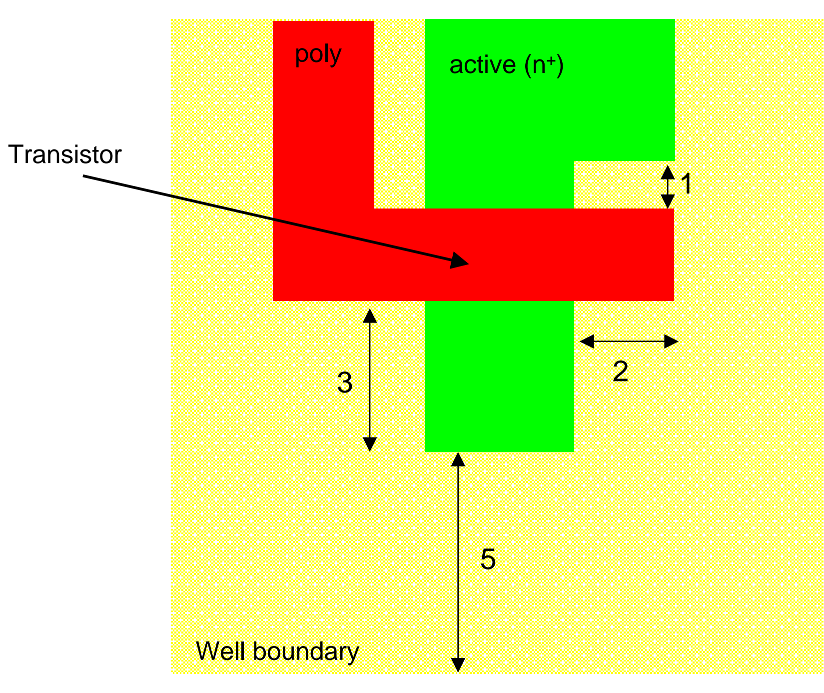
Layer	Color	Representation
Well (p,n)	Yellow	
Active Area (n^+,p^+)	Green	
Select (p^+,n^+)	Green	
Polysilicon	Red	
Metal1	Blue	
Metal2	Magenta	
Contact To Poly	Black	
Contact To Diffusion	Black	
Via	Black	

Intra-Layer Design Rules (λ)

