

Analog Layout Design

Learn a little about semiconductors and integrated circuits

THIS DOCUMENT INTRODUCES AN OVERVIEW METHODOLOGY FOR ANALOG LAYOUT DESIGN, IT GIVES VISUAL REPRESENTATION OF SEMICONDUCTOR DEVICES AND INTEGRATED CIRCUITS

SHUBHAM KUMAR

Date : 1/May/2023

CONTENTS

1. Linux Command.....	3.
2. IC Fabrication	5.
3. MOSFET Working	25
4. IC design process.....	26
➤ Layer introduction.....	30
➤ DNW layer.....	39
5. Layout introduction.	
➤ Poly	41.
➤ Device Fingers	43
➤ Metal Routing.....	45.
6. Basic importance of layout	
➤ Layout matching.....	48
➤ Dummy.....	59
➤ Current Mirror & Differential pair.....	61
➤ Length of diffusion (LOD effect).....	68
➤ Well proximity.....	70
➤ STI	71
7. General Layout Guidelines.....	72
8. Devices Available in layout	
➤ Resistor.....	74
➤ Capacitor.....	77
➤ Inductor.....	87
➤ Diode.....	89
➤ Transistor.....	91
➤ MOSFET.....	92
• Planer.....	93
• Finfet.....	95
• FdSoi.....	101
➤ Guardring.....	109

9. Routing Importance	
➤ Signal Routing matching.....	112
➤ Physical design routing process.....	118
➤ Noise and Shielding.....	121
➤ Buffer	126
➤ Double patterning (metal coloring).....	128
10. Logic cells	131
11. Parasitic Capacitance and Resistance.....	138
12. DRC and LVS	142
13. SXCUT and PSUB2 layer.....	144
14. Sheet Resistance.....	145
15. EM (due to current drop).....	150
16. IR (due to voltage drop).....	151
17. Antenna error.....	153
18. Latch-Up error.....	160
19. ESD and Clamp.....	167
20. I/O pads / Bond Pads & Wire bonding.....	174
21. Electrical Rule Check (ERC)/Softcheck.....	176
22. Density.....	177
23. High Speed Layout Design.....	179
24. Tie-Hi, Tie-Lo.....	189
25. Level Shifter.....	190
26. SERDES.....	191
27. Buffer Chips / Data Buffer.....	194
28. Complete ESD explanation	196
29. Example of Analog Circuit	224
30. Wafer Level Packaging.....	225
31. Industry example.....	238

1. LINUX COMMAND

- **pwd** : show work directory: e.g. if you are at your home directory then it will print something like /home/<username>
- **ls** :list directory: This command will list the items of a directory. If you don't specify a directory then it will list work directory, the place where you currently are.
- **ls -a**: will show the hidden files
- **ls -al** : will list the files and directories with detailed information like the permissions, size, owner, etc
- **cd** : change directory. It will change your work directly as you specify. You will have to specify a directory. eg: **cd /home/Document/**
- **cd..**: change directory one level up.
- **cp** :Copy Command. It will copy a file or directory. It is similar to Copy-Paste in GUI.
Ex:cp (file name) (location), **cp file.txt new_file.txt**
- **mv** :Move or rename files. It is like renaming a file or cut-paste in GUI.
- **rm, rm -r, rm -rf** : remove file or directory. ex: **rm file.txt**
- **rmdir**: remove empty directory. This will not remove content of the directory but it will delete a directory if it is empty.
- **mkdir** : make directory. It will create a new directory. We will need to specify a name. ex: **mkdir file/**
- **chmod** : It is used to change file permission. chmod stands for change mode in Linux., ex: chmod 775 (file name), **chmod775 file.txt**
basic permission of a file can have : **r(read), w(write), x(execute)**
- **clear** : clear Screen: Use clear command to clear the terminal screen.
- **grep**:The command “grep” is used to search for a text in the specified file/folder.
ex:grep “linux” file.txt
- **zip** :The command “zip” is used to compress one or more files and store them in a new file with .zip extension.
- **unzip** :The command “unzip” is used to decompress a .zip file and extract all the files within to current directory.

- **history** :The simple command “history” displays the list of all commands entered since the user started the session.
- **evinceand acroread** : This command is for pdf file open.
- **du -sh**: check the file size.
- **touch** : The touch command allows to update the access and modification times of the specified files
- **find**: used to search for files in a directory or a hierarchy based.
Ex: **find [flags] [path] -name [expression]**
- **ping**: used for network connectivity test, ex: **ping google.com, ping 8.8.8.8**
- **whoami**: used to display the username currently in use
- **passwd**: used to change the password of user account
- **kill**: for terminate the tool, ex: **kill 532452, kill firefox**
- **exit** : exit from terminal.
- **shutdown** : To power off your computer immediately.

2. IC FABRICATION

What is IC?

Integrated circuits (IC) sometimes called a chip or microchip is a semiconductor wafer on which thousands or millions or trillions of tiny resistors, capacitors and transistors are fabricated.

An IC can function as an oscillator, amplifier, timers, computer memory or microprocessor.

An IC is categorized as either linear (analog) or discrete (digital) depending on its intended applications.

Moore's law

In the year 1965 Gordon Moore Intel Co-Founder and Chairman predicted that transistors would continue to shrink, allowing

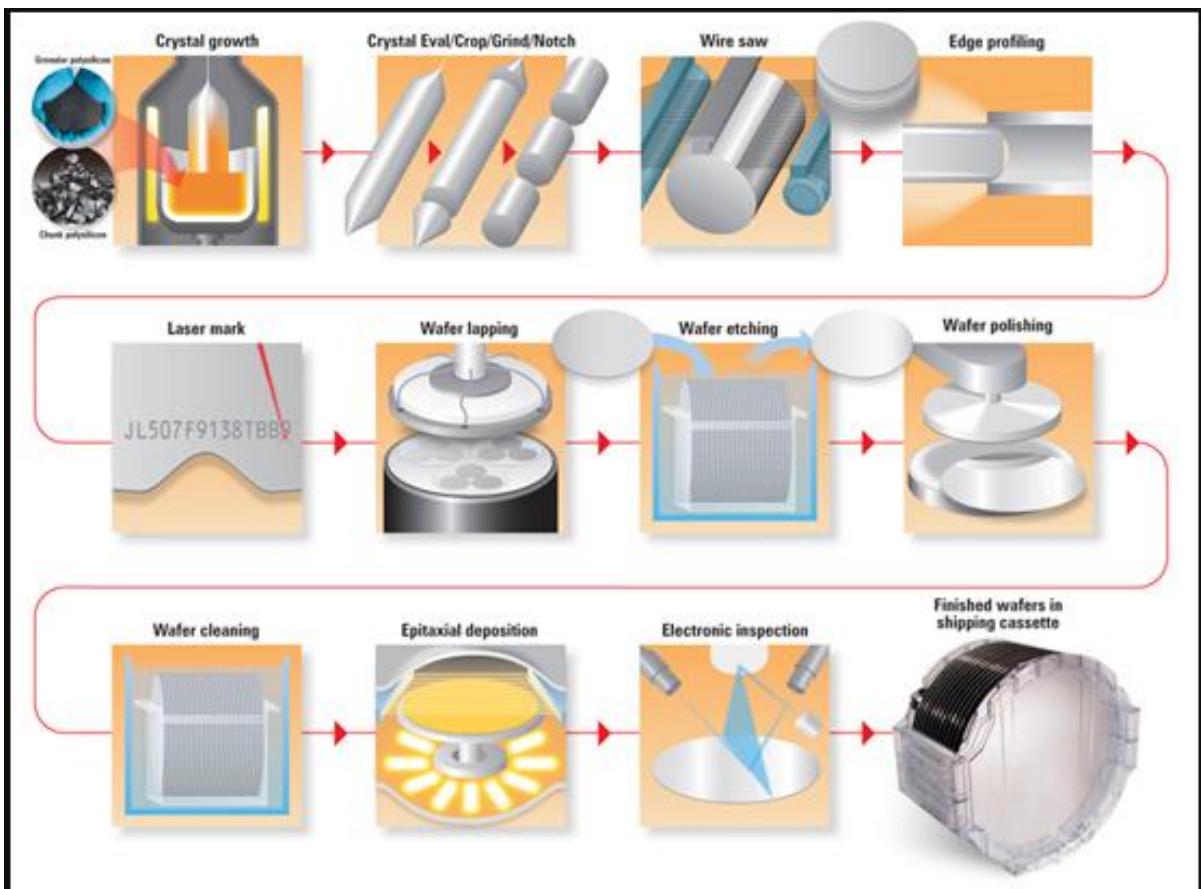
- Doubled transistors density every 18-24 month
- Doubled performance every 18-24 month
- The period often quoted as 18 month

Basic steps of IC fabrication

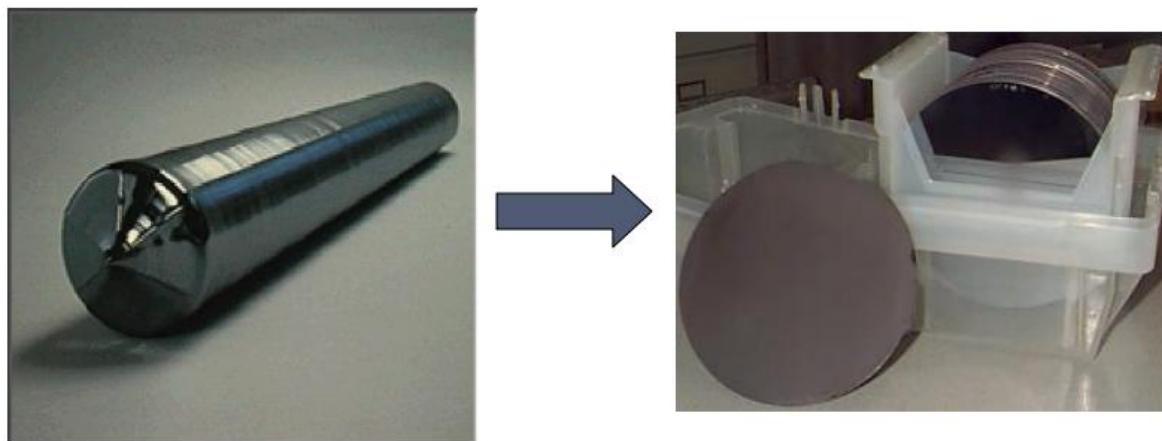
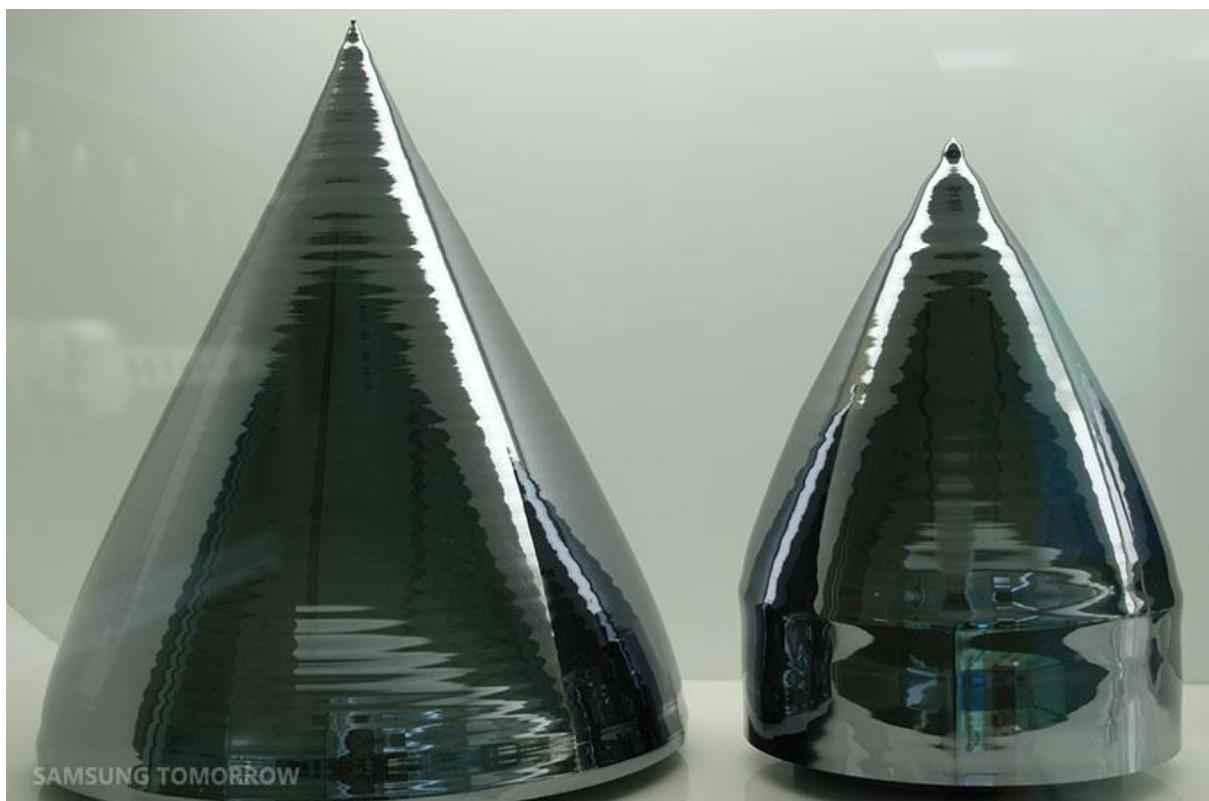
- Wafer production
- Epitaxial growth
- Photolithography
- Masking
- Etching
- Doping
- Atomic diffusion
- Ion implantation
- Metallization
- Assembly and packaging

Wafer Production

- The Wafer is round slice of semiconductor material such as silicon. Silicon is preferred due to its characteristics; it is more suitable for manufacturing IC.
- It is base or substrate to entire chip
- First purified polycrystalline silicon is created from the sand. Then it is heated to produce molten liquid.
- A small piece of solid silicon is dipped on the molten liquid.
- Then the solid silicon is dipped on the molten liquid.
- A thin round wafer of silicon is cut using wafer slicer. Wafer slicer is a precise cutting machine and each slice having thickness about 0.01 to 0.025 inches.
- When wafer is sliced, the surface will be damaged. It can be smoothening by polishing. After polishing the wafer. It must thoroughly clean and dried.
- The wafers are cleaned using high purity low particle chemicals. The silicon wafers are exposed to ultra pure oxygen

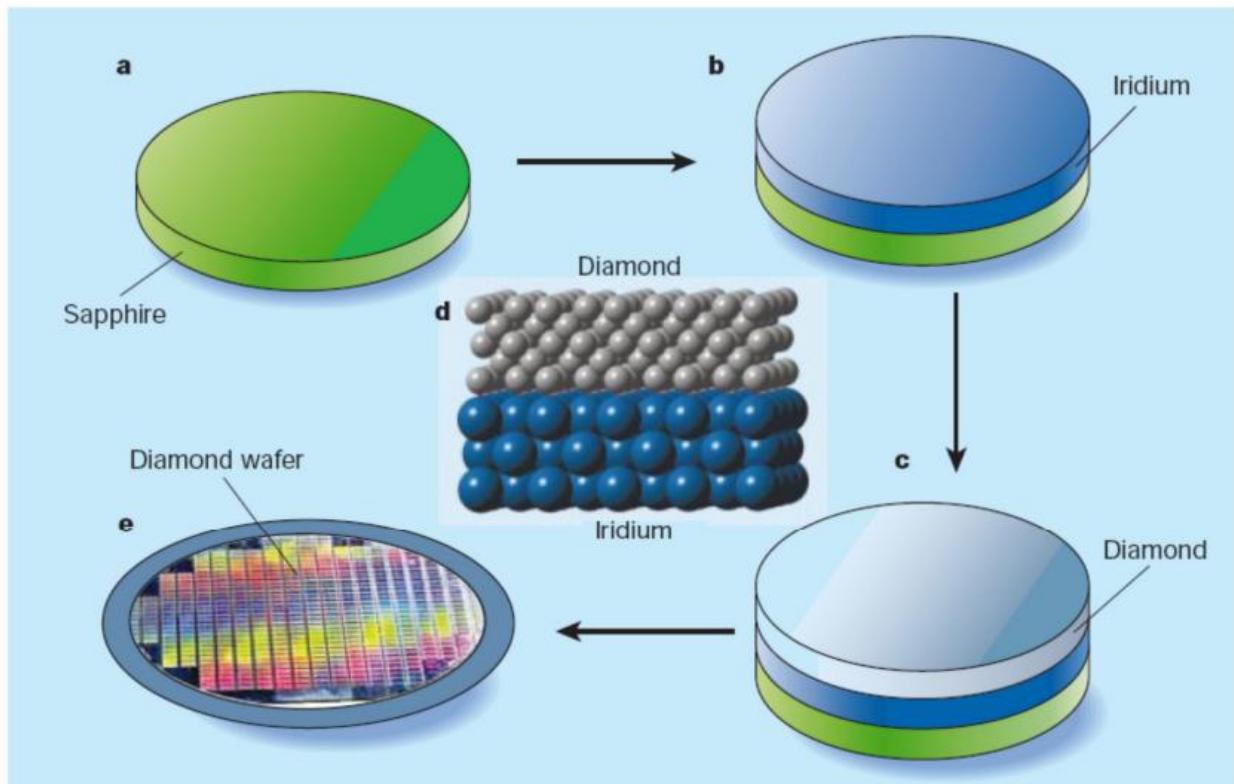


Starting substrates are cut and polished from huge single crystals of silicon. Not done by the foundries.



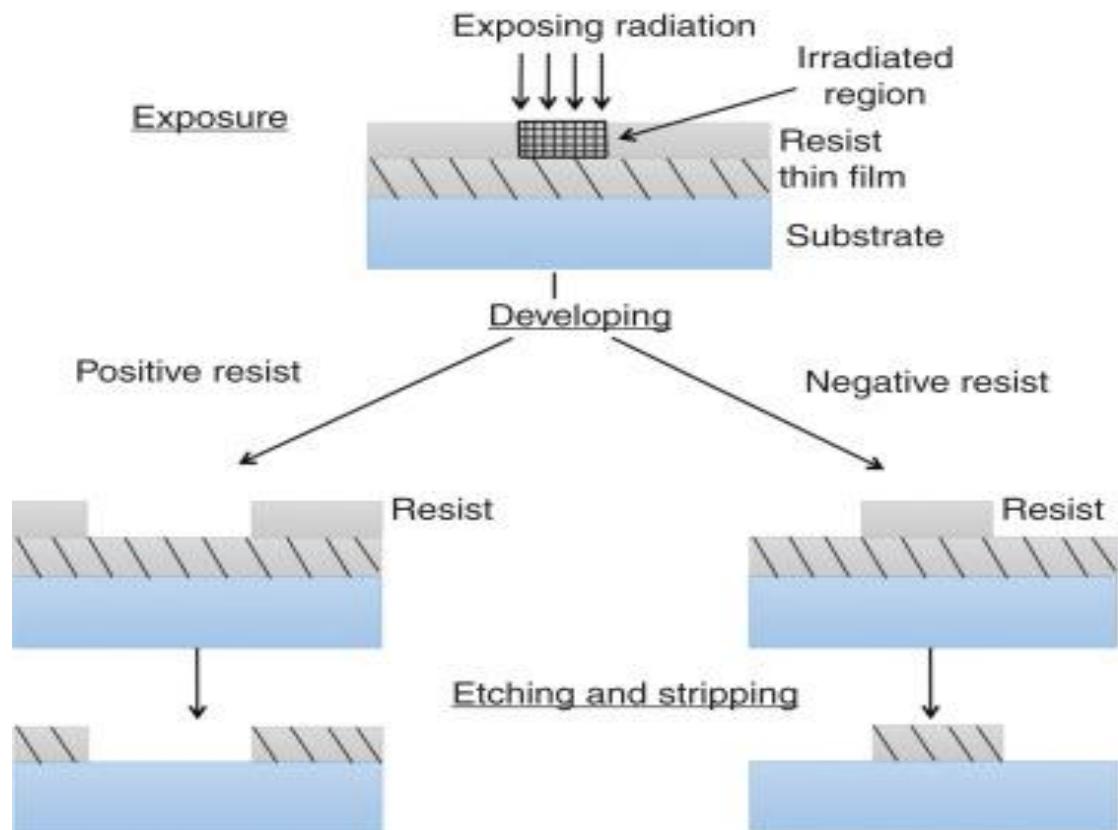
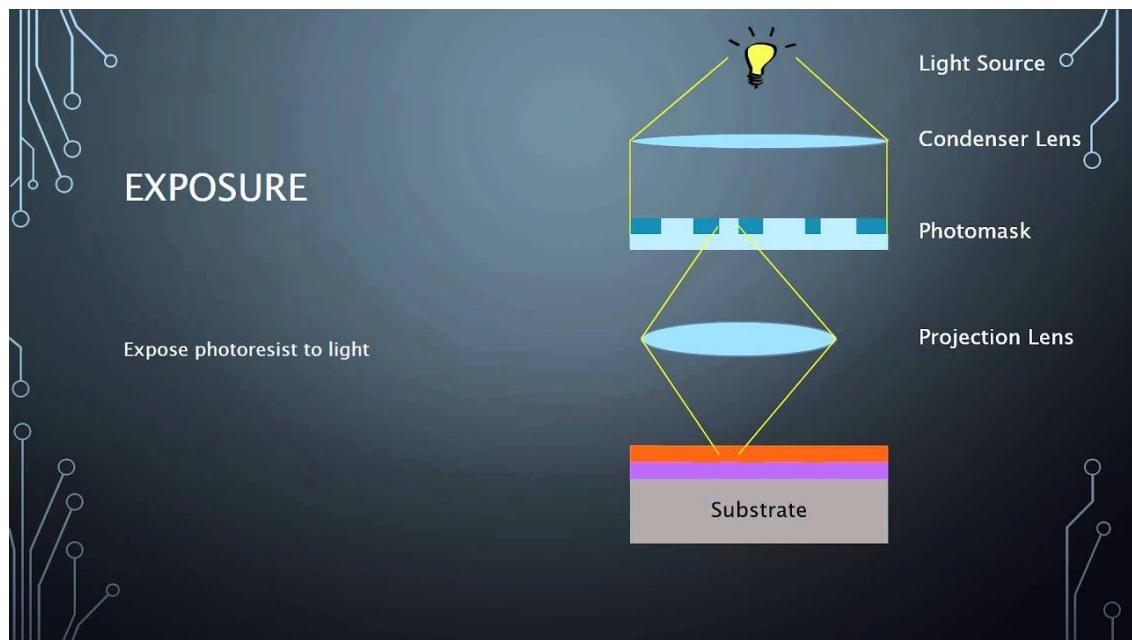
Epitaxial growth

- It means the growing of single silicon crystal upon original silicon substrate.
- A uniform layer of silicon dioxide is formed on the surface of wafer.



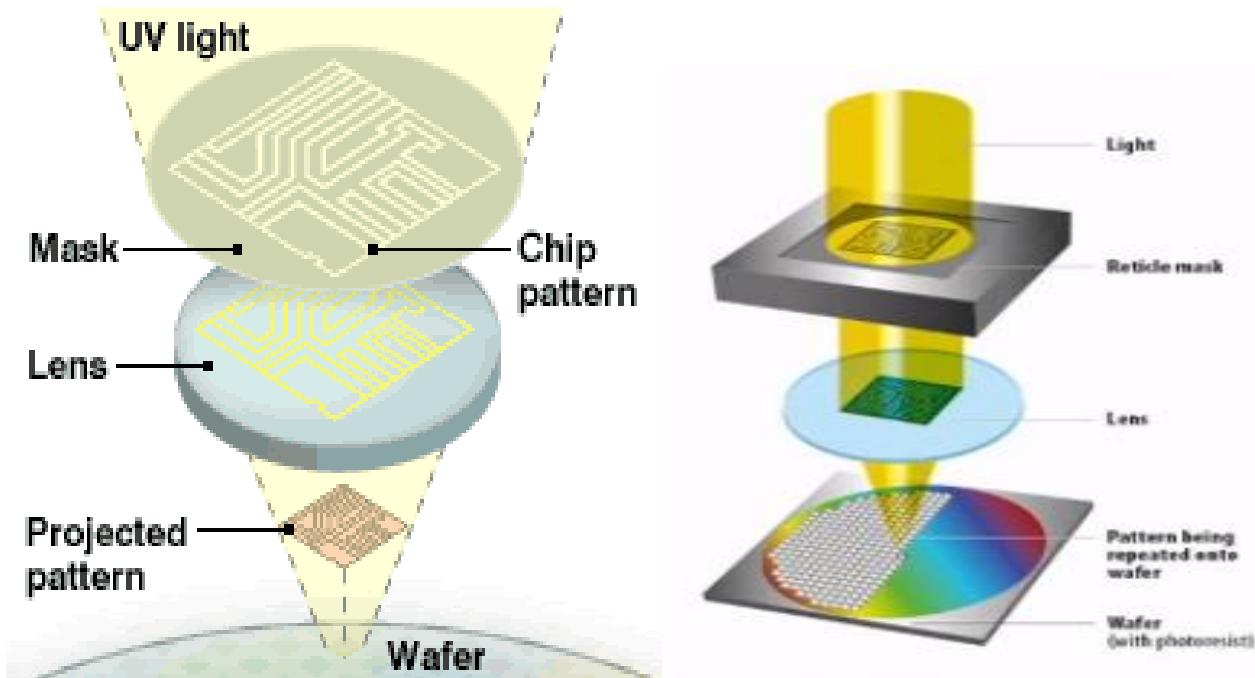
Photolithography

- Photolithography is a technique that is used to define the shape of micro-machined structures on the wafer.



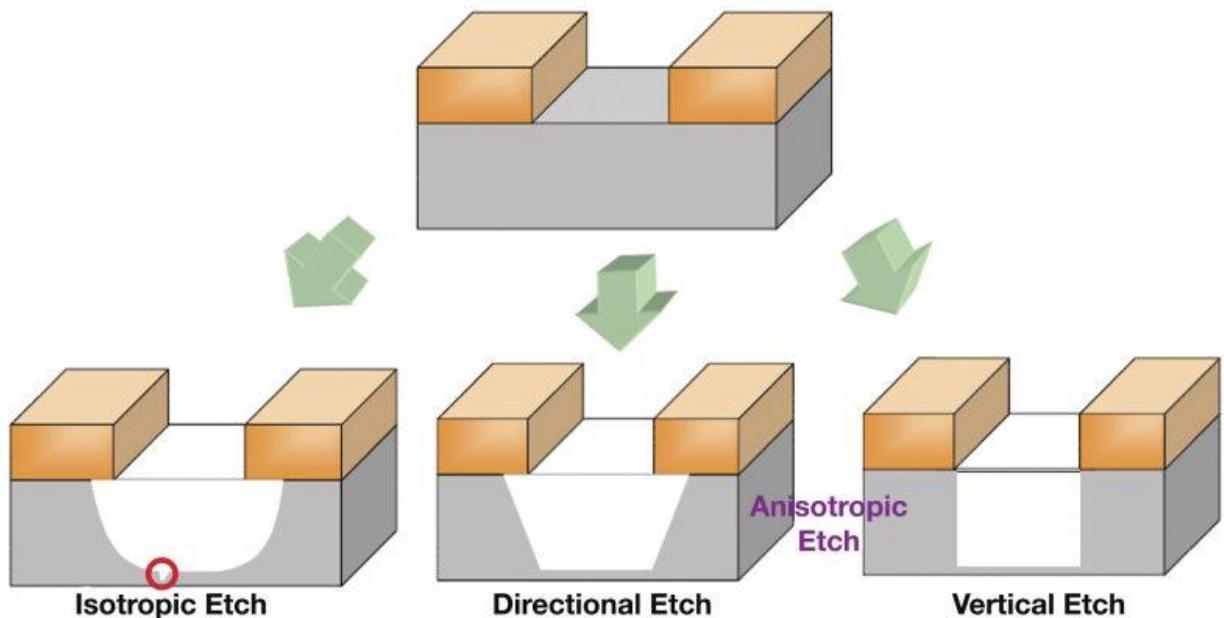
Masking

- To protect some area of wafer when working on another area, a process called photolithography is used.
- The process of photolithography includes masking with a photographic mask and photo etching.
- A photoresist film is applied on the wafer.
- The wafer is aligned to a mask using photo aligner.
- Then it is exposed to ultraviolet light through mask.
- Before that the wafer must be aligned with the mask.



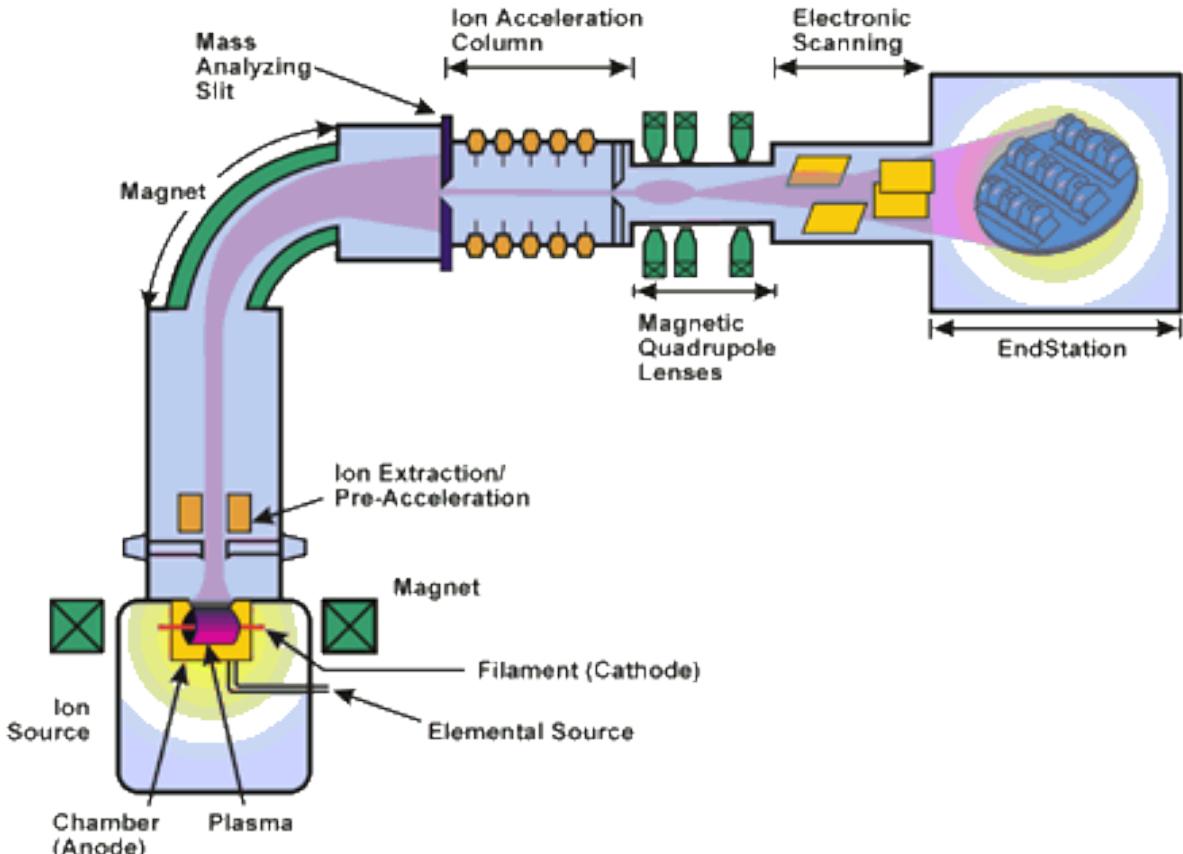
Etching

- Selectively removing unwanted material from the surface of the wafer.
- The pattern of the photo-resist is transferred to the wafer by means of etching agents
- The parts of material are protected by this etching mask.



Doping

- To alter the electrical character of silicon, atom with one less electron than silicon such as boron and atom with one electron greater than silicon such as phosphorous are introduced into the area.
- The P-type(boron) and N-type (phosphorous) are created to reflect their conducting characteristics.



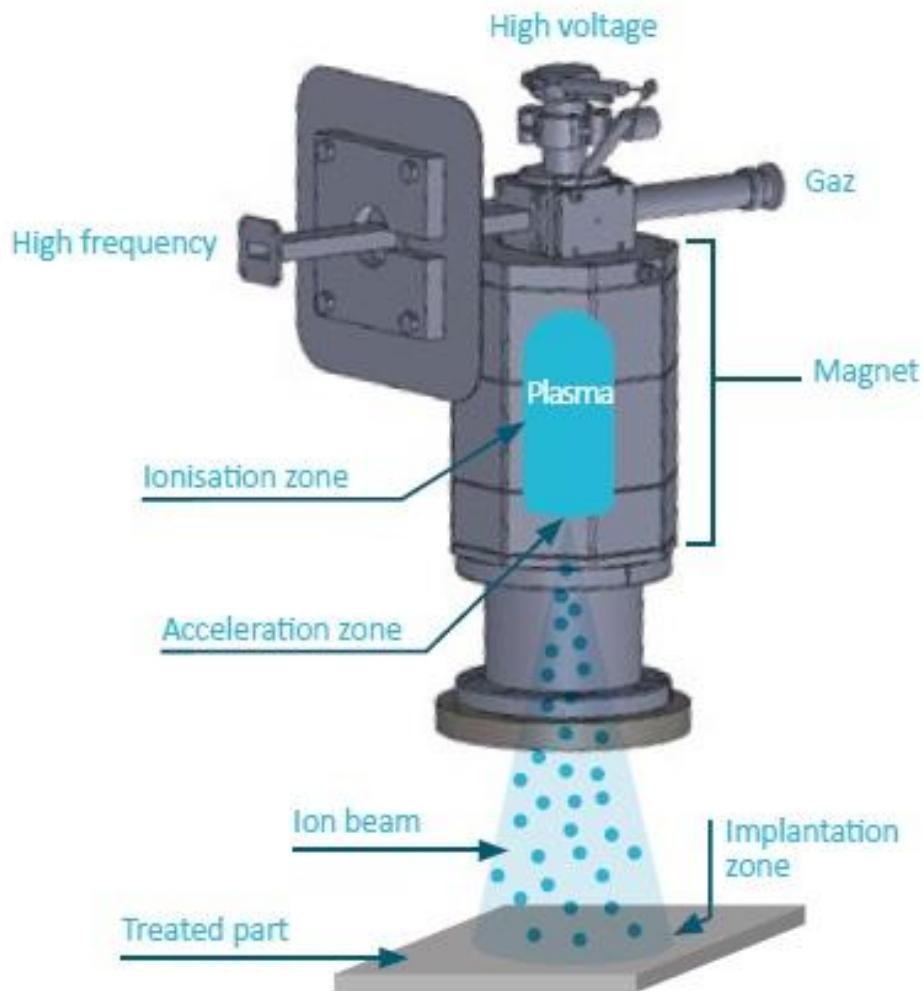
Atomic diffusion

- Diffusion is defined as the movement of impurity atom in semiconductor material at high temperature.
- Then the wafers are heated at a temperature of about 1500-2200F. The inert gas carries the dopant chemical. The dopant will get deposited on the wafer.
- This method can only be used for large areas. For small areas it will be difficult and it may not be accurate.

Ion implantation

- Ion implantation is a low-temperature process by which ions of one element are accelerated into a solid target, thereby changing the physical, chemical, or electrical properties of the target.
- Ion implantation is used in semiconductor device fabrication and in metal finishing, as well as in material science research.

- The ions can alter the elemental composition of the target (if the ions differ in composition from the target) if they stop and remain in the target.
- Ion implantation also causes chemical and physical changes when the ions impinge on the target at high energy.



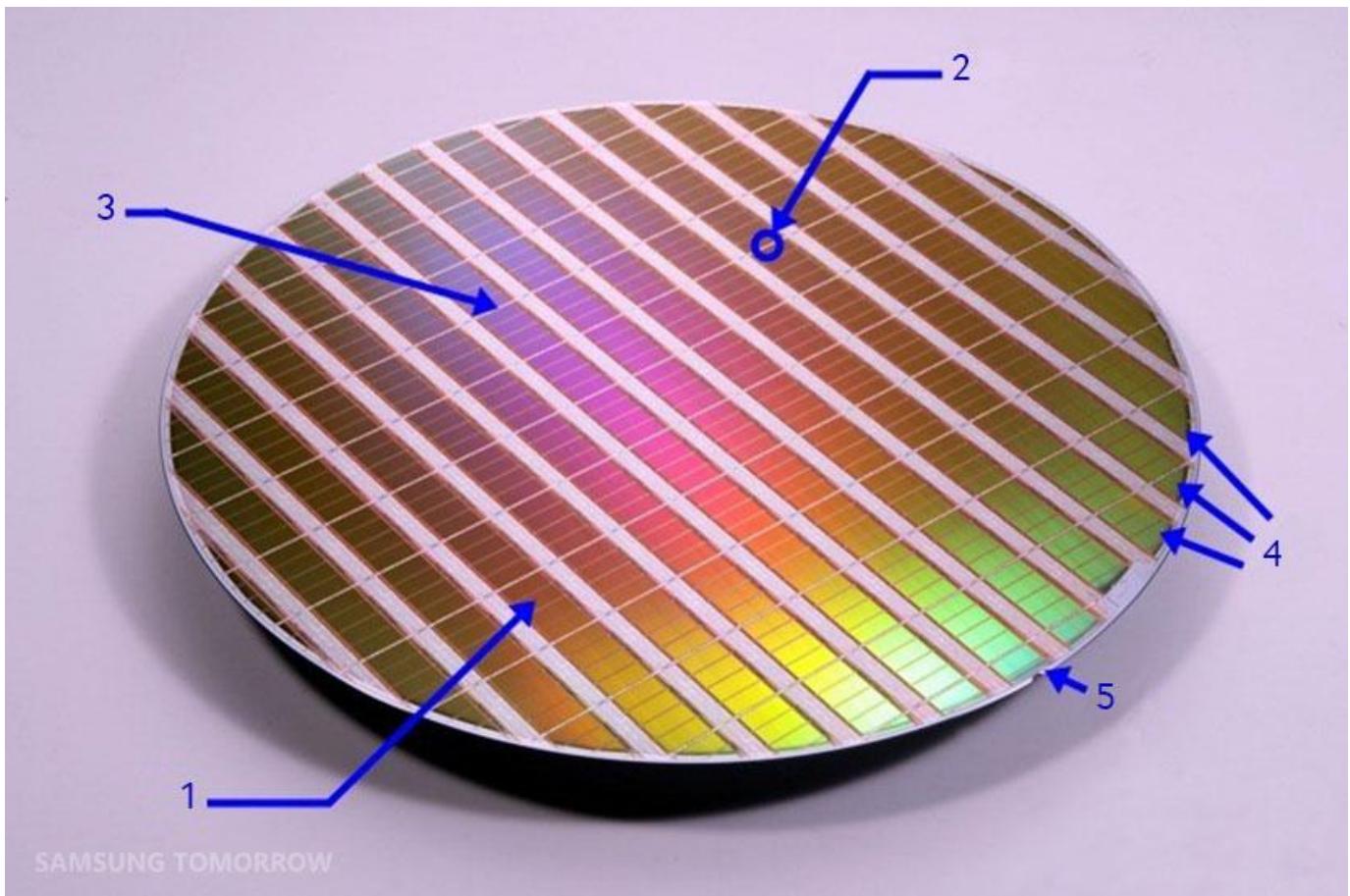
Metallization

- It is used to create contact with silicon and to make interconnections on chip.
- A thin layer of aluminum is deposited over the whole wafer.
- Aluminum is selected because it is a good conductor, has good mechanical bond with silicon, forms low resistance contact and it can be applied and patterned with single deposition and etching process.
- It must be patterned and etched to form the actual wires connecting individual device into a circuit.

Assembly and packaging

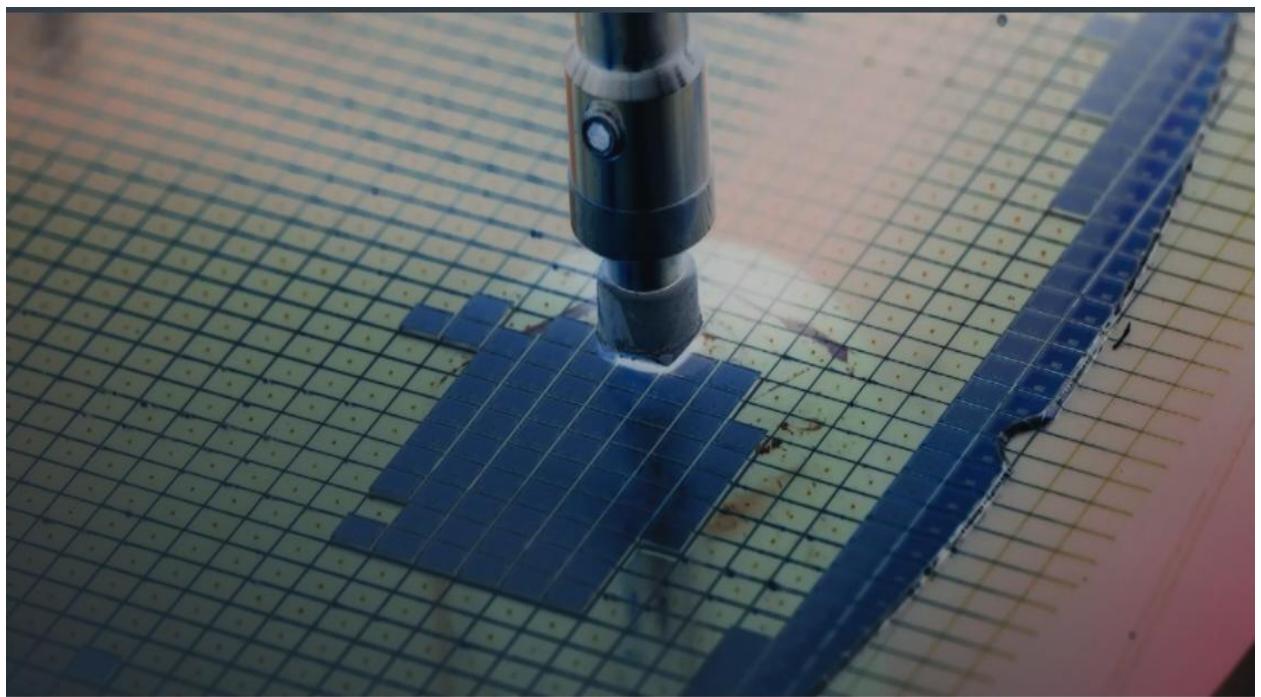
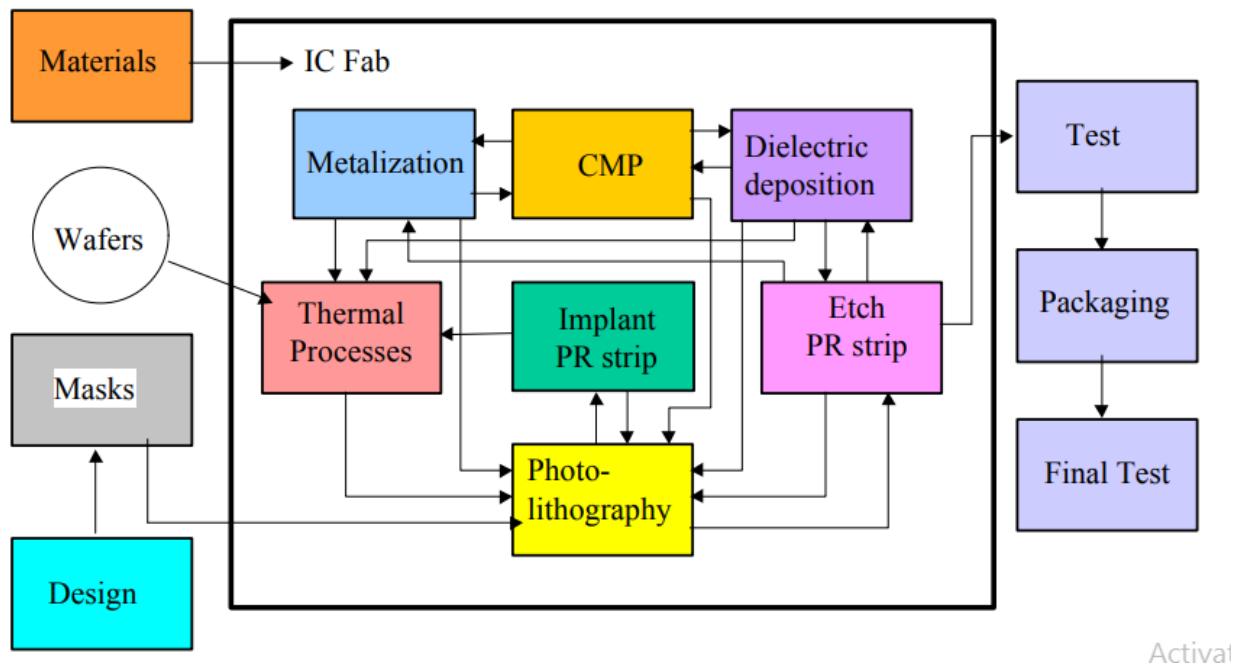
- Each of the wafers contains hundreds of chip. These chips are separated and packaged by a method called scribing and cleaving.
- The wafer is similar to a piece of glass. A diamond saw cut the wafer into single chip.
- The diamond tripped tool is used to cut the lines through the rectangular grid which separated the individual chip.
- The chip that are failed in the electrical test are discarded.
- Before packaging, remaining chips are observed under microchip. The good chip is then mounted into a package.

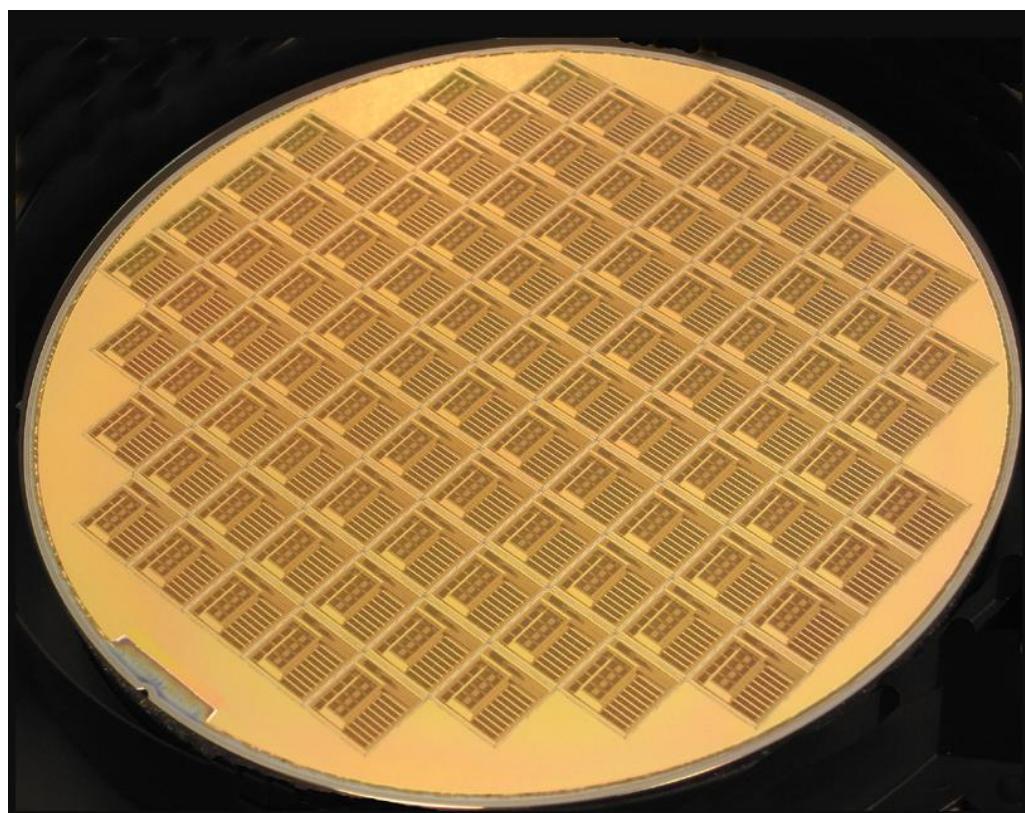
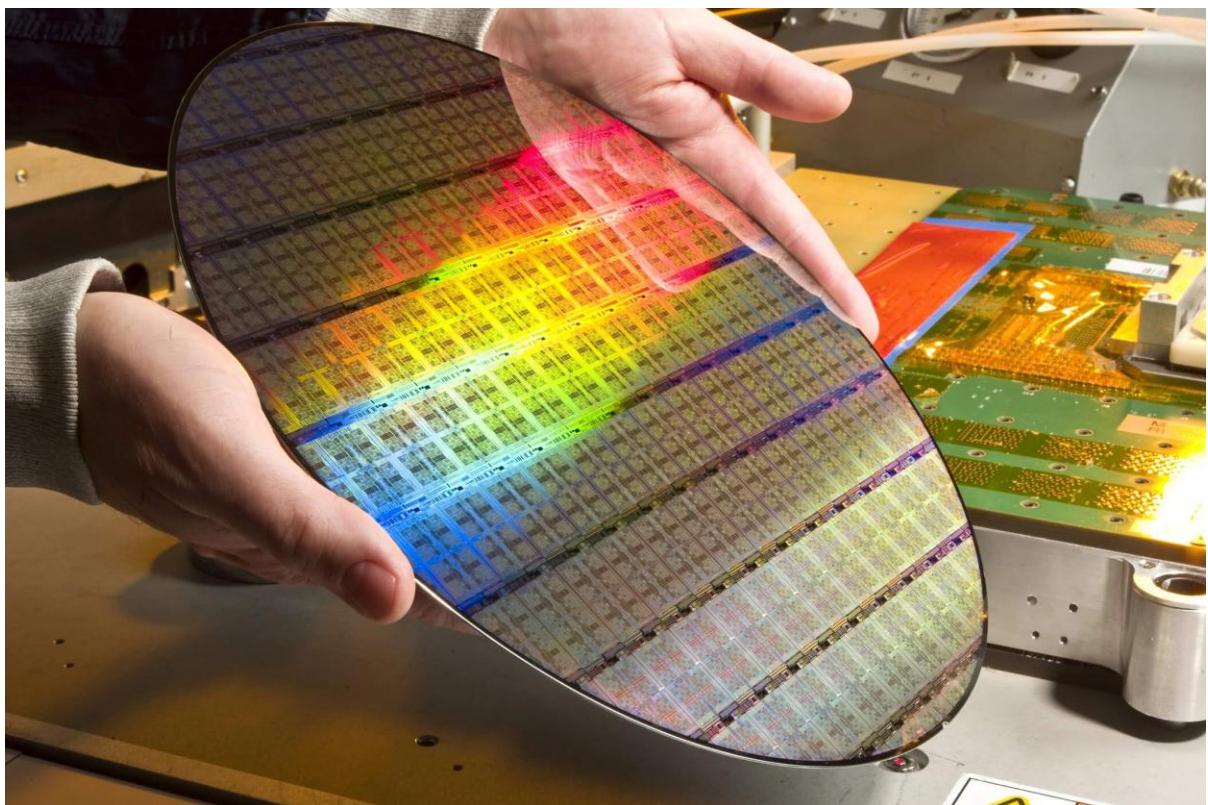
Wafer and chips:



1. **Chip**: a tiny piece of silicon with electronic circuit patterns
2. **Scribe Lines**: thin, non-functional spaces between the functional pieces, where a saw can safely cut the wafer without damaging the circuits
3. **TEG (Test Element Group)**: a prototype pattern that reveals the actual physical characteristics of a chip (transistors, capacitors, resistors, diodes and circuits) so that it can be tested to see whether it works properly
4. **Edge Die**: dies (chips) around the edge of a wafer considered production loss; larger wafers would relatively have less chip loss
5. **Flat Zone**: one edge of a wafer that is cut off flat to help identify the wafer's orientation and type

Wafer Process Flow



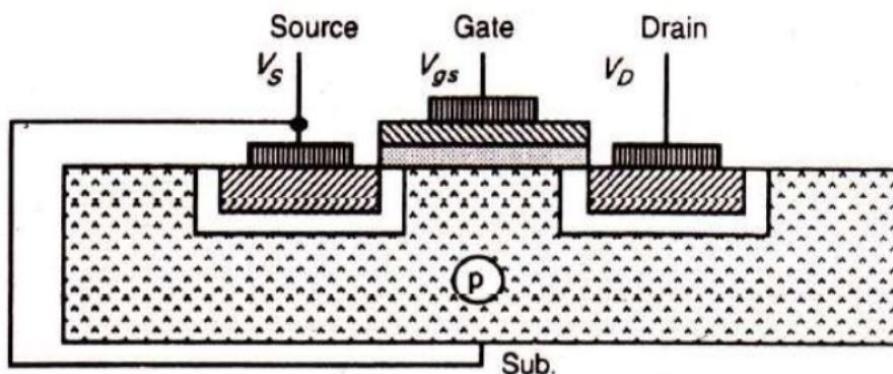


Shubham Kumar
krshubham018@gmail.com

Video link:

1. <https://youtu.be/c9arR8T0Qts>
2. <https://youtu.be/fwNkg1fsqBY>
3. <https://www.computerhistory.org/revolution/digital-logic/12/288/2220>

NMOS FABRICATION



Representation of elements

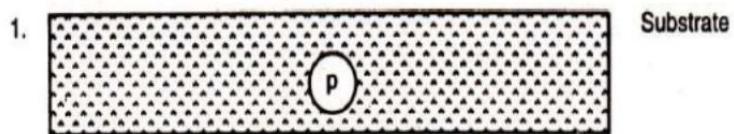
KEY	
	Metal
	Polysilicon
	Oxide
	n-diffusion
	p-diffusion
	p-substrate
	n-substrate
	Depletion

NMOS Fabrication Steps

- There are a huge number of fundamental fabrication steps utilized as a part of the generation of present-day MOS ICs.
- A similar procedure can be utilized for the planned of NMOS or PMOS or CMOS devices. The most commonly used material could be either metal or poly-silicon.