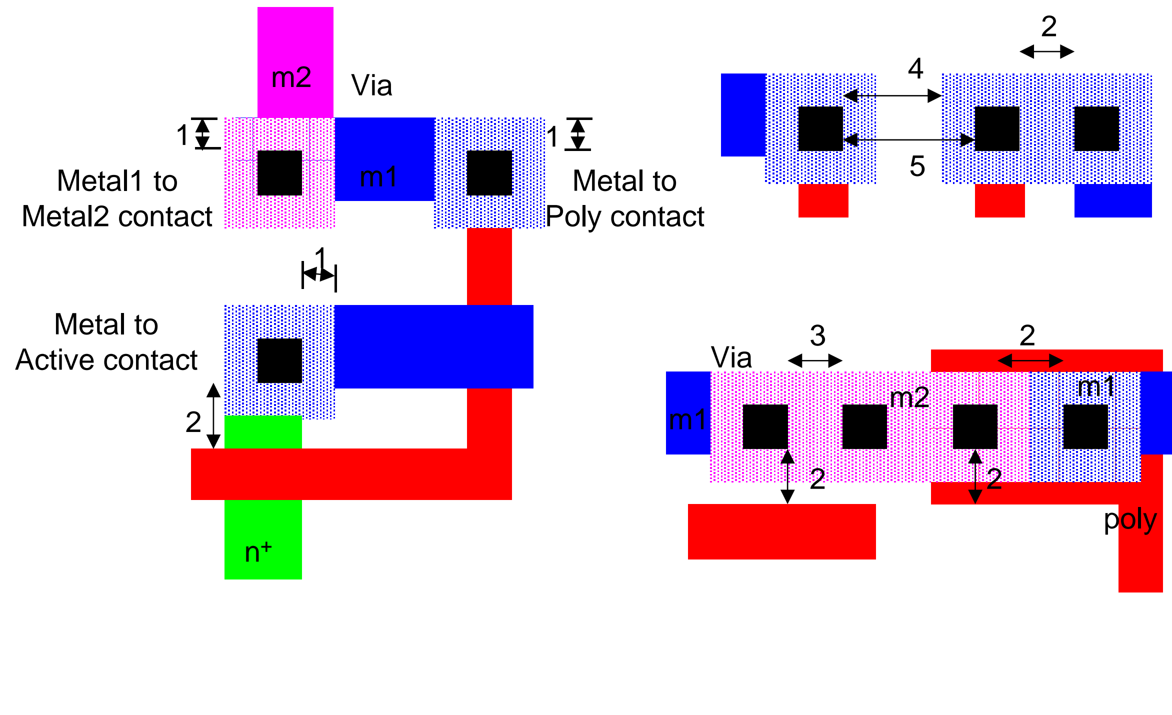


Minimum dimensions and distances

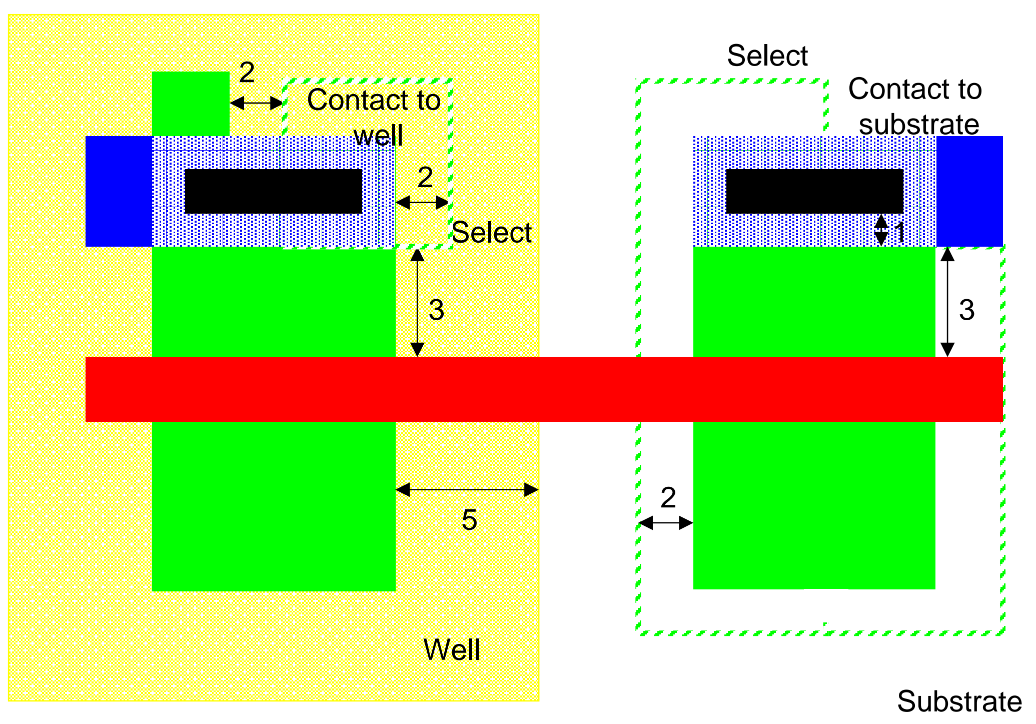
Inter-Layer Design Rules - Transistor Layout (λ)