STEP 1:-

- A thin layer of silicon wafer is converted into p type material by doping with Boron material.



STEP 2:-

- A thick layer of silicon dioxide (SiO_2) is grown on entire p-type substrate



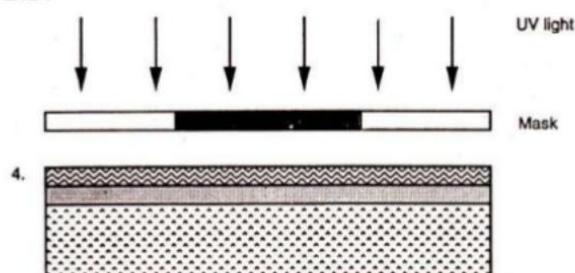
STEP 3:-

- The surface is now covered with a photoresist which is deposited over the thick layer of sio₂.



Step 4:-

- The photoresist layer is then exposed to ultraviolet light through a mask which defines those regions into which diffusion is to take place together with transistor channels.



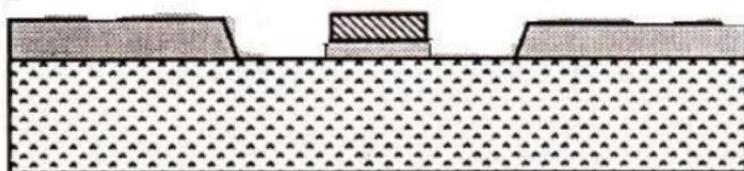
Step 5:-

- These areas are etched away together with the underlying silicondioxide so that the wafer surface is exposed in the window defined by the mask.



Step 6:-

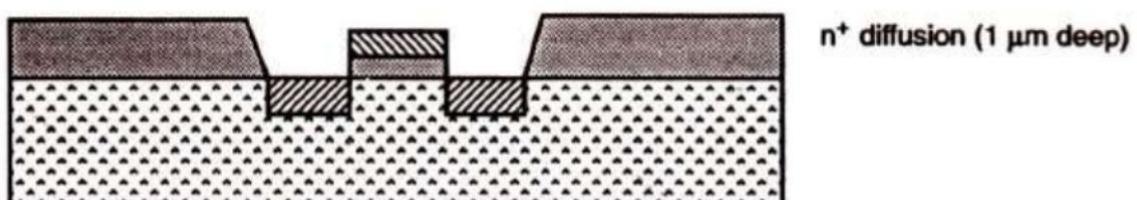
- The remaining photoresist is removed and a thin layer of SiO₂ (0.1 micro meter typical) is grown over the entire chip surface and then polysilicon is deposited on top of this to form the gate structure.
- A photoresist is deposited over the entire the polysilicon layer and expose UV light through the mask2



Step 7:-

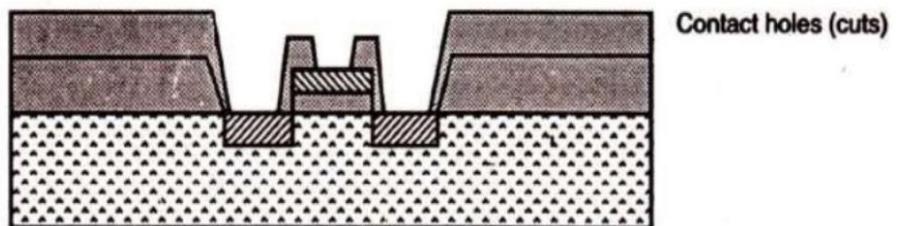
- Diffusion is achieved by heating the wafer to a high temperature and passing a gas containing the desired n-type impurity (for example, phosphorus)

Note: The polysilicon has an underlying thin oxide which acts as a mask during diffusion. This is called self-aligning.



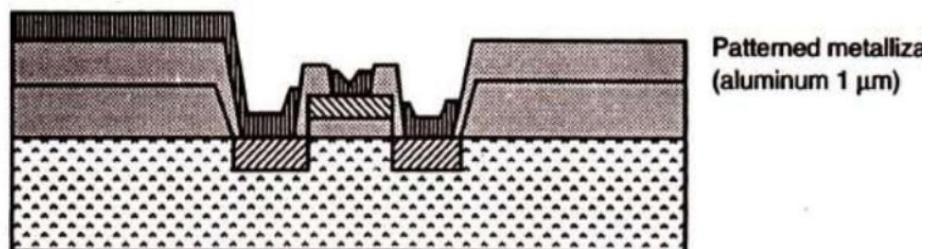
Step 8 :-

- A Thick oxide (SiO_2) of 1(micro meter) thickness is grown over all and photoresist material is placed on it.
- Expose the UV light through mask-3 on the selected areas of the polysilicon gate ,drain and source areas are etched to create the contact cuts.

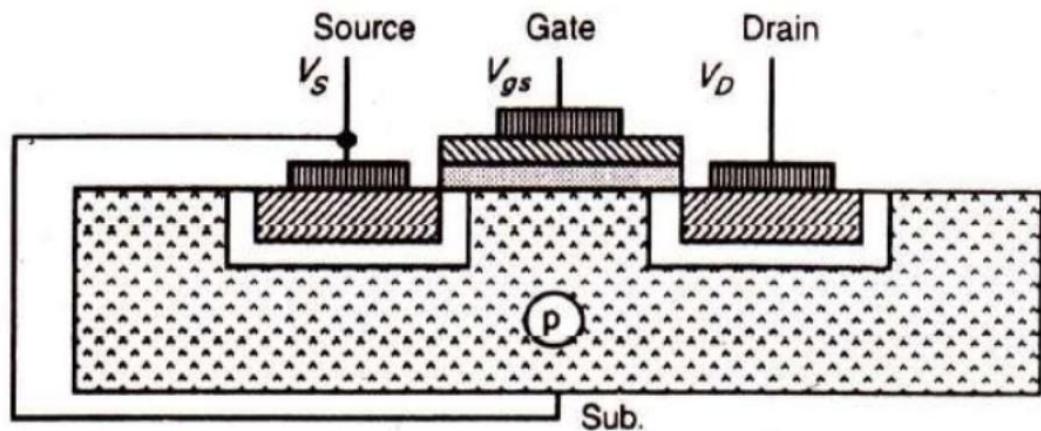


Step 9 :-

- Now a metal(aluminium) is deposited over its surface of 1(micro meter) thickness.
- Again a photo resist material is grown all over the metal and Expose the UV light through mask-4 are etched to from the required interconnection pattern.



- The final NMOS structure is



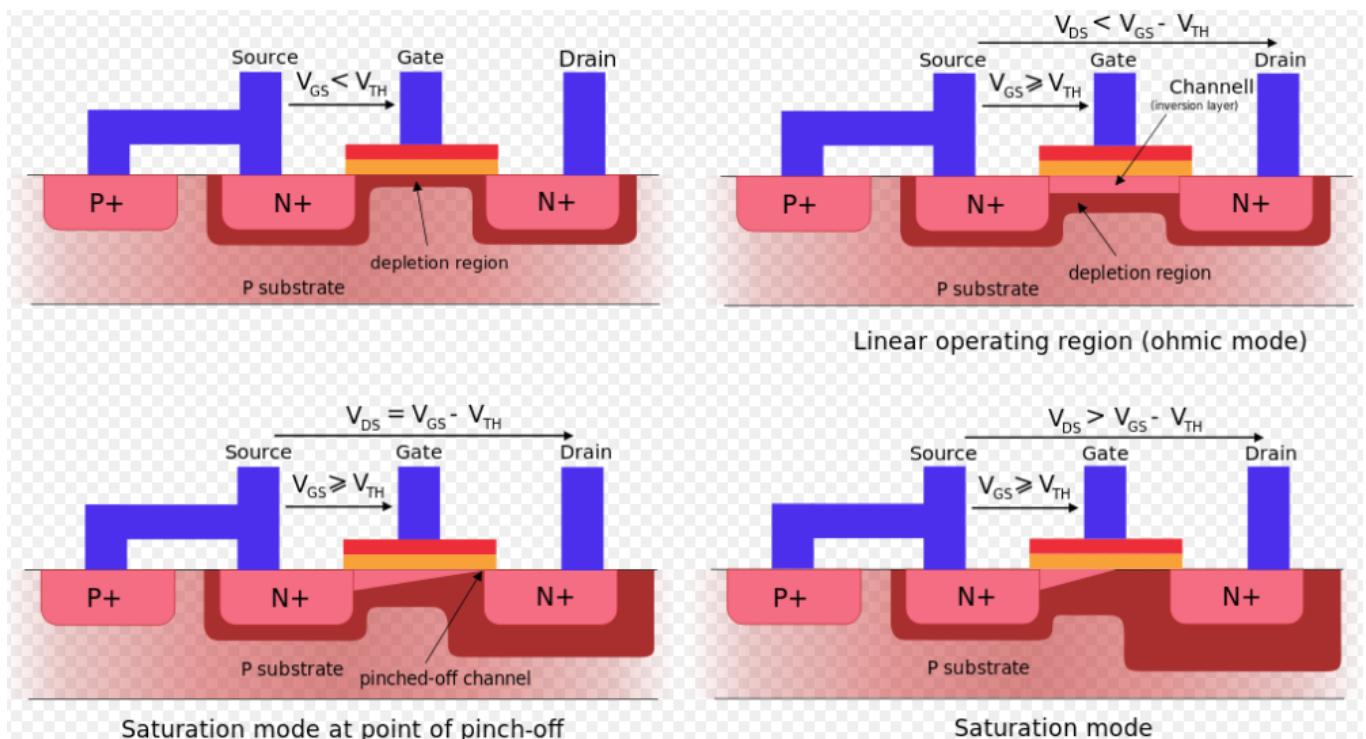
Video link :<https://youtu.be/1Lad28K3Xi0>

3. MOSFET WORKING

The main principle of the MOSFET device is to be able to control the voltage and current flow between the source and drain terminals. It works almost like a switch and the functionality of the device is based on the MOS capacitor. The MOS capacitor is the main part of MOSFET.

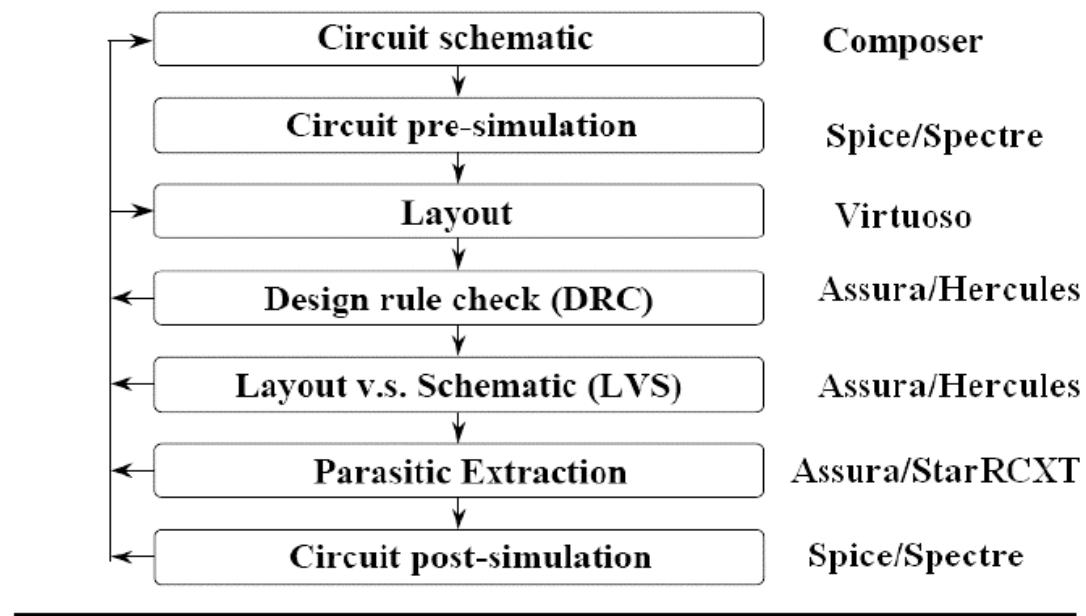
NMOSFET: In the case of an N-type switch, the body or back gate terminal is connected to the most negative supply (usually GND in single power supply systems) and the gate is used as the switch control.

Whenever the gate voltage exceeds the source voltage by at least a threshold voltage, the MOSFET conducts.

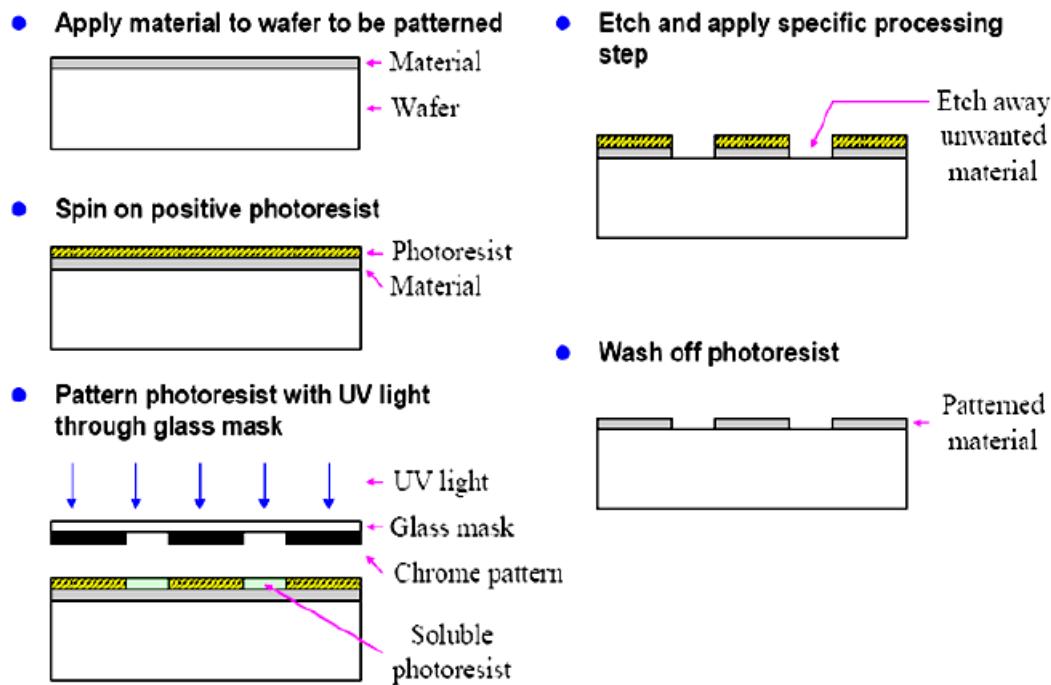


4. IC DESIGN PROCESS

Fully Custom IC Design Flow

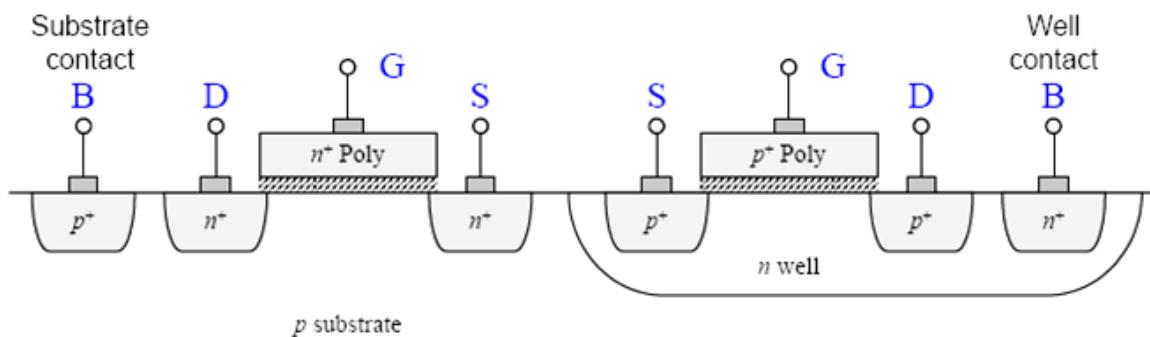


IC Photolithographic Process



CMOS P-Substrate Process Flow

- Cross section view



- NMOS process steps

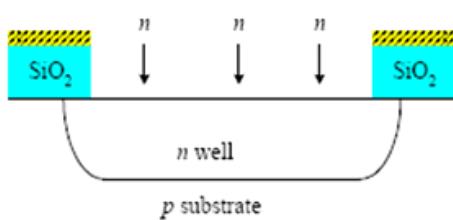
- Active region, poly gate, p+/n+ implant, metal contact/line

- PMOS process steps

- N well, active region, poly gate, p+/n+ implant, metal contact/line

Process Flow v.s. Layout

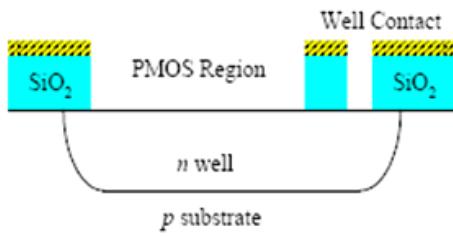
- N well



- NWELL layer (NW)



- Active region

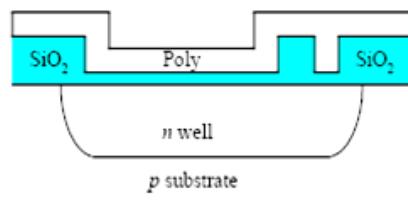


- DIFF layer (OD)

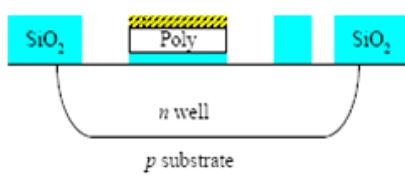
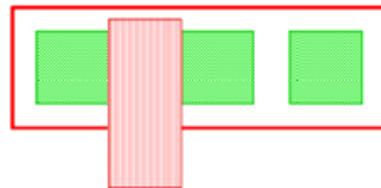


Process Flow v.s. Layout

- Poly gate

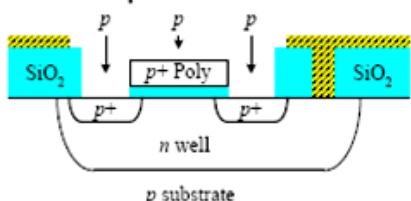


- POLY1 layer (PO)

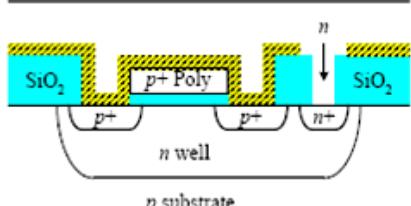


Process Flow v.s. Layout

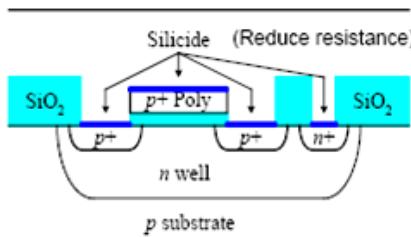
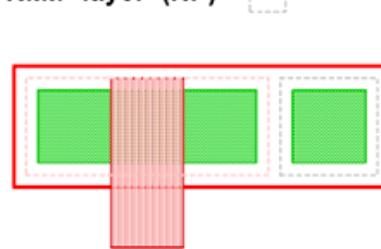
- P+/N+ implant



- PIMP layer (PP)

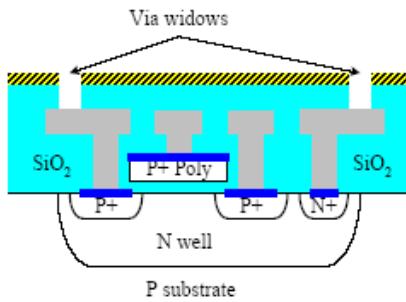


- NIMP layer (NP)

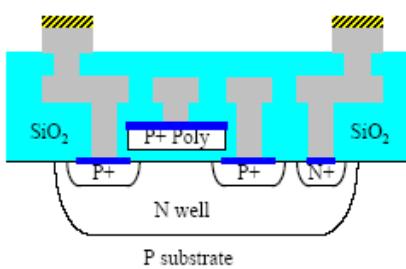
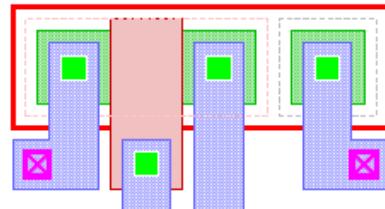


Process Flow v.s. Layout

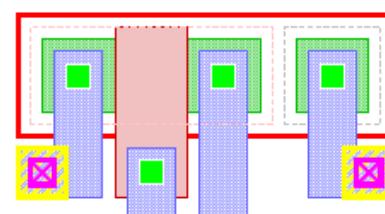
- Metal vias/lines



- VIA12 layer (VIA1)

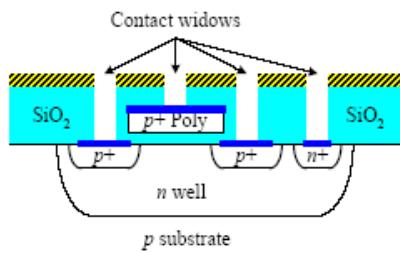


- METAL2 layer (M2)

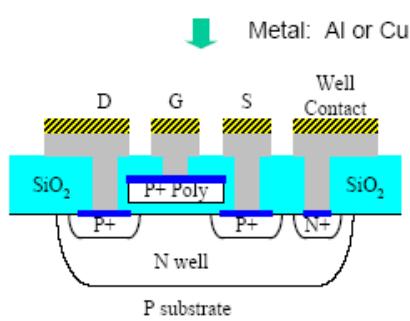
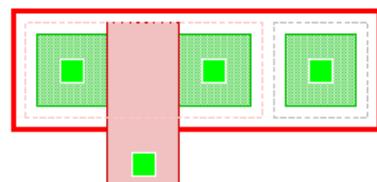


Process Flow v.s. Layout

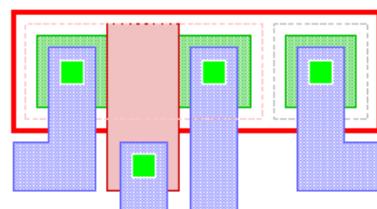
- Metal contacts/lines



- CONT layer (CO)



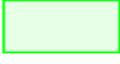
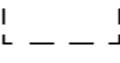
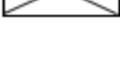
- METAL1 layer (M1)



Layer introduction :

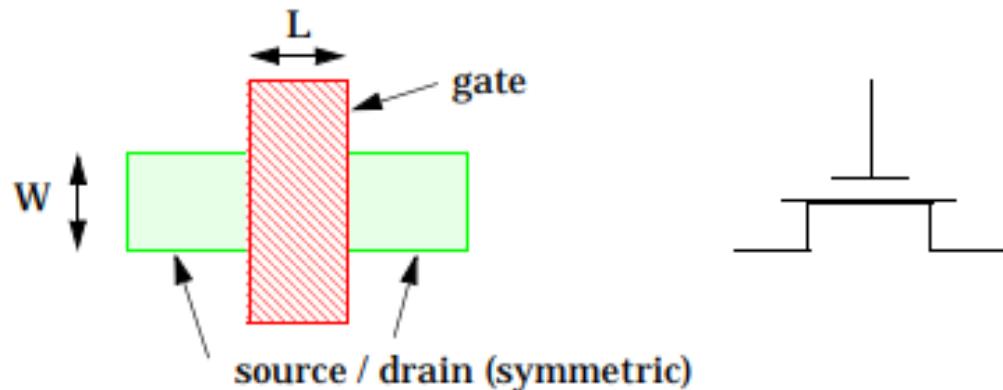
Some basic layers are used in any layout are as follows:

- ✓ N-Well
- ✓ PPlus / NPlus
- ✓ OD
- ✓ Poly
- ✓ Metals
- ✓ Contact

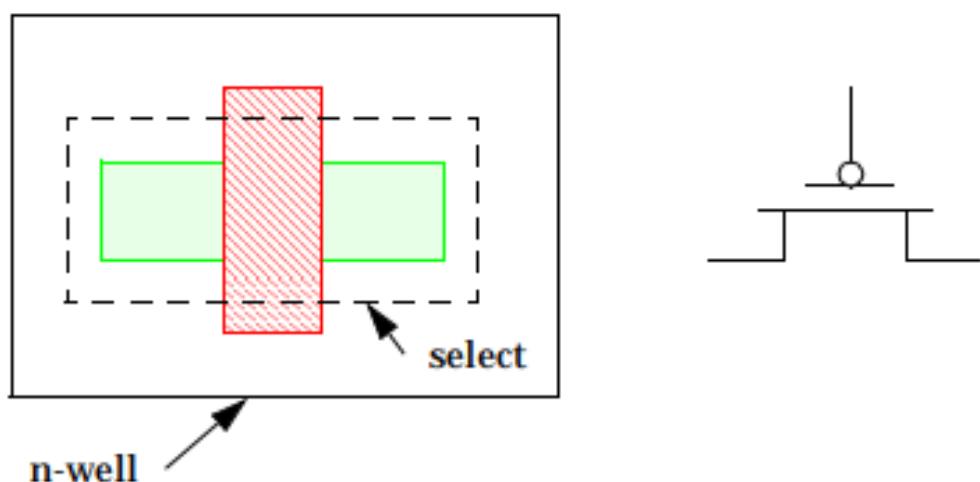
Layer	Representation	Color Convention (EECS 105)
n-well		purple
active		green
select (p ⁺)		brown
polysilicon		red
metal		blue
contact		black

MOSFET Layout

- *polysilicon crossing active* results in an NMOS device:

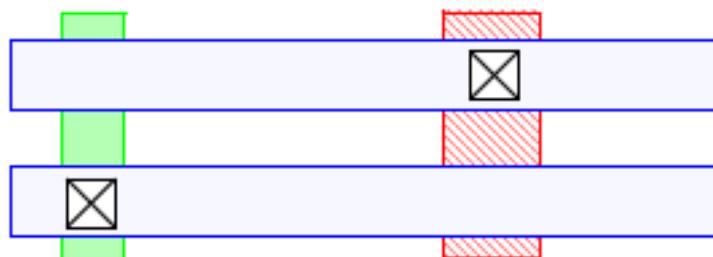


- PMOS devices are placed in *n-wells*:

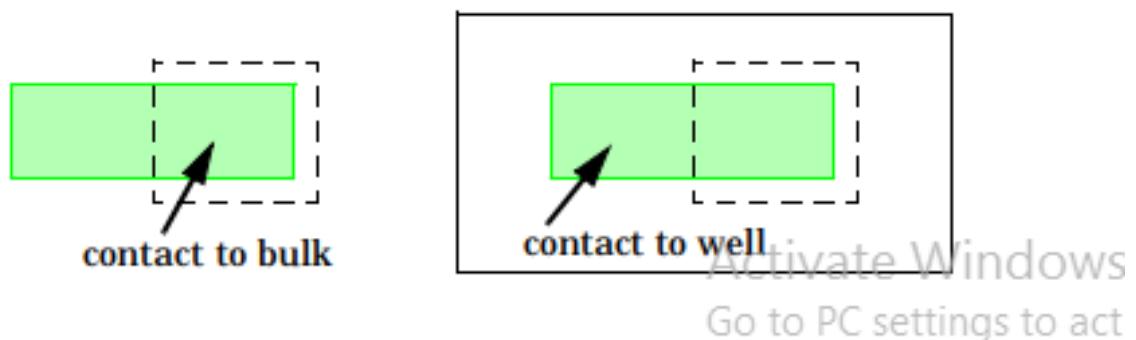


Electrical Connectivity

- *active*, *polysilicon*, and *metal* can be used for interconnects (wires)
- *metal* has much lower resistivity than either *active* or *polysilicon*
- *metal* is separated from *active* or *polysilicon* by an (insulating) oxide; a *contact* is needed for electrical connections between these layers



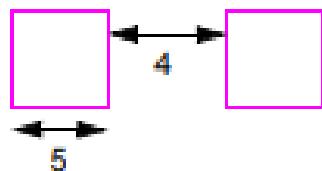
- *active* and *polysilicon* cannot be connected directly (without *metal*)
- use p-doped *active* (*select mask*) as contact to the bulk
- use n-doped *active* (no *select mask*) as a contact to *n-wells*



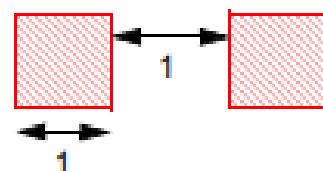
Layout Rules (EECS 105 Technology)

minimum dimensions and separations (in mm, not to scale):

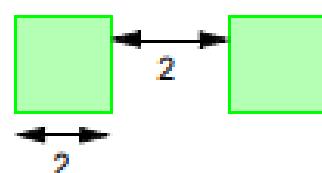
n-well



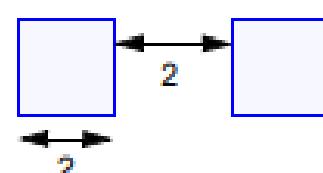
polysilicon



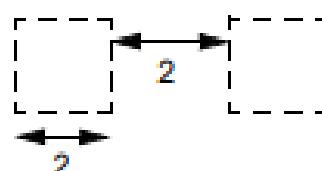
active



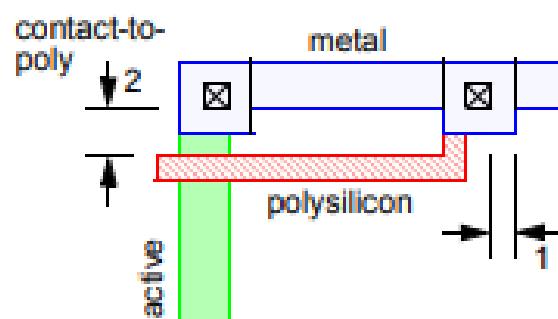
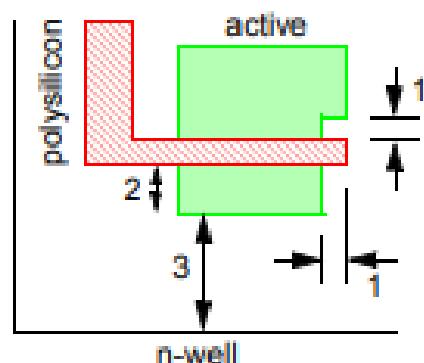
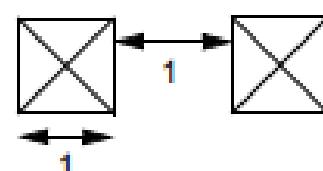
metal



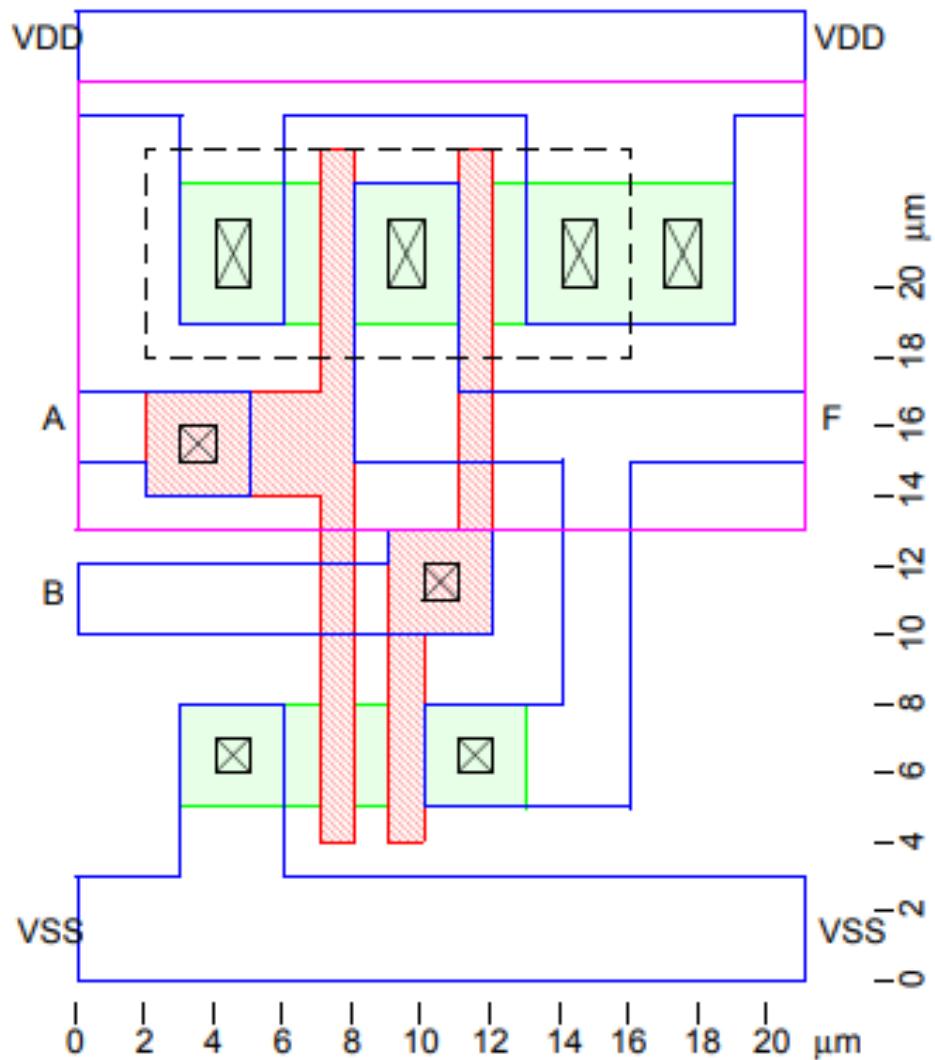
select



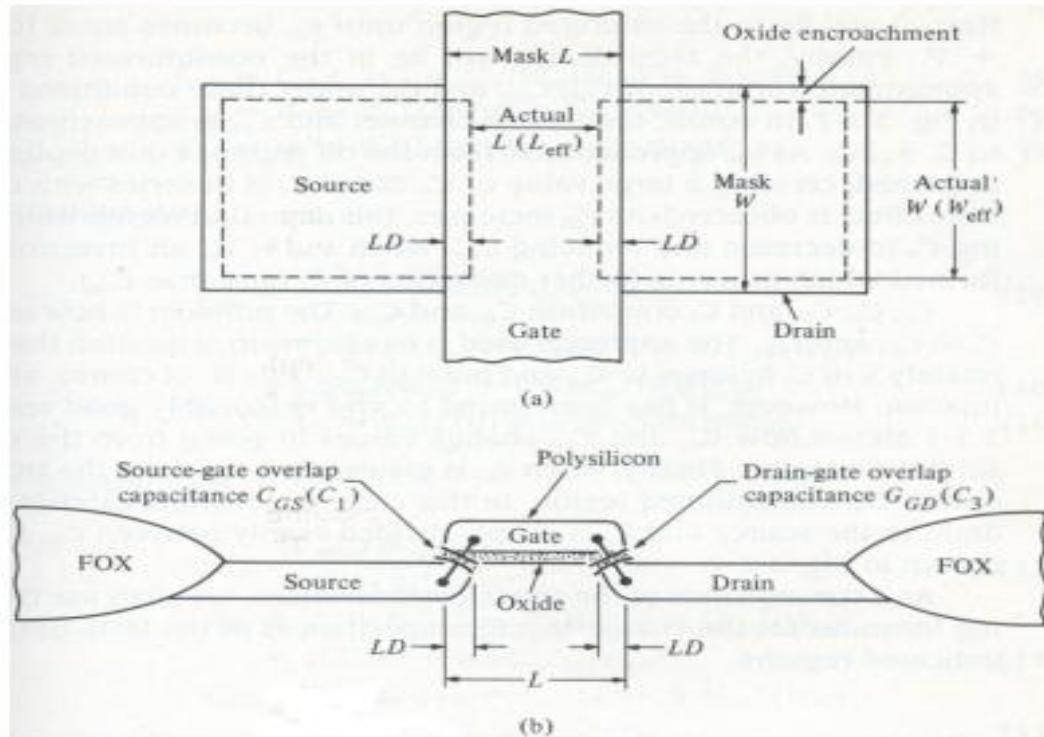
contact



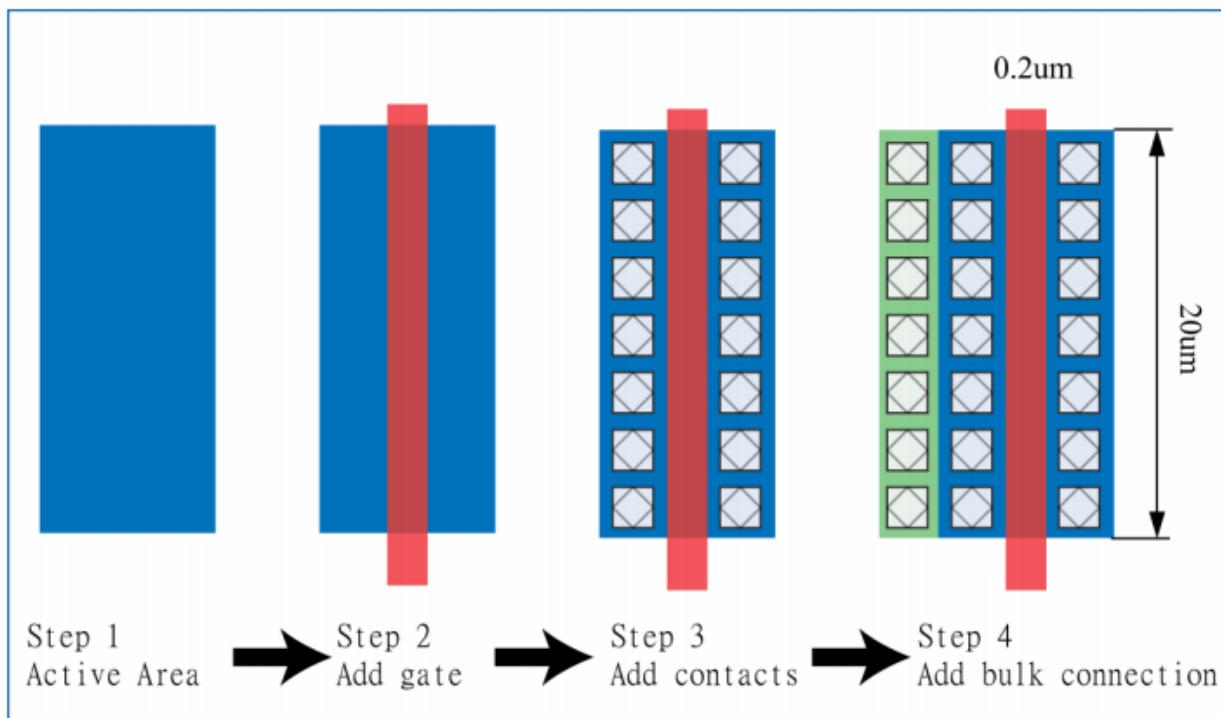
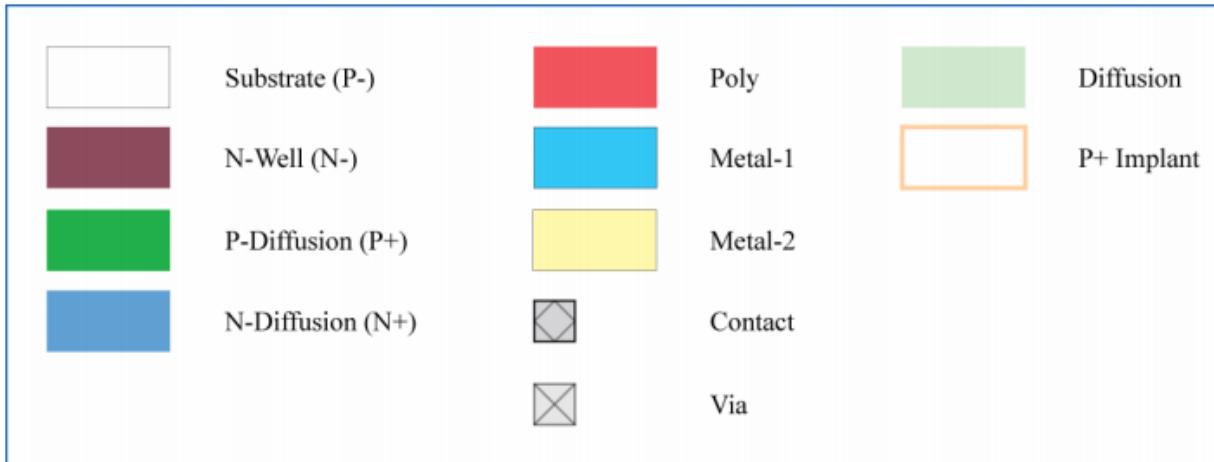
Layout Example

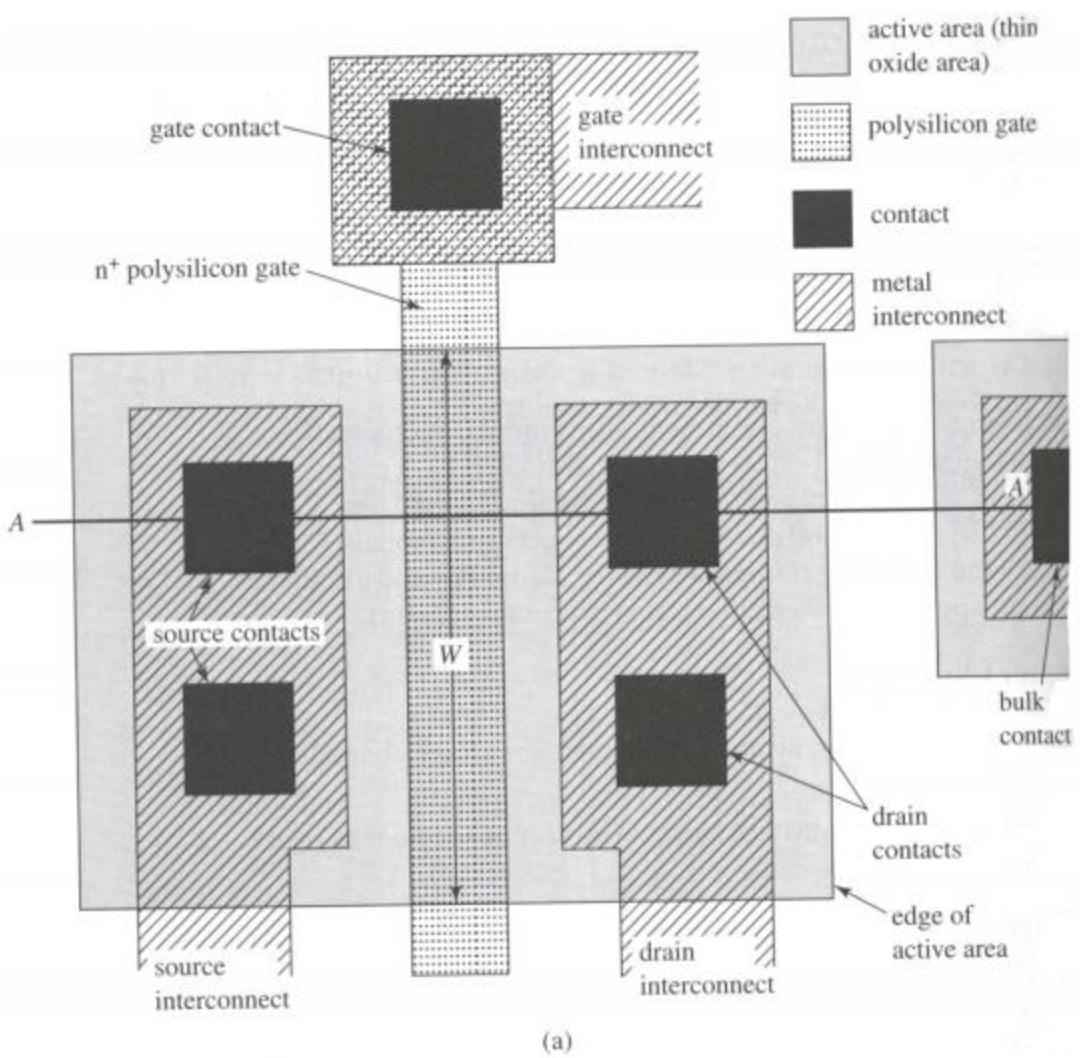


Effective channel length : (it decides the technology name)



MOSFET layers in layout

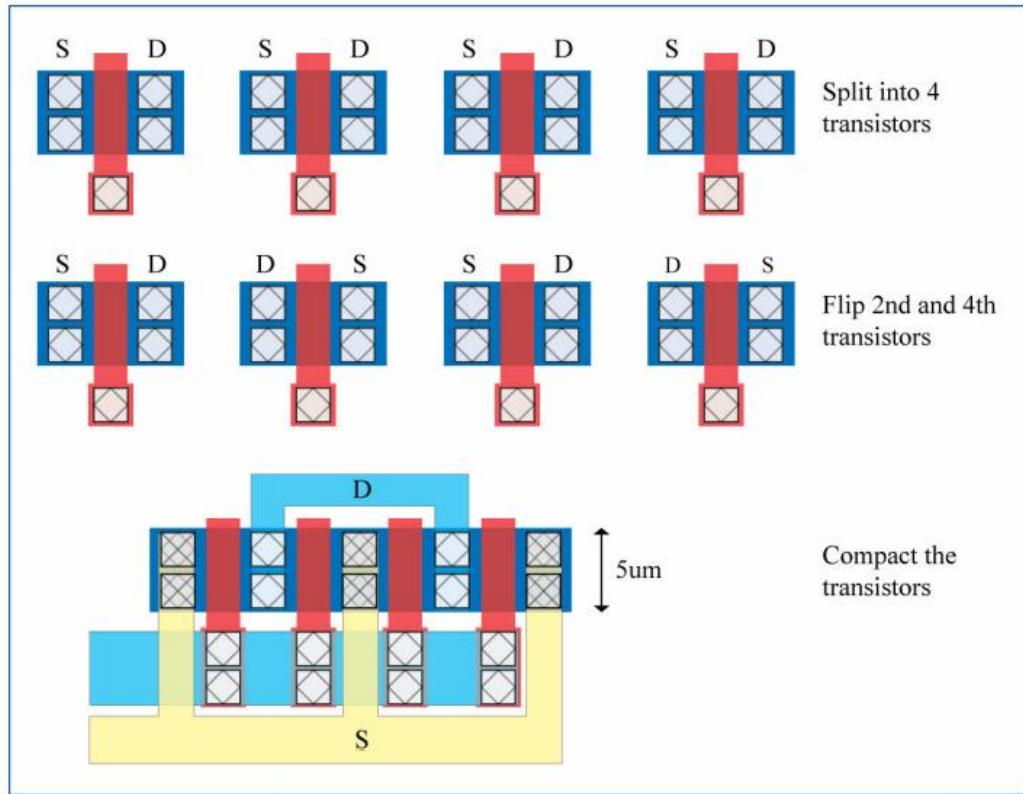




(a)

gate oxide

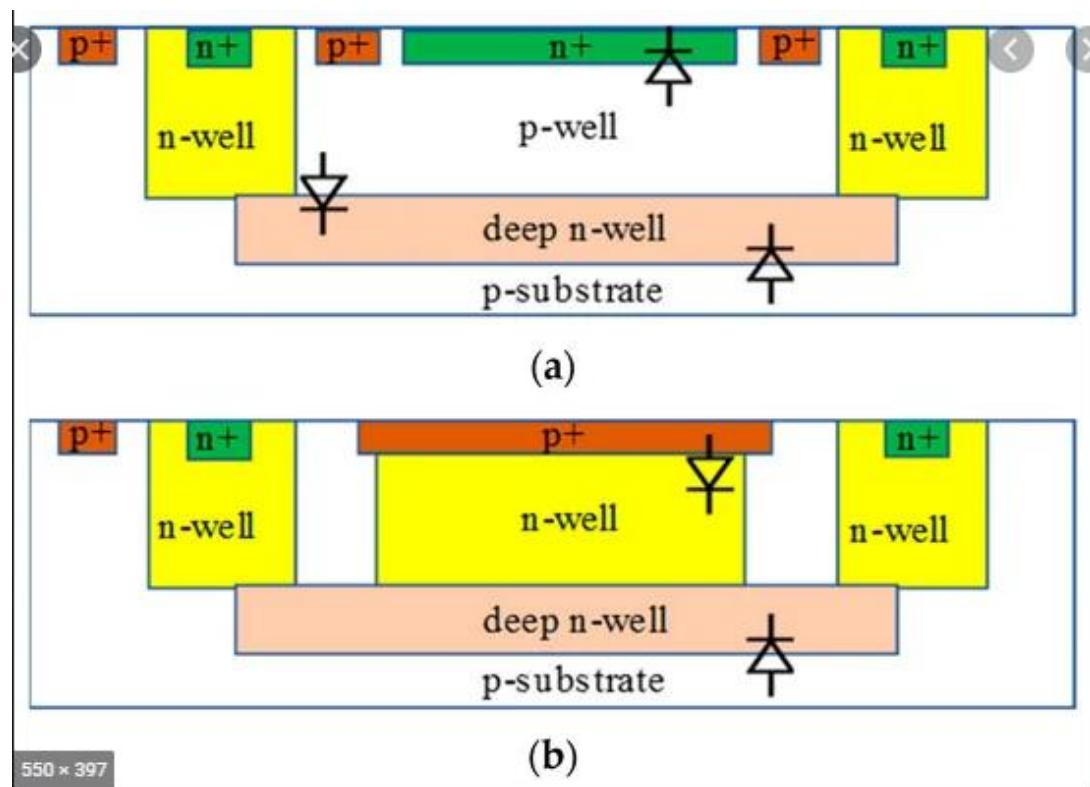
(b)