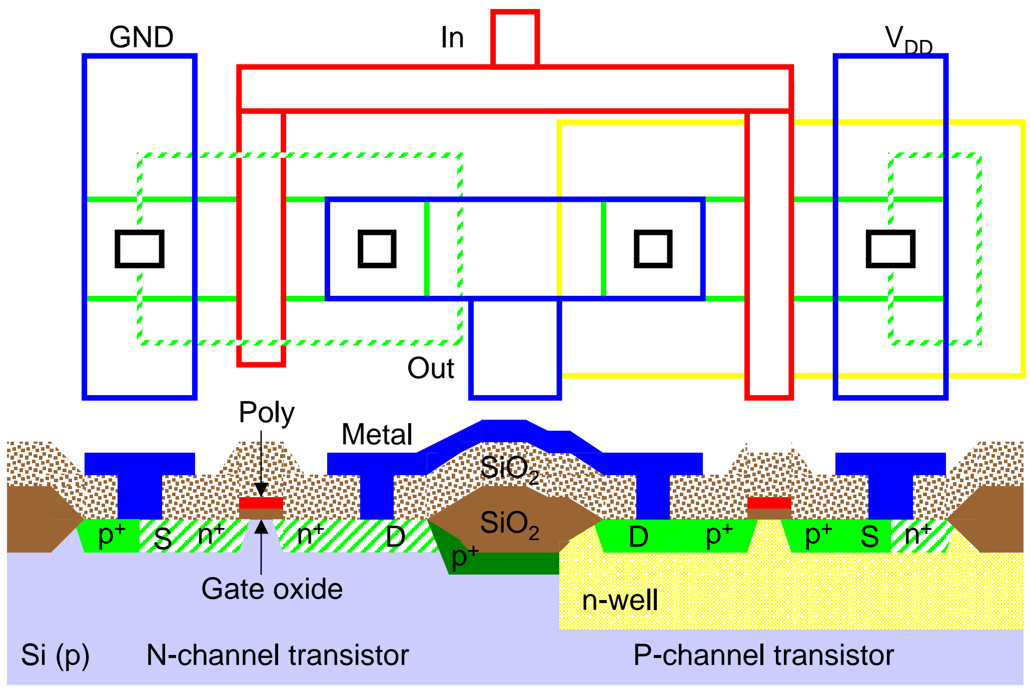
Inter-Layer Design Rules - Contact and Via (λ)



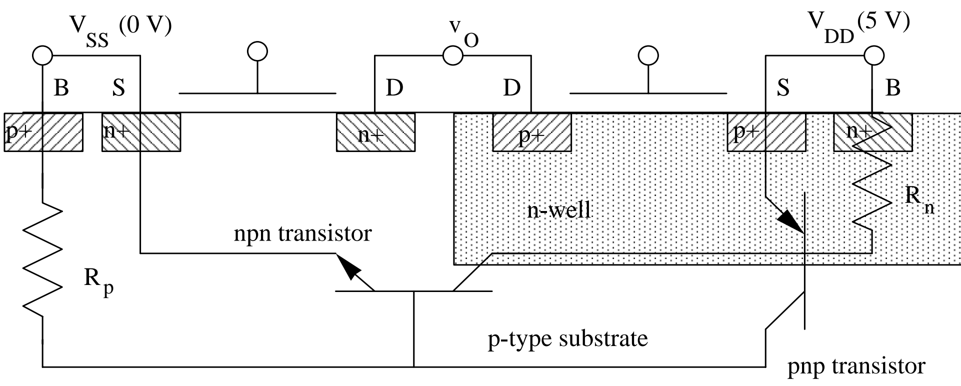
Select Layer (λ)



CMOS Inverter Layout



CMOS Latchup



- The parasitic bipolar transistors can destroy the CMOS circuitry
- The bipolar devices are normallly inactive
- The collector of each bipolar transistor is connected to the base of the other in a positive feedback structure
- The latchup effect can occur when:
  1. Both bipolar transistors conduct
  2. Product of gains of the 2 transistors in the feedback loop exceeds unity ($\beta_P\beta_N > 1$)