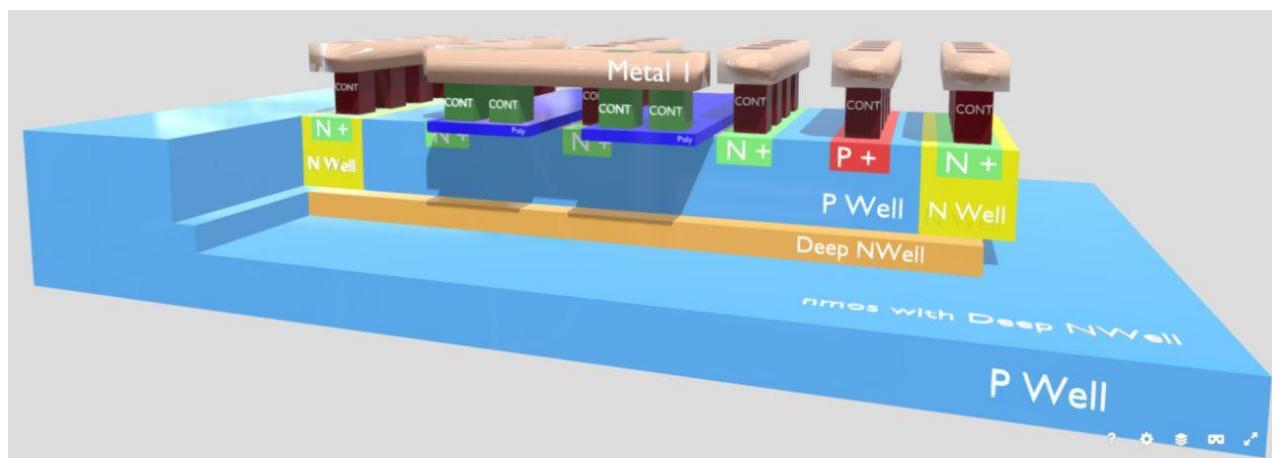
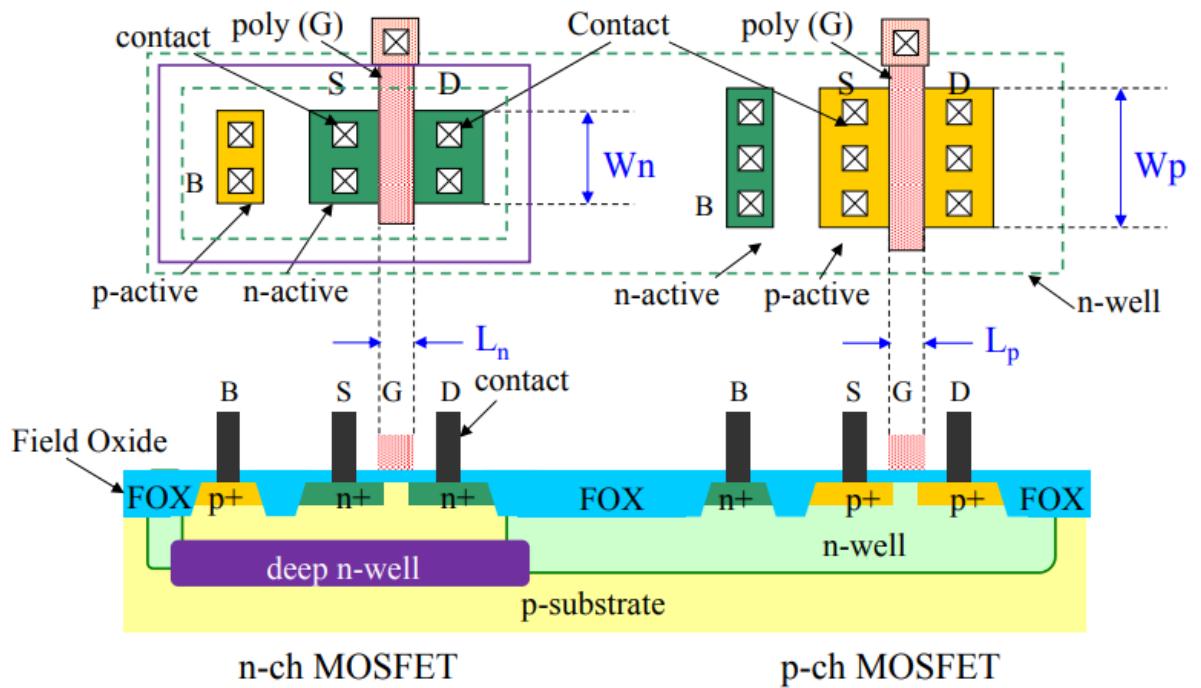


DNW layer (It is used to isolate P-well):

Generally DeepNwell is used to isolate NMOS from the substrate of other NMOS.

Mostly this layer is used to substrate noise coupling injected by Digital Logic in Mixed Signal environment through bulk or substrate.





Video link :<https://youtu.be/fpKx4hj8MYM>

3D view link: <https://skfb.ly/oGTIz>

5.LAYOUT INTRODUCTION

Poly :Polysilicon is a high purity, polycrystalline form of silicon.

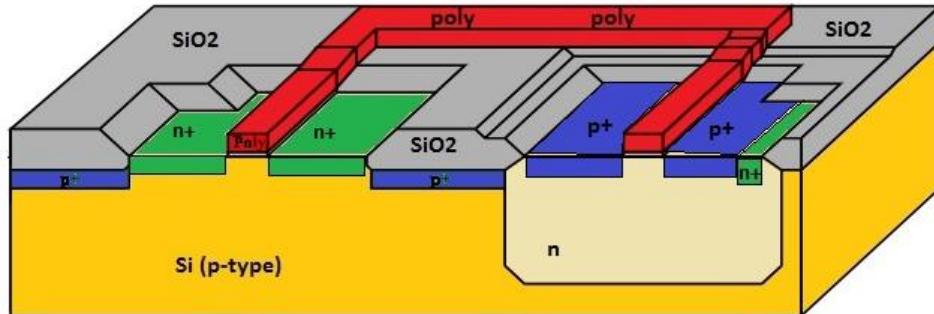
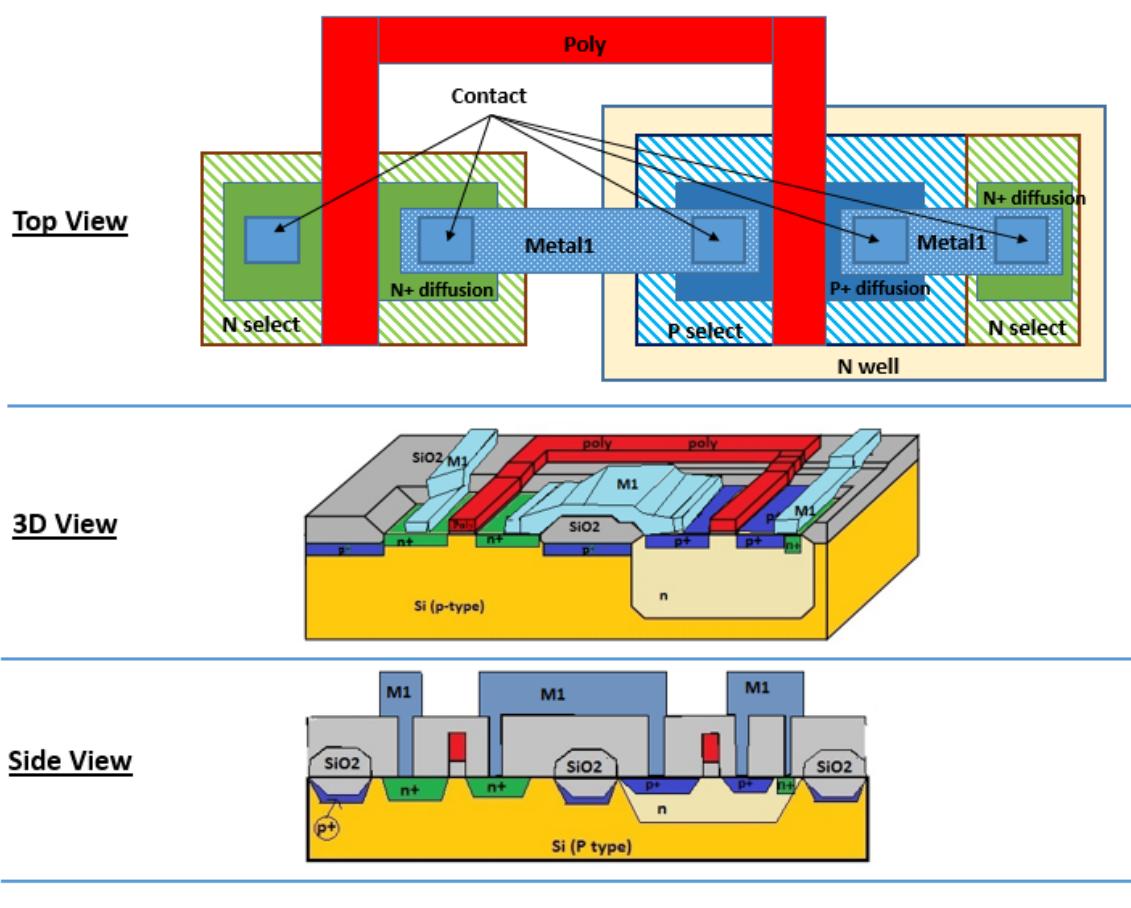


Fig : 3D view of the silicon wafer after depositing the poly layer



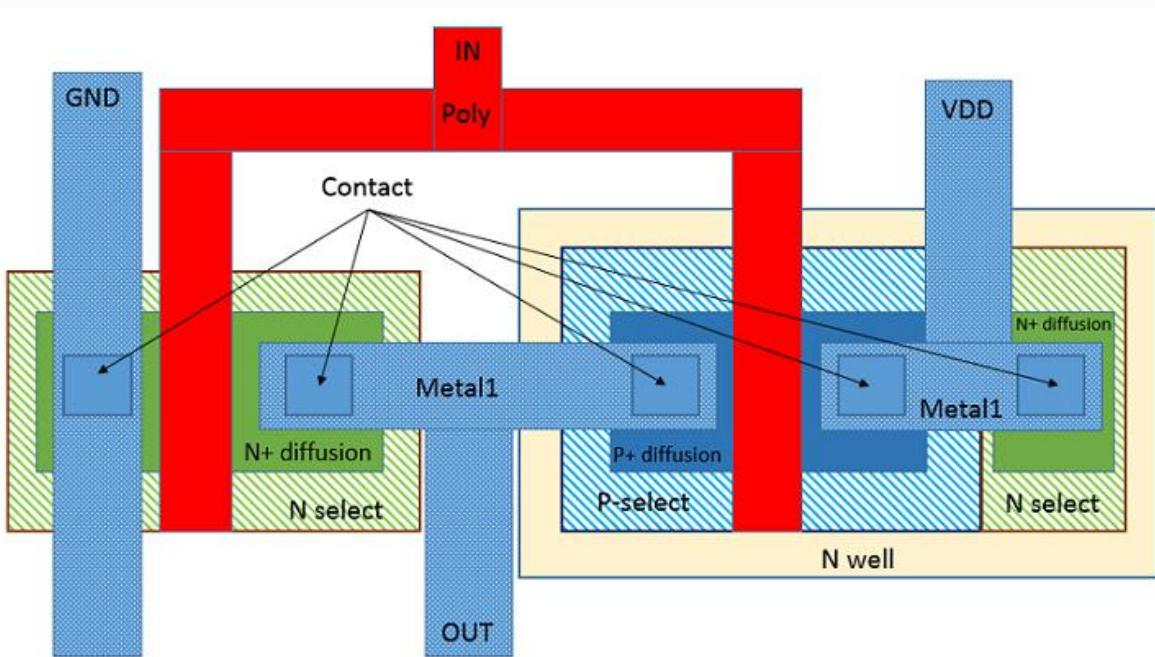


Fig :Layout view of CMOS inverter

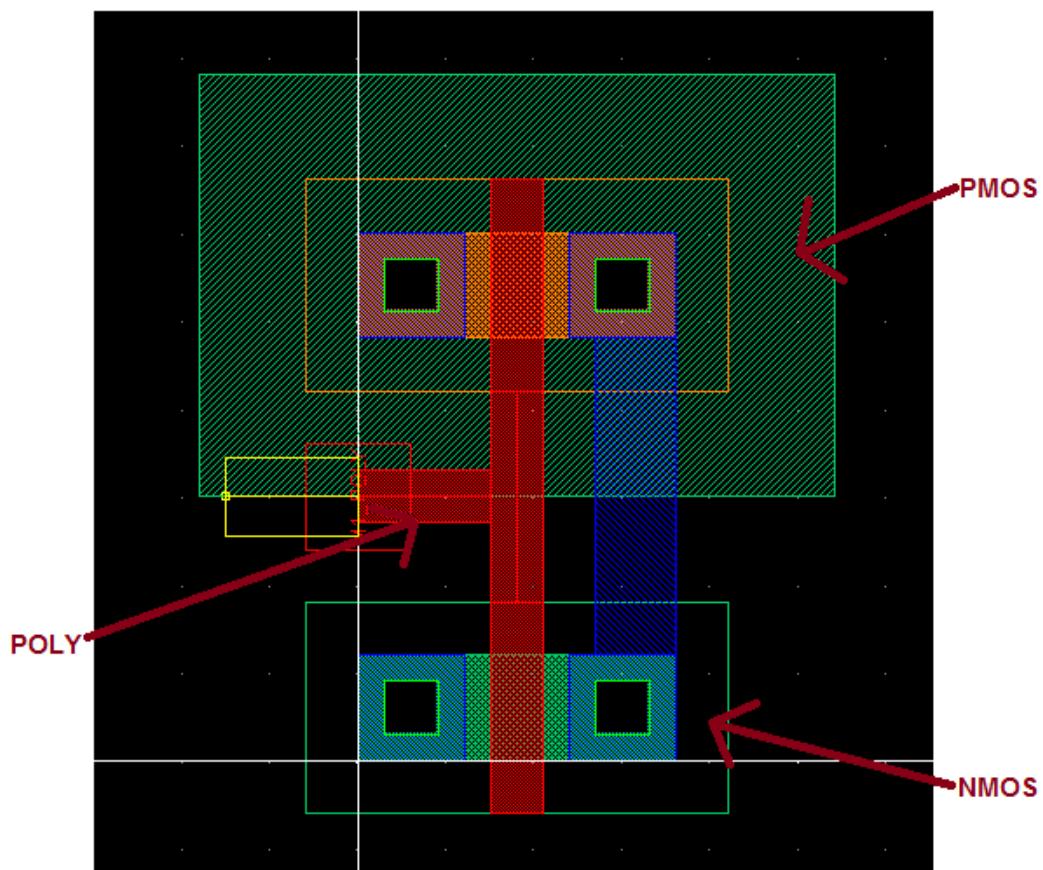
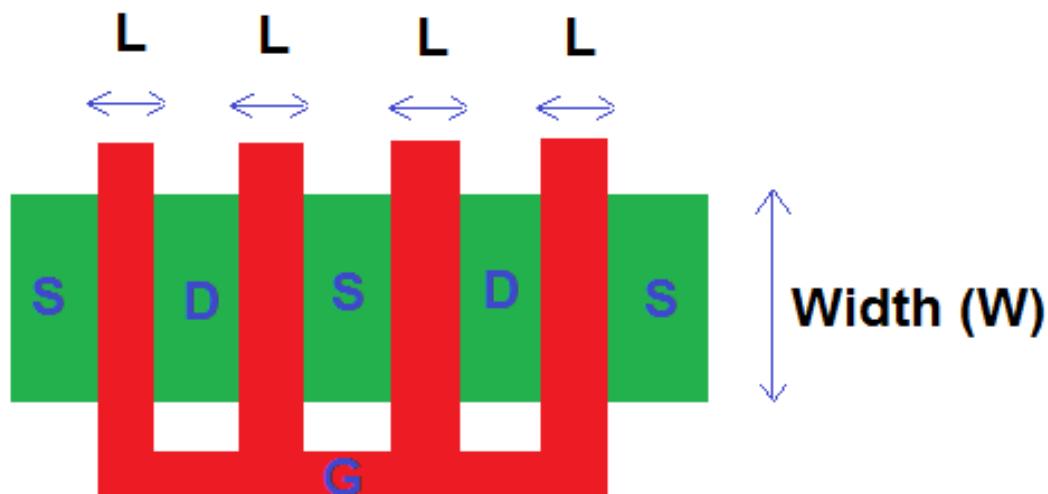
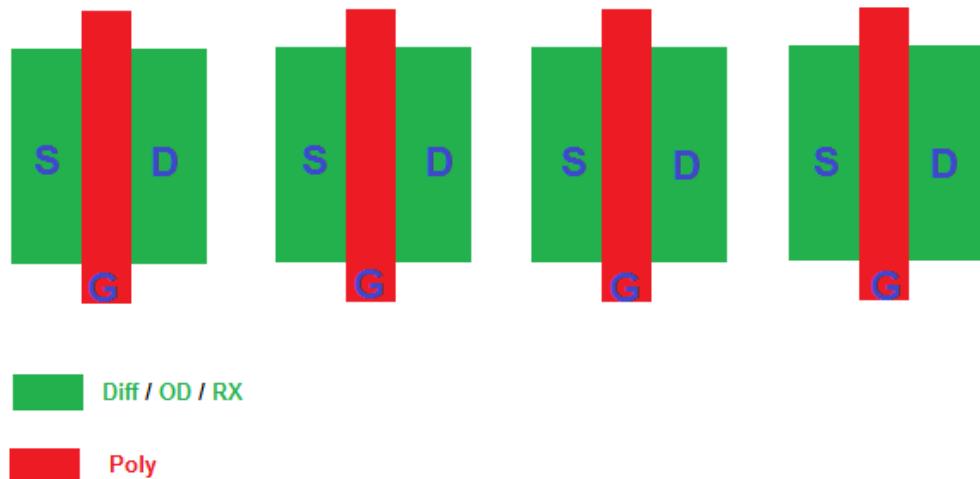


Fig : Layout view of two Active Region (PMOS and NMOS)

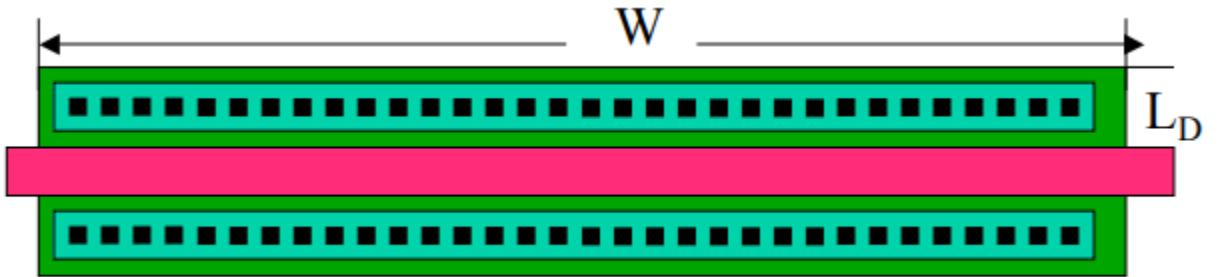
Device Fingers: Simply it is a multiplier or parallel connected device. If 4 parallel connected device is there so we can say the device have 4 fingers.



$$L_{\text{eff}} = L, W_{\text{eff}} = 4W$$

Fig : Four finger devices with drain and source merge

- ❖ Why should increase fingers and reduction of width is good?

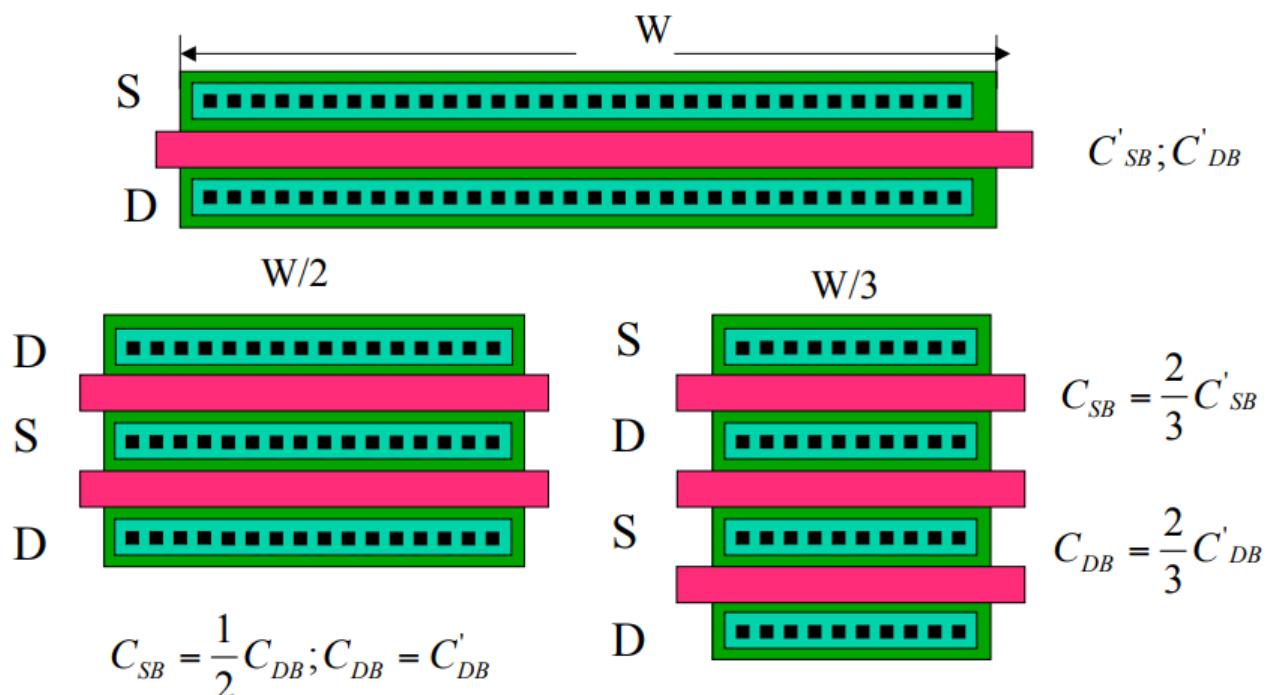


Capacitance diffusion substrate

$$C_{SB} = C_{DB} = (W + 2l_{diff})(L_D + 2l_{diff})$$

Resistance of the poly gate

$$R_{gate} = L_{gate} \cdot R_{sq, poly}$$



$$C_{SB} = \frac{1}{2} C_{DB}; C_{DB} = C'_{DB}$$

Metal Routing:

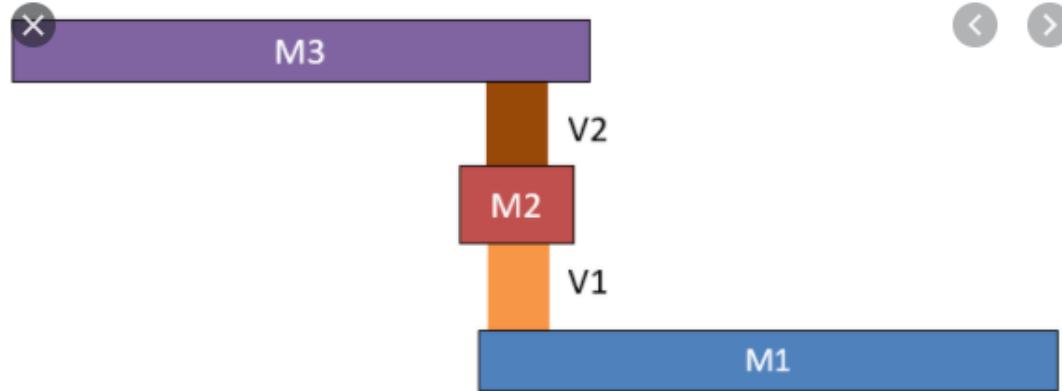


Fig : Connection in between Metal 1(M1) to Metal3 (M3)

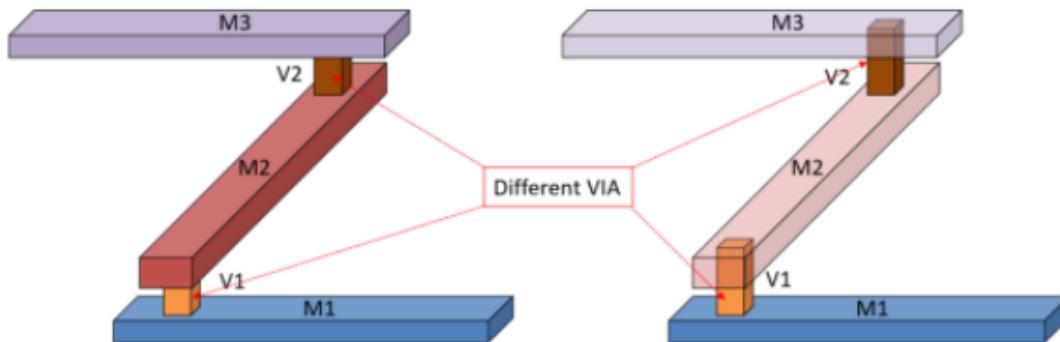


Fig :3D view of VIA and Metal connection

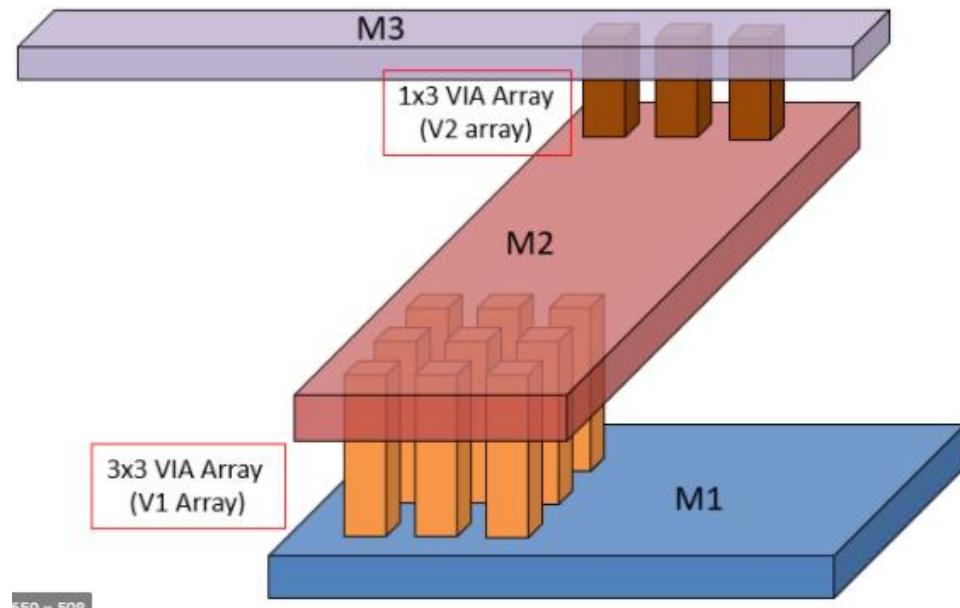


Fig :3D view of VIA and Metal connection

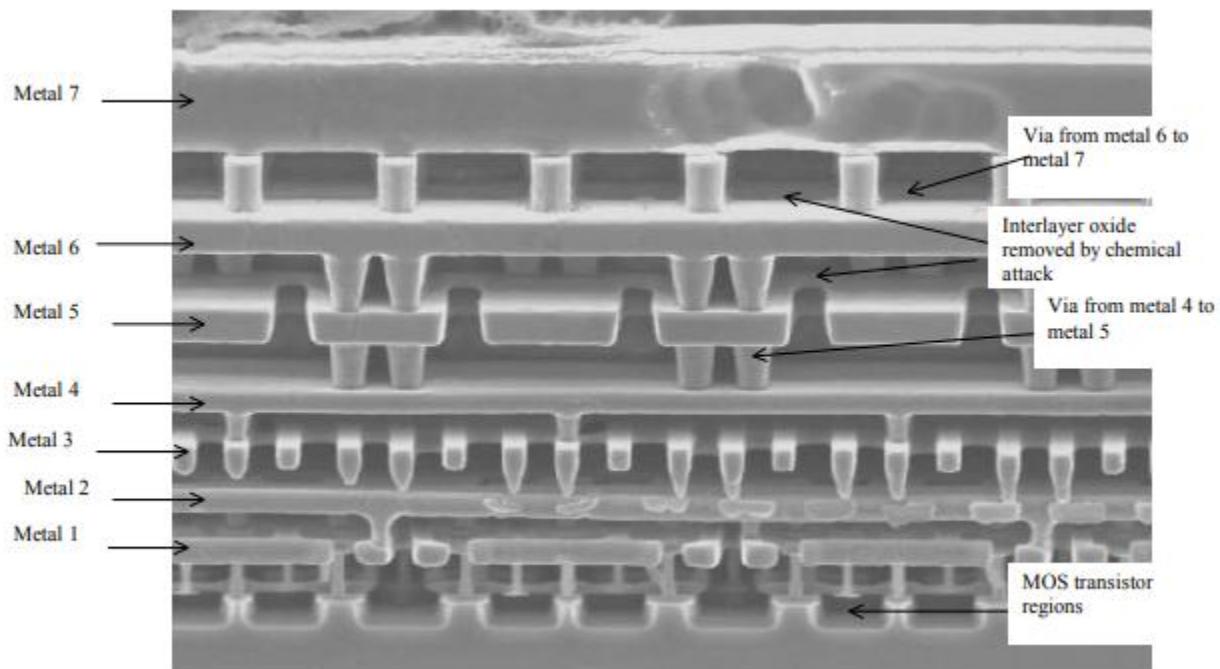
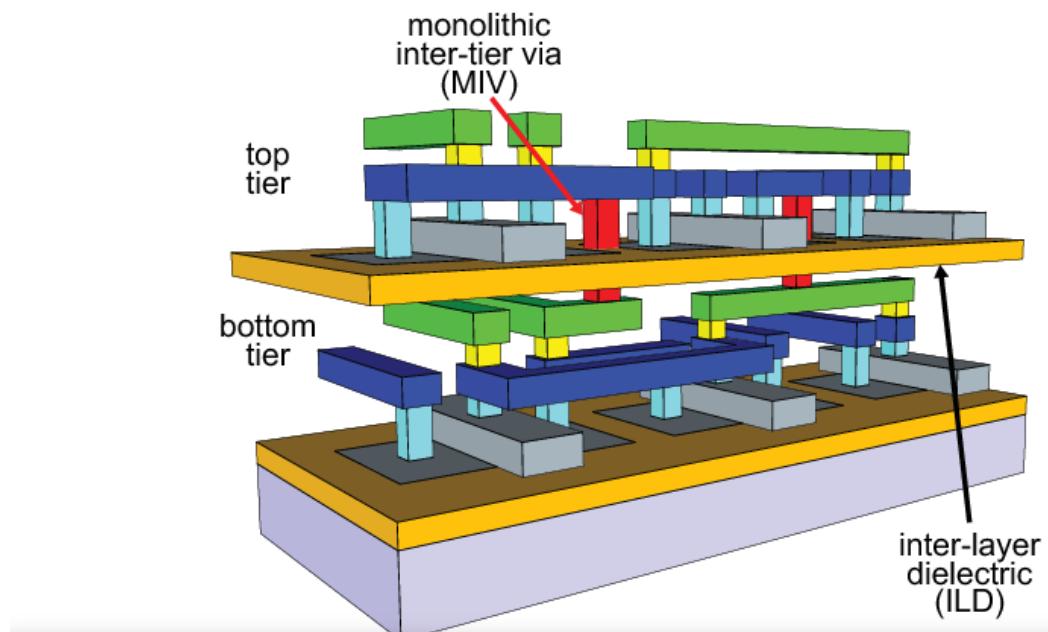
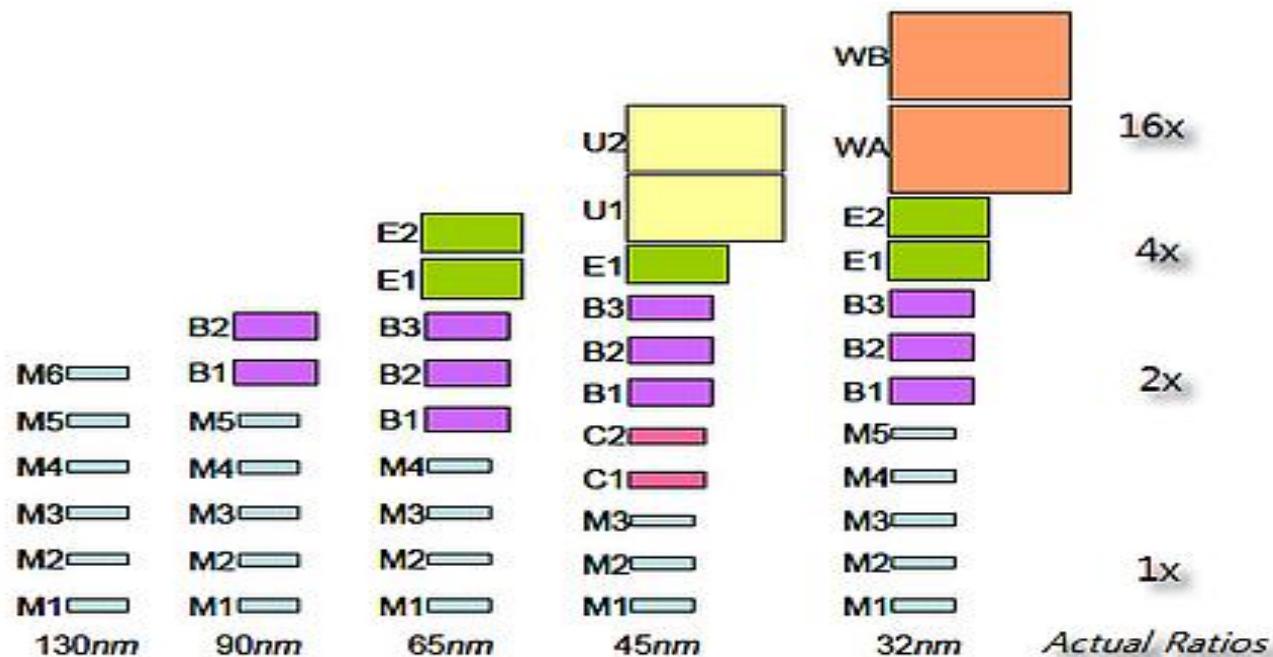


Fig :Cross-section of a $0.12\mu\text{m}/120\text{nm}$ technology (Courtesy Fujitsu)

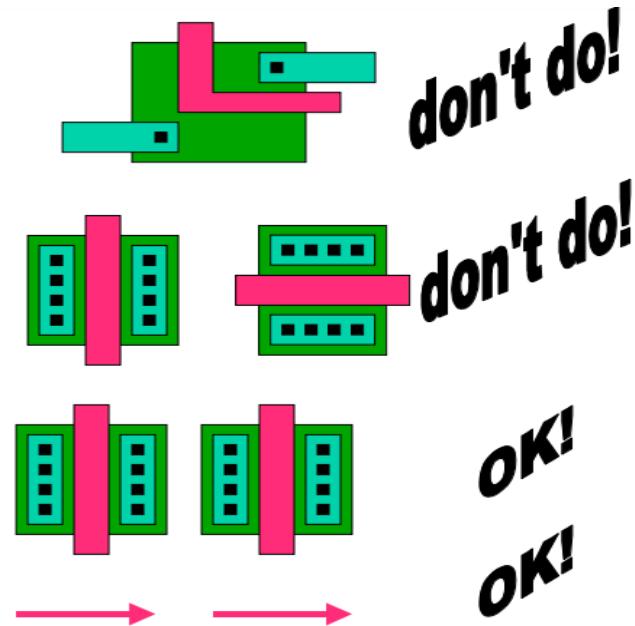
- **Metal ratio:** It is define with respect to first metal (metal1)

Assume that metal1 height is 1um and metal2 is 2Xso metal2 height is 2um.
Metal1 height is very with respect to technology.

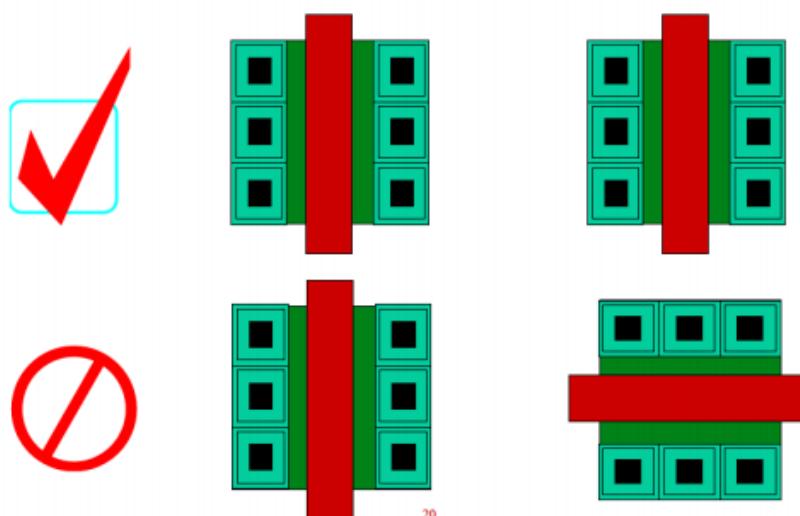


6.BASIC IMPORTANCE OF LAYOUT

Layout Matching :

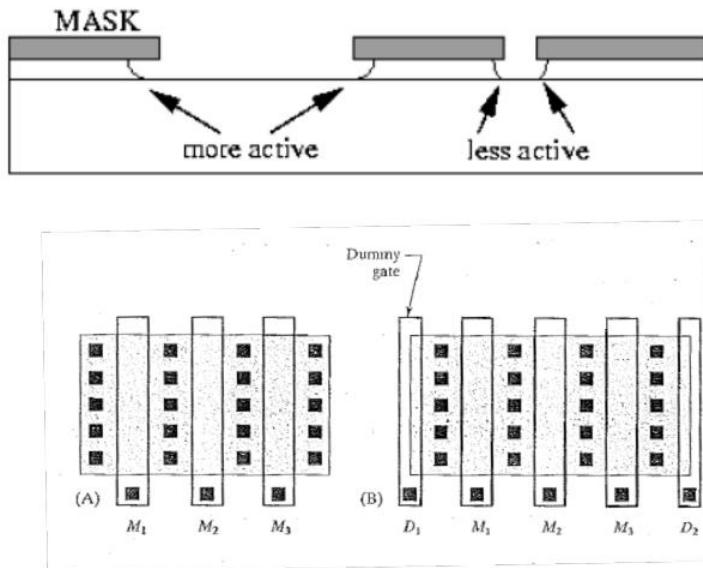


- Transistor orientation is very important

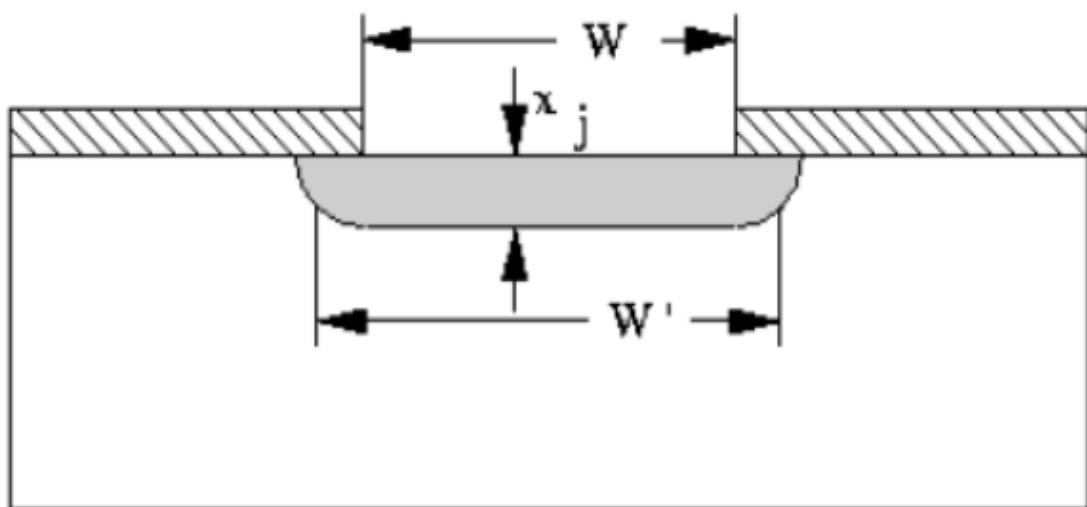


20

- **Etch Effect**: Poly silicon does not always etch uniformly.
 - Large openings etch faster than small openings in mask.
 - Solution is to use dummy structures.



- **Diffusion Effects** : Diffusion widens implanted region can affect doping of neighboring devices.
Solution : Increase distance and use dummy structures that affect all transistors the same.



- **Stress Effects :**

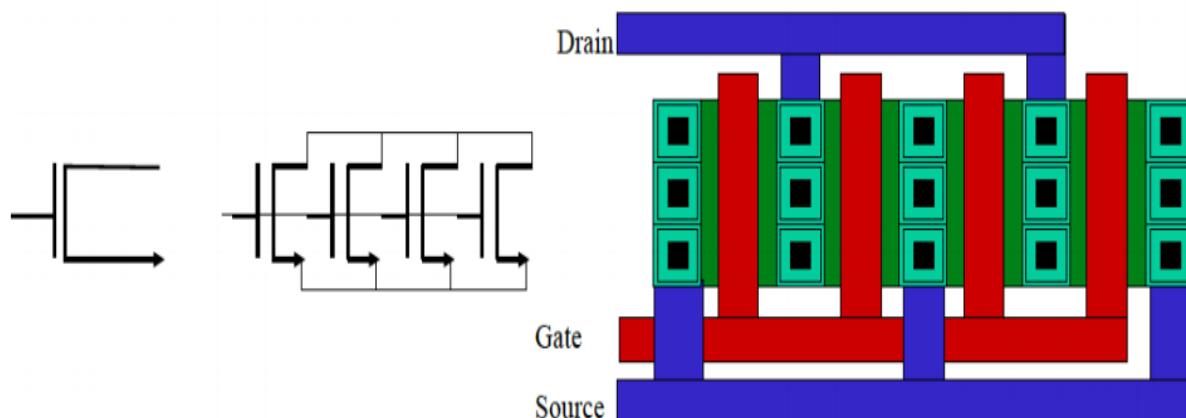
- a. The fabrication under high temperatures may leave residual stresses in chip.
- b. Packaging can cause stress in chip.

- Solution :**

- a. Keep critical matched devices in center of chip or on enterlines
- b. Avoid using corners for matched devices

- **Dealing with Large Transistors :**

- a. These can be split into many parallel fingers.
- b. Contact space is shared amongst transistors.
- c. Parasitic capacitance is reduced.



Transistor Matching :

- **Why Matching is needed:**

- ✓ What if your headphones started playing songs with more volume in left than the right? Sounds bad right. So the driver sensitivity needs to be matched for both left & right side.
- ✓ Matching in context of analog layout is an act of creating a similar environment for a set of transistors, so that they exhibit similar electrical properties. (i.e, trans conductance, current gain, drain capacitance etc.).
- ✓ Actually due to **process variation** in fabricating the transistors (i.e like non uniform doping, oxide thickness, etching etc.), no 2 transistors have exactly the same electrical properties. By applying techniques such as *Interdigitizer* & *Common centroid*, it is possible to match transistors reasonably.
- ✓ In a differential amplifier, matching would greatly enhance the performance by improving the ability to reject noise in its differential input. Therefore it “Common mode rejection ratio(CMRR)” would increase. In a current mirror, matching would accurately mirror the reference current.

- **Device matching type:**

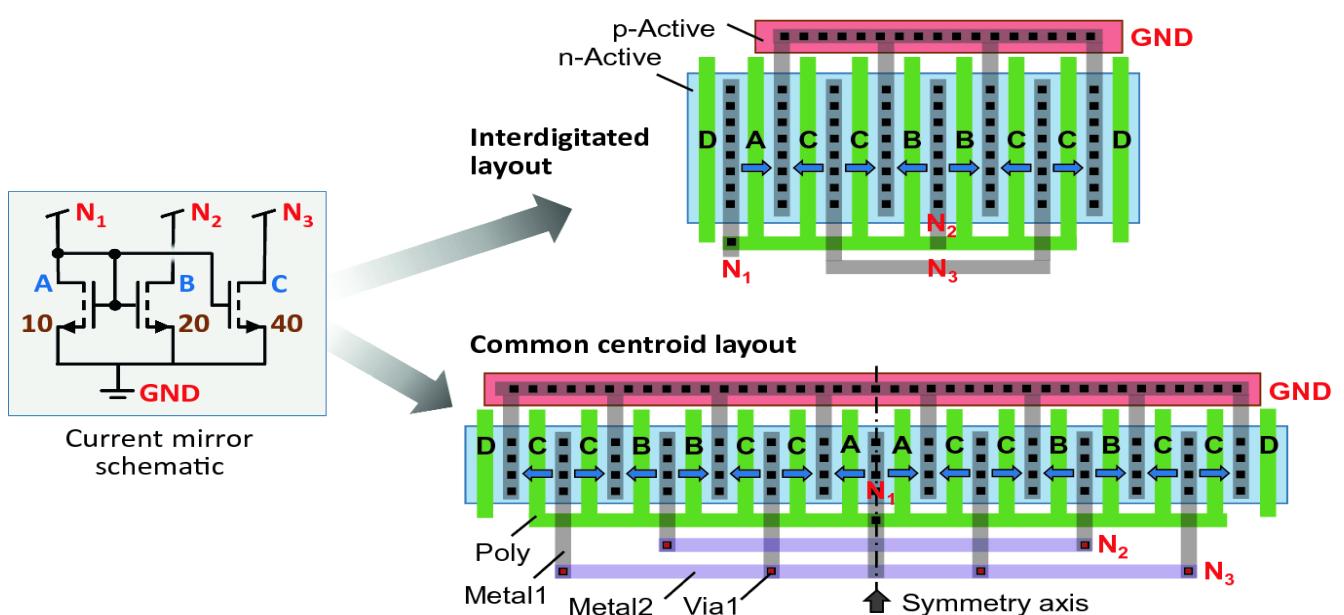
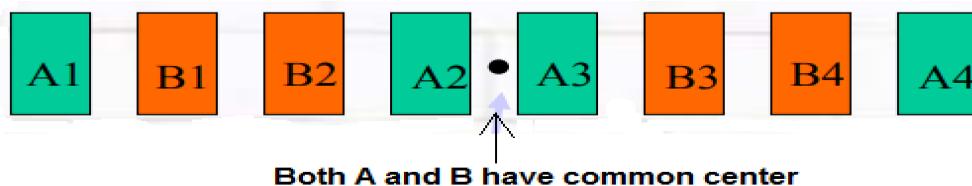
1. Inter-Digitization Techniques
2. Common Centroid

- **Inter-Digitization Techniques:**

- We have to match two components A and B (A and B can be anything like Mosfet, transistor, resistor, and capacitor).
- Lets split A and B into 4 small components i.e. A1-A4 and B1-B4
- Inter-digitization technique: Placing alternate components.

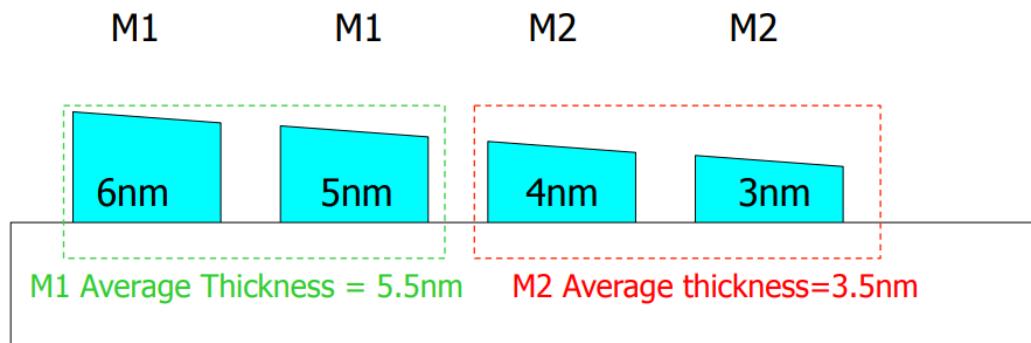


Common centroid technique: Placing components such that both components have same centroid.

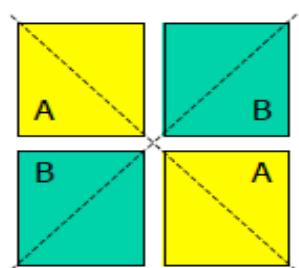
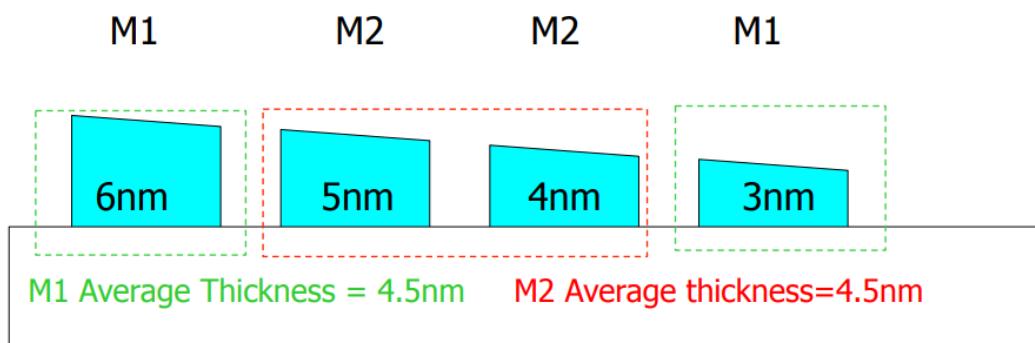


- Common Centroid Layouts :

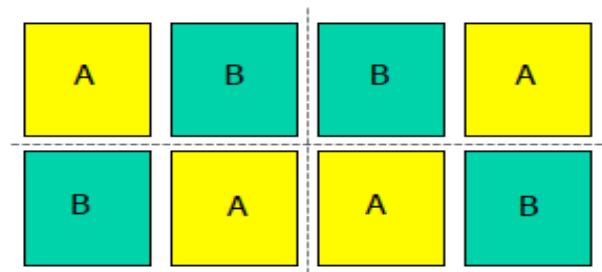
Matching won't be good!!!



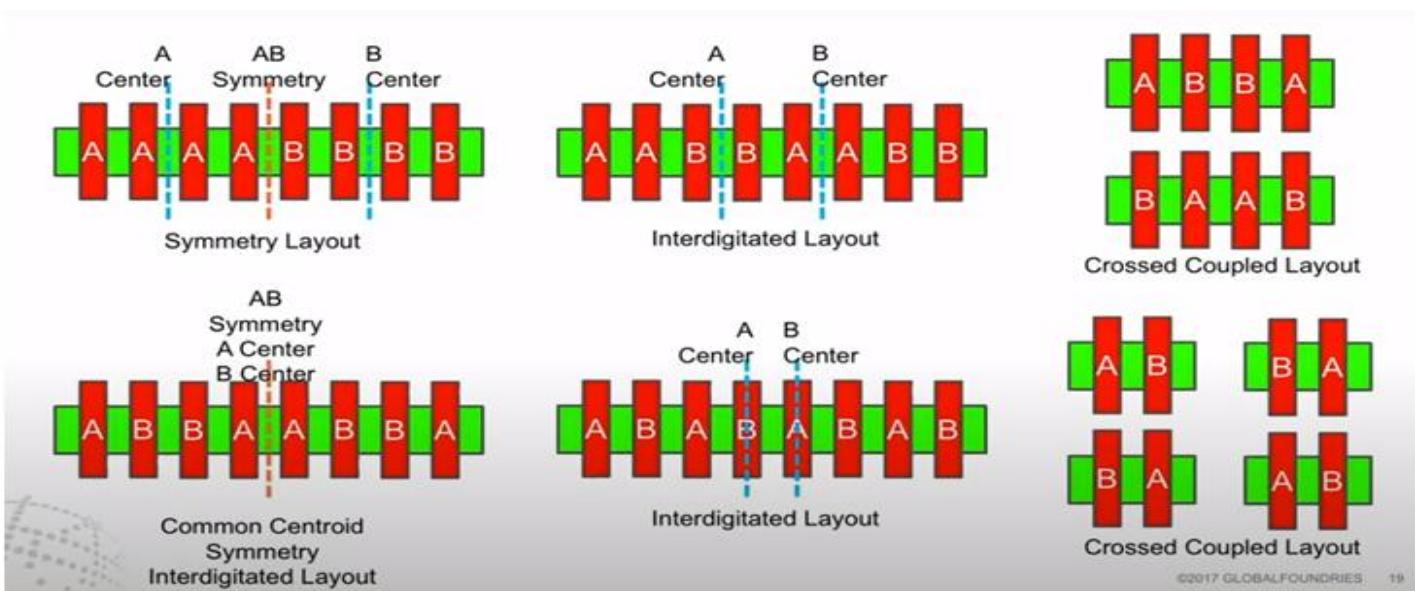
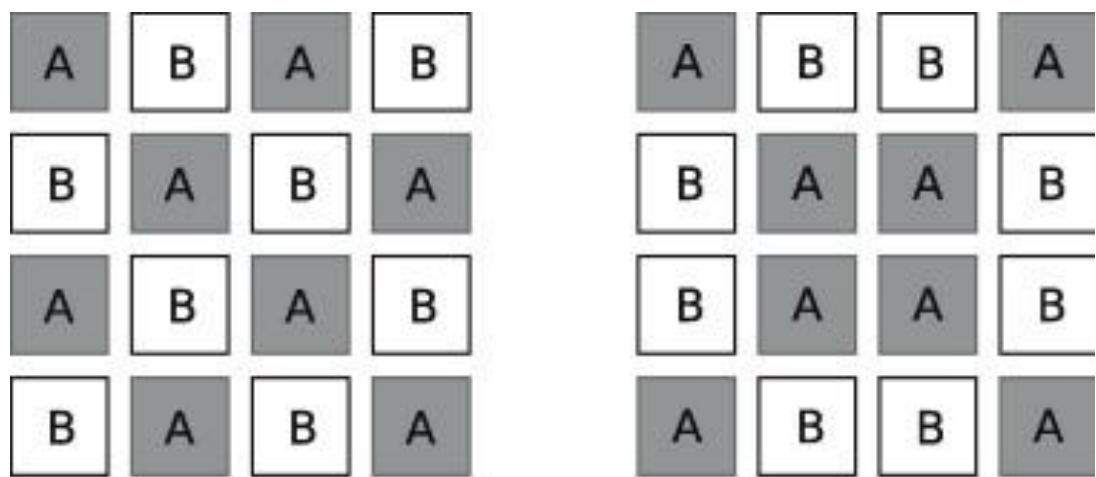
- Solution: Break and distribute parts of a transistor so as to cancel out the effects of oxide / doping gradient profiles.



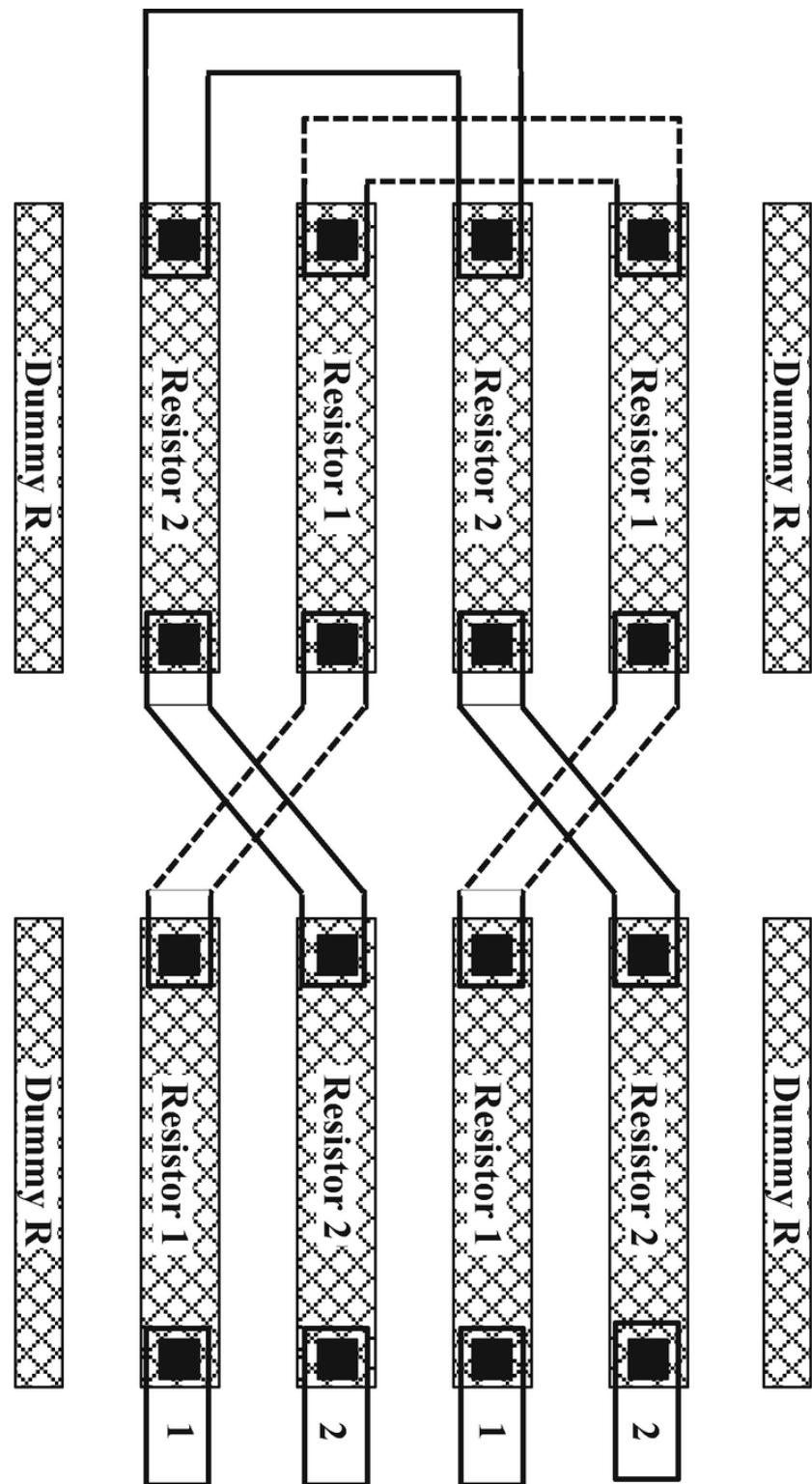
Cross coupling



Tiling (more sensitive to high-order gradients)



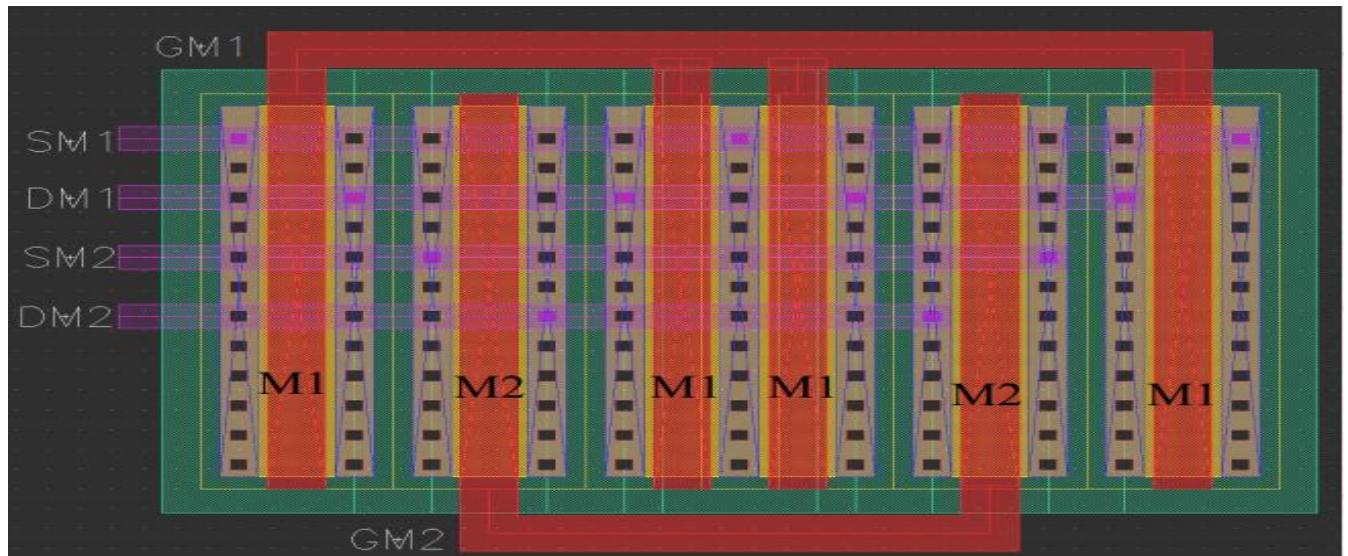
Resistor matching:



- Two Example of common centroid technique

1. Matching between M1 and M2 of differential amplifier

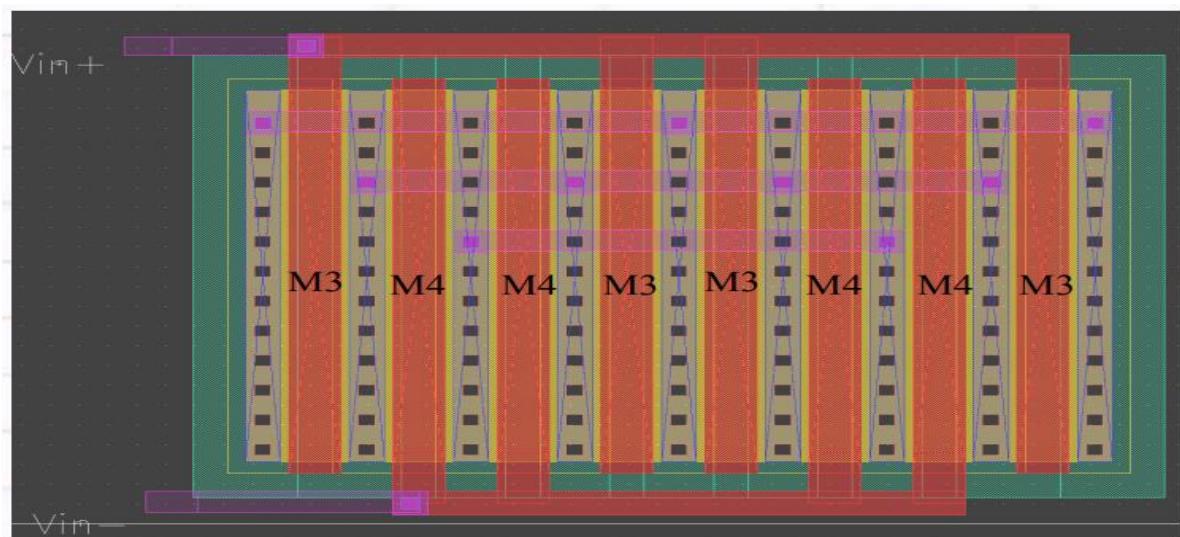
$$(W/L)_{M1} = 2 (W/L)_{M2}$$



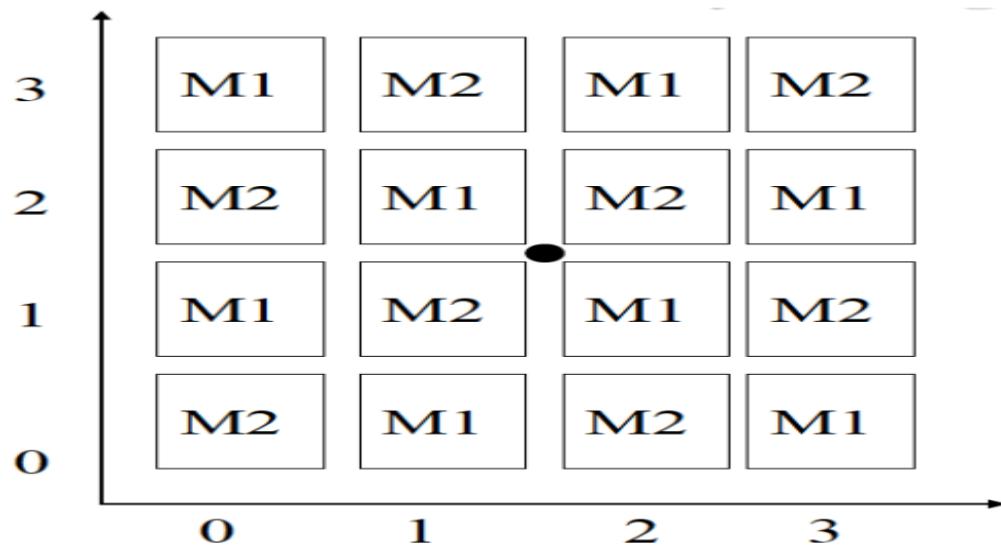
An alternative approach of M1 M1 M2 M2 M1 M1 can also be used!!!

2. Matching between M3 and M4 of differential amplifier

$$(W/L)_{M3} = (W/L)_{M4}$$



- Averages process variations :



❖ **Why Special attention on Matching?**

- ✓ A large variety of analog circuits rely on matching of transistors.
Circuits like differential pair rely on gate tosource voltage matching while current mirrors rely oncurrent matching.
- ✓ Most integrated resistors and capacitors have atolerance of about 20% to 30%. But ratio of two similarcomponents can be controlled to a tolerance of 15 or even0.1% by proper matching of the components.0 1% by proper matching of the components.

- Reasons for Mismatch :

Mismatch in integrated circuits are generally of two types :

- a. **Random mismatches** due to microscopic fluctuations in dimensions, doping, oxide thickness and other parameters that influence component values
- b. **Systematic mismatches** which are caused by :
 - Process biases
 - Mechanical stress
 - Temperature gradients
 - Polysilicon etch rates etc

- **Rules for MOS transistor matching :**

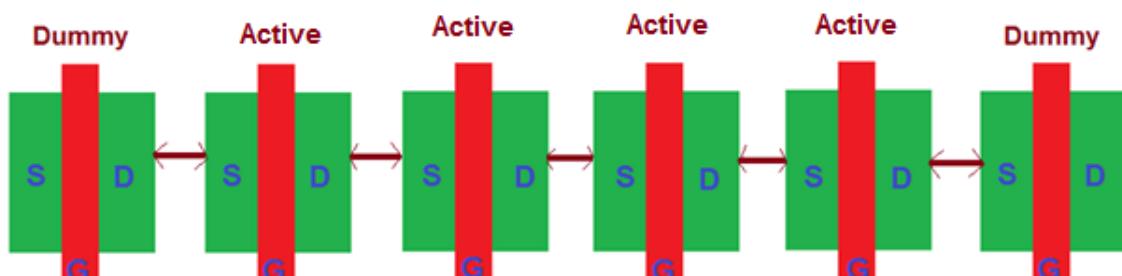
- ✓ Place transistors in close proximity.
- ✓ Orient transistors in the same direction.
- ✓ Keep the layout of the transistors as compact as possible
- ✓ Whenever possible use Common centroid layouts.
- ✓ Place transistors segments in the areas of low stress gradients.
- ✓ Place transistors well away from the power devices.
- ✓ For current matching keep overdrive voltage large.
- ✓ For voltage matching keep overdrive voltage smaller.

- **Rules for resistor and capacitor matching :**

- ✓ Construct matched resistors of same type.
- ✓ Make matched resistors of the same width.
- ✓ Orient matched resistors in the same direction.
- ✓ Place matched resistors in close proximity.
- ✓ Place the matched resistors in such a way that their centroids coincide i.e. interdigitate arrayed resistors.
- ✓ Place dummies on either end of the resistor array.
- ✓ Connect matched resistors to cancel thermoelectric effects.

Diffusion/OD Break matching devices:

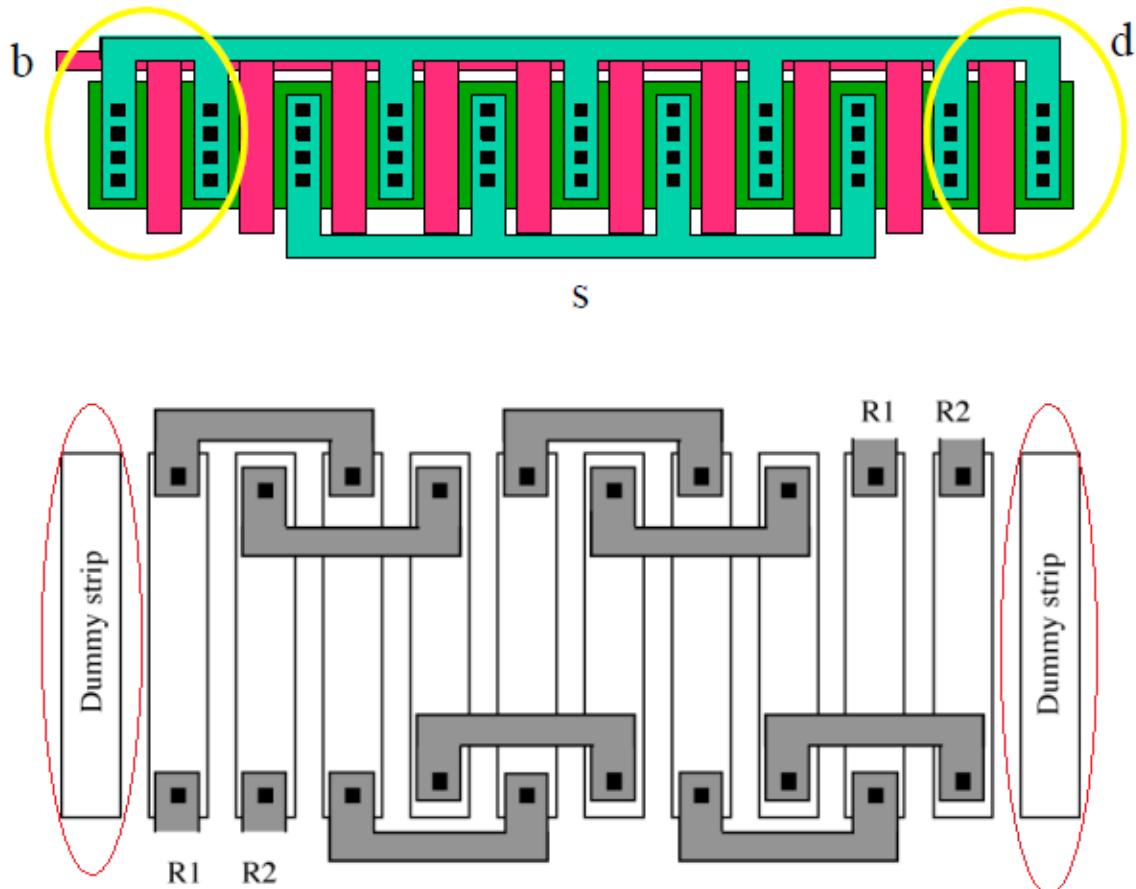
All device (MOS, resistor, capacitor...) should be placed in identical space and all device should see same environment. As shown in below fig.



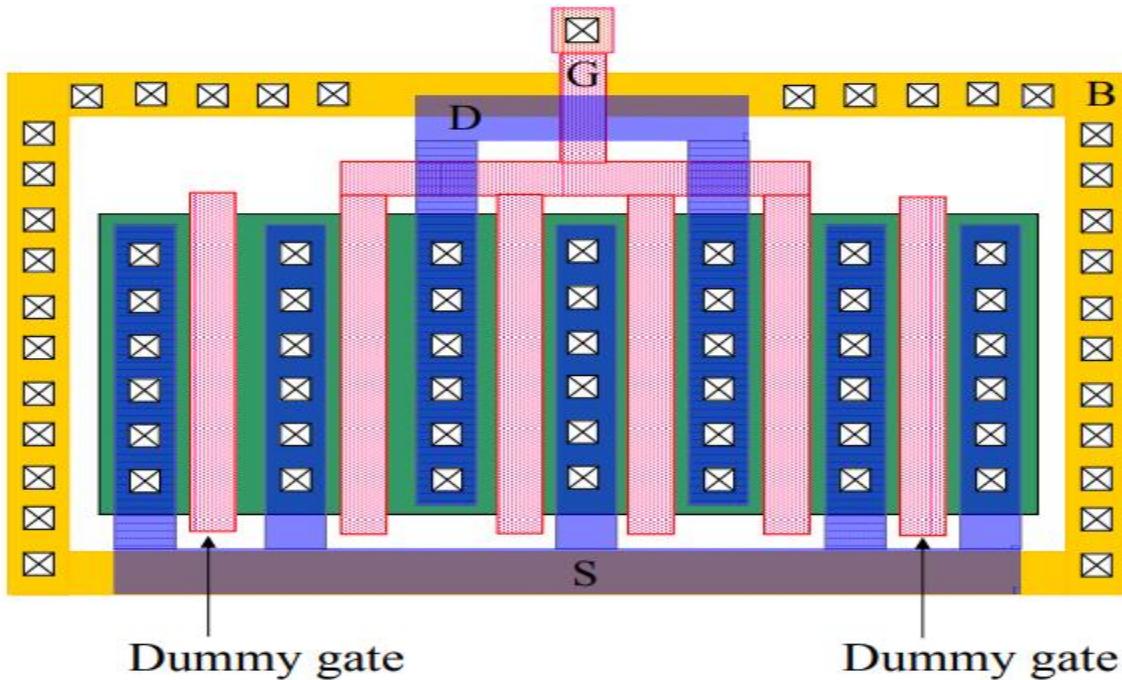
DUMMY:

- **Why Dummy Device is need :**

- ✓ To minimize the etch effects during fabrication. If they are not placed then it can so happen that one side is slightly more parameter ($V_{the.g}$) than the other transistor. Hence the matching will not be identical for all the device. For a good matched layout dummy device must be used. Generally dummy devices are added to resistors as well so that they surround the complete resistor.



The dummy pattern may be formed to reduce the production tolerance.



Document link for deeper concussion of matching:-

<https://core.ac.uk/download/pdf/84832221.pdf>

Video links for LayoutMatching:

1. need of layout matching : https://youtu.be/i_ejWt7wyEA
2. Layout Matching :https://youtu.be/92m_8kd0J0c
3. InterdigitaionVs Common Centroid:https://youtu.be/qkKQmtxG_7A
4. How to add dummy:<https://youtu.be/5ZRpTDAIpI8>

Current Mirror :

The simple current mirror can be obtained from the one transistor current source by using a second transistor in diode connection that generates the necessary gate-source voltage of the transistor in the output stage. The gate source voltage is set by the diode geometry and the injected input or reference current. Since the gates and the sources of the two transistors are connected together the gate-source voltage of the current source will be equal to the gate source voltage of the diode. The schematic of the circuit is given in below Figure.

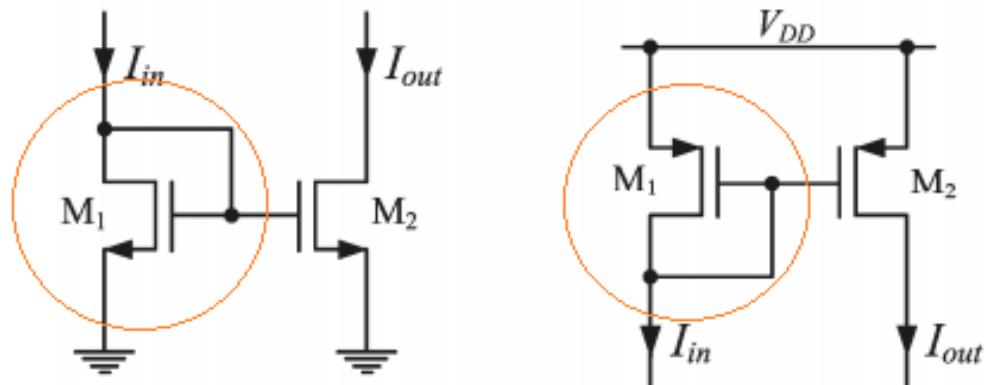
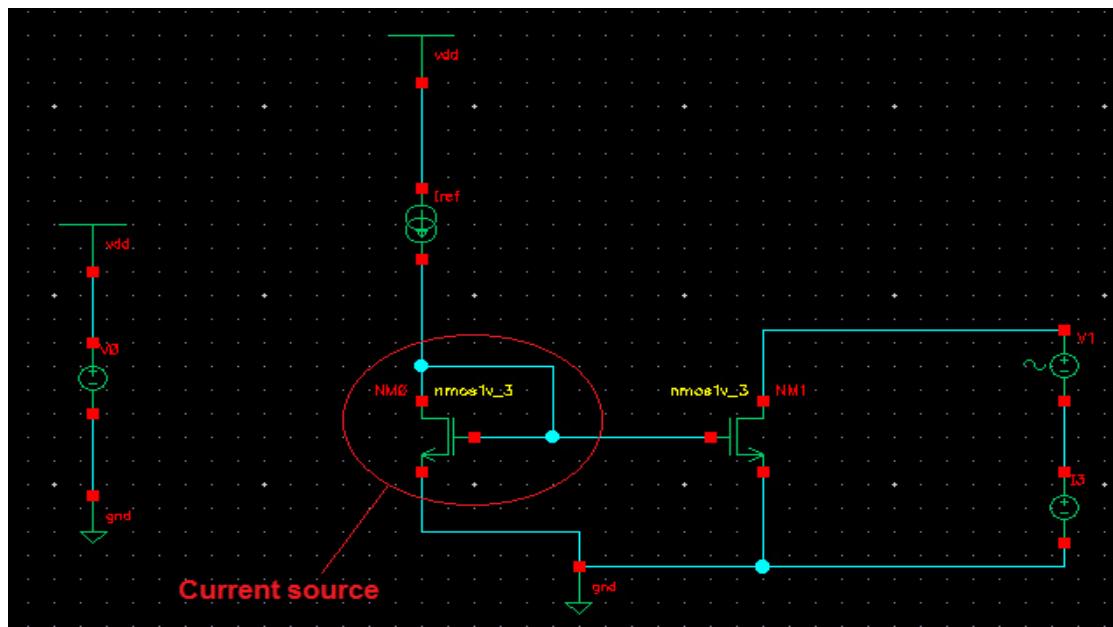


Figure . NMOS and PMOS implementations of a simple current mirror

- Earlier the constant current source was use resistor than it replace by a mosfet with a constant gate voltage than it replace by a diode connected mosfet/device.
- It is 2 terminal devices.
- The current source will force current to the diode connected MOSFET NM0 which intern develop voltage **VGS** across gate of NM0. This **VGS** is given to the gate of Mosfet NM1. If the W/L ratios are same then it will the same current as current source. So, here we are able to mirror the current from Mosfet NM0 to mosfet NM1.



Cascode MOS current mirror

The cascode current mirror is derived from the simple current mirror by cascading both branches of the circuit. The schematic of the resulting structure is given in Figure.

NMOS and PMOS implementations of a cascode current mirror The bulk terminal of the transistors is connected either to the lowest potential, or to VDD for PMOS transistors. In Figure the substrate connections of the transistors M3 and M4 facilitate the identification of VBS voltages when calculating the input and the output resistances.

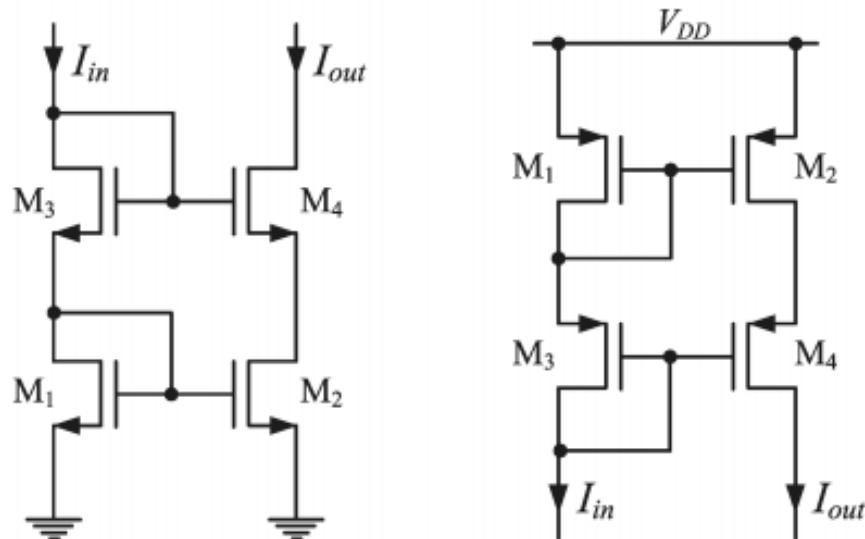
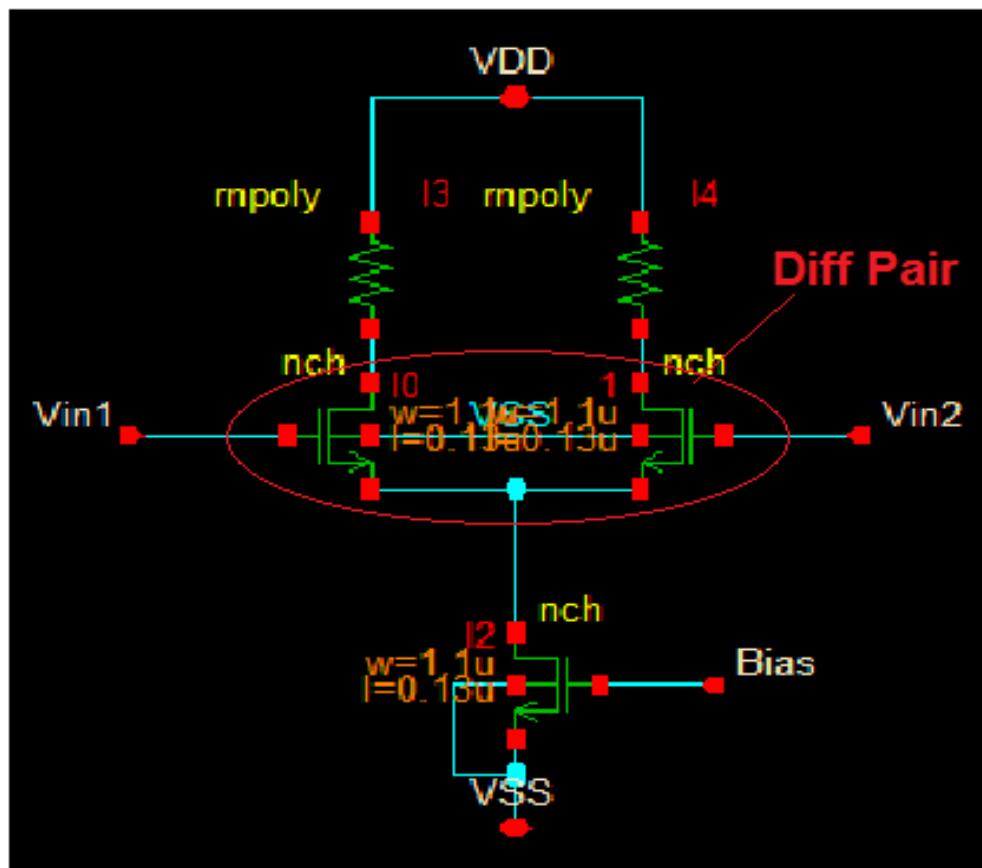


Figure . NMOS and PMOS implementations of a cascode current mirror

Video link: <https://youtu.be/GZvLm-mMJlY>

- **Differential Pair (input Pair) :**

- a. Differential amplifiers apply gain not to one input signal but to the *difference* between two input signals. This means that a differential amplifier naturally eliminates noise or interference that is present in both input signals
- b. Differential amplification also suppresses common-mode signals-in other words, a DC offset that is present in both input signals will be removed, and the gain will be applied only to the signal of interest (assuming, of course, that the signal of interest is not present in both inputs). This is particularly advantageous in the context of IC design because it eliminates the need for bulky DC-blocking capacitors.
- c. The subtraction that occurs in a differential pair makes it easy to incorporate the circuit into a negative-feedback amplifier



Video link : <https://youtu.be/dvuG414ZsTQ>

❖ How does mismatch affect the performance of the circuit ?

$$I_1 = \frac{1}{2} \mu n C_{ox} (W/L)_1 \cdot (V_{GS} - V_{t1})^2$$

$$I_2 = \frac{1}{2} \mu n C_{ox} (W/L)_2 \cdot (V_{GS} - V_{t2})^2$$

Where

I = Drain current

μn = Mobility carriers of device either minority carrier or majority carrier

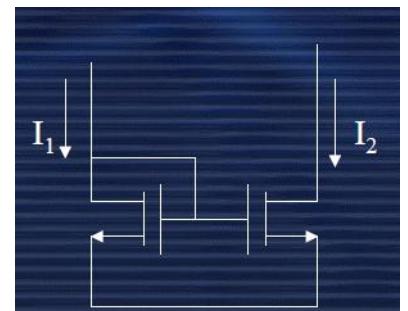
C_{ox} = Thickness of oxide

W = Width of device (OD)

L = Length of device (OD)

V_{GS} = gate source voltage

V_T = Threshold voltage



Defining average and mismatch quantities, we have

$$I = (I_1 + I_2)/2 ; \Delta I = I_1 - I_2; W/L = [(W/L)_1 + (W/L)_2]/2$$

$$V_t = (V_{t1} + V_{t2})/2 ; \Delta V_t = V_{t1} - V_{t2}$$

Substituting these expressions and neglecting higher order terms we obtain :

$$\Delta I/I = [\Delta(W/L) / (W/L)] - [\Delta V_t / (V_{GS} - V_t)^2]$$

Thus from the above equation we can see that the mismatch in the current depend upon

1. Mismatch in the (W/L) values of the transistors.
2. Mismatch in the threshold values of the transistors which increases as the overdrive voltage ($V_{GS} - V_t$) is reduced.

❖ Input offset voltage of a differential pair:

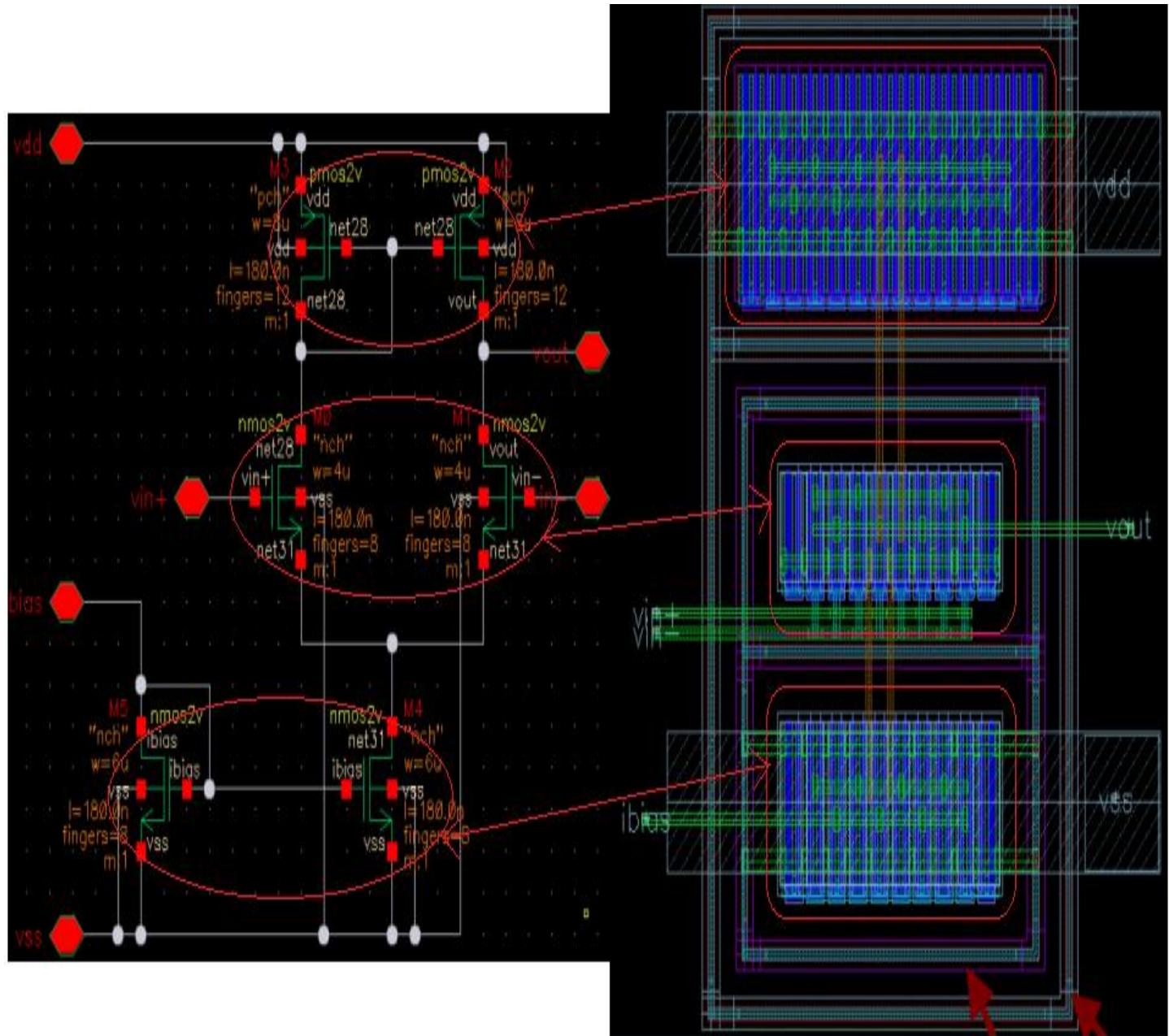
$$V_{OS} = \Delta V_t + \frac{(V_{GS} - V_t)}{2} \left[-\frac{\Delta R_L}{R_L} - \frac{\Delta(W/L)}{W/L} \right]$$

Thus we see that the offset voltage depends upon two parameters :

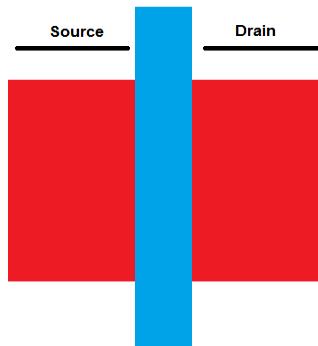
1. The first component is the threshold voltage mismatch of the transistors . This depends upon the layout and it can be reduced by careful layout.
2. The second component of the offset scales with the overdrive voltage and is related to mismatch in the load elements and mismatch in the W/L values.

- **Industrial Quality :**

- Includes multiple guardrings.
- Full common centroid.
- Dummy structures.



LOD Effect (Length of oxide diffusion) :



In CMOS source area & drain area of a single finger device is equal.

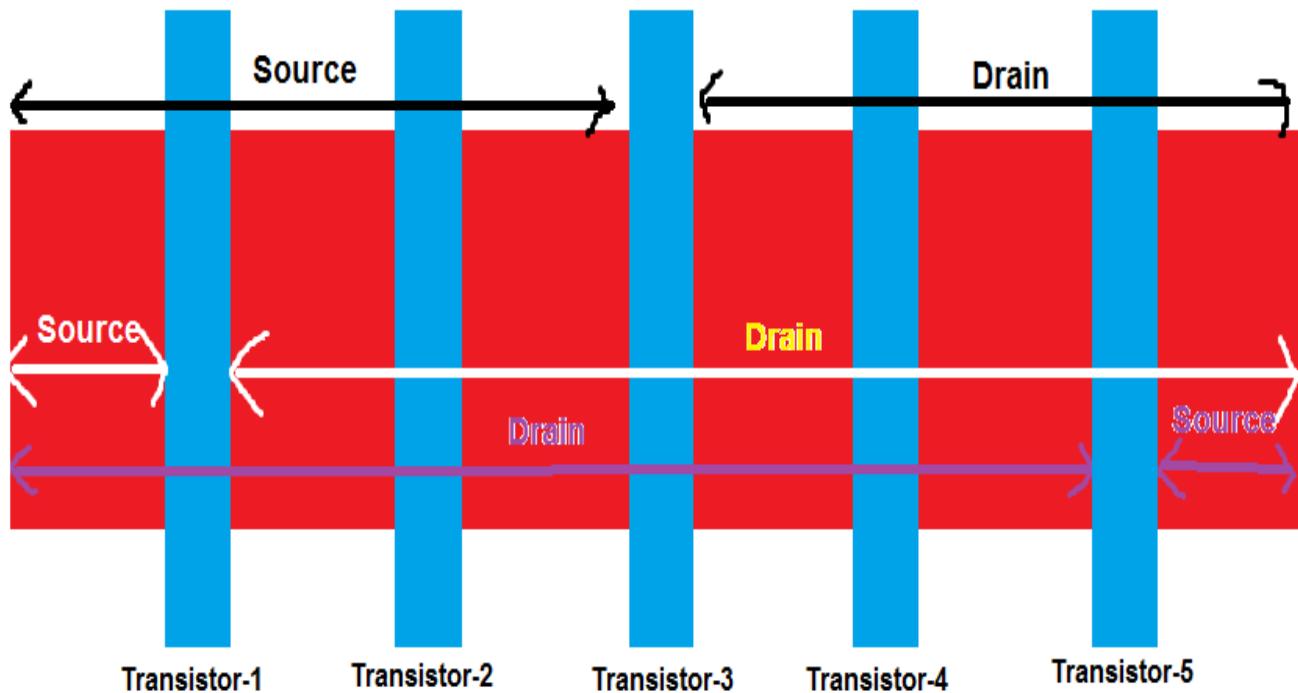
In case of multiple finger the area of source and drain are not equal for corner transistor.

Example – For below device

Transistor-3 : The source area is equal to drain area.

Transistor-1 : The source area is smaller than drain area.

Transistor-5 : The drain area is larger than source area.



Due to above condition, unintentionally we increase the size of device, cause of this the electron mobility & holes mobility effect threshold voltage V_T of the transistor.

The transistor property may vary upto $\pm 10\%$ due to affect in V_T .

How V_T change :

$$I_D = \frac{1}{2} \mu n \cdot C_{ox} \cdot (W/L) \cdot (V_{gs} - V_T)^2$$

Where

I_D = Drain current

μn = Mobility carriers of device either minority carrier or majority carrier

C_{ox} = Thickness of oxide

W = Width of device (OD)

L = Length of device (OD)

V_{gs} = gate source voltage

V_T = Threshold voltage

Due to increase in mobility charge carrier the threshold voltage is varying $\pm 10-20\%$

Example:

Assume in PMOS we fixed $V_T = 0.7v$

Due to increase in mobility extra majority charge carrier present in PMOS due to LOD and device becomes turn on in 0.3v or 0.4v.

In PMOS majority charge carrier is holes

Similarly

Assume in NMOS we fixed $V_T = 0.7v$

Due to increase in mobility extramajority charge carrier present in NMOS due to LOD and device becomes turn on in 0.9v or 1.0v.

In NMOS majority charge carrier is holes.

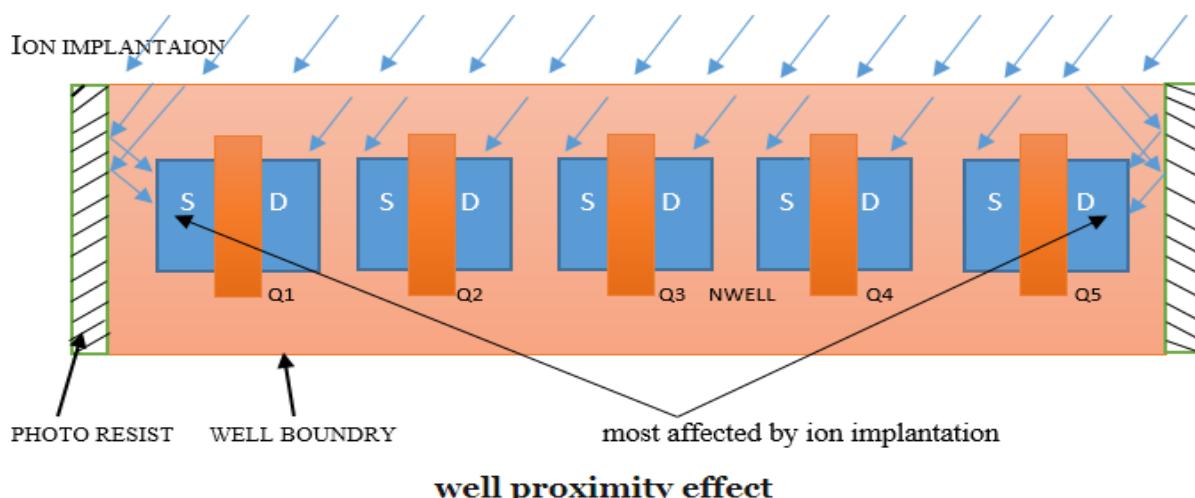
- **Sources of layout proximity effect**

- Well proximity effect
- Unintentional stressors f: Shallow trench isolation (LOD effect)

Well proximity effect (WPE) :

The transistors that are close to the well edge have different performance than ideally placed transistors, because of this effect the transistor speed can vary by +- 10%. the V_t of transistor got changed.

The transistor placed to the well boundary so it will get many problems during ion implantation. Implanted ion is coming to the well boundary and reflected/scatter from the well boundary to transistors Q1 & Q5 boundary and ions are deposited on the Q1 Q5 boundary. Ion particles are scattered/reflected due to photoresist on both side of Nwell. These ions are deposited only those transistors who are near to the well boundary, so any one of the terminals of transistors gets affected by ion implantation and the rest of the transistor will get uniform ions.

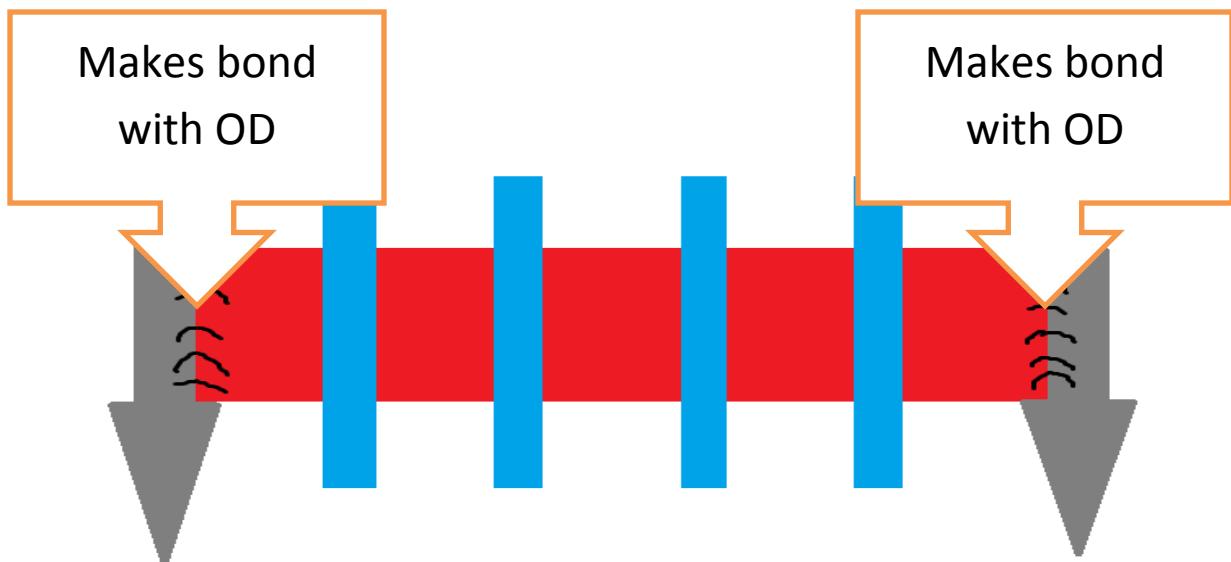


Solution:

- Use dummy surrounded to the active device
- The Well boundary should be 1.5 or 2 um away from the active device

Shallow trench isolation(STI) :

Silicon is dielectric material so it makes bond with active OD. So STI covalent bond share with active OD covalent bond & unintentionally the mobility carrier is changed, cause of this V_T also changed and the performance of the device affected.



Video link: <https://youtu.be/cd5qdHtxogs>

7.GENERAL LAYOUT GUIDELINES

Layout Steps :

- a. **Floor planning** - Division of the entire die area among subcomponents to facilitate interconnection and effectively utilize the area.
- b. **Placement** - Placing the modules in the layout.
- c. **Routing** - Connecting the modules with different metal layers.

Some important guidelines which should follow in layout:

- ✓ Add dummy device surrounding any critical matched devices
- ✓ Uniform Poly and OD spacing
- ✓ Align/overlap dummy poly can guarantee uniform od spacing
- ✓ Use mosaic layout
- ✓ Big transistor array or poly resistor array will have high OD and/ Poly density, can affect matching. Refer to design manual for density rule.
- ✓ Avoid OD/Poly for signal routing
- ✓ Adequate Well/Bulk/Substrate ties
- ✓ All OD/Poly/NW in rectangular shape
- ✓ Use even MOS fingers for device merging
- ✓ Evaluate LOD for MOS unit in an array Especially for PMOS unit when used in headroom sensitive design or clock driver
- ✓ Avoid shared OD for MOS with different length
- ✓ If possible, connect gate from both end for lower gate resistance.

Guidelines for analog layout :

- ✓ Use of more number vias
- ✓ Fingering and proper orientation
- ✓ Device matching
- ✓ Symmetrical and common centroid layout design
- ✓ Use of Guard ring and substrate trapping

Standard cell design methodology :

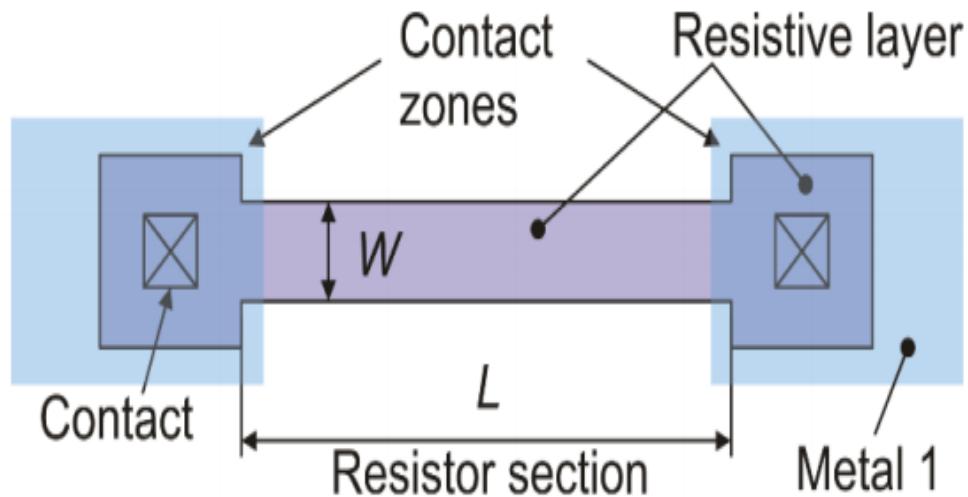
- ✓ VDD and GND should abut (standard height)
- ✓ Adjacent gates should satisfy design rules
- ✓ NMOS at bottom and PMOS at top
- ✓ All gates include well and substrate contacts

❖ Things to remember :

- ✓ Keep sufficient spacing between power blocks and sensitive blocks.
- ✓ Two high frequency carrying pins should not be side by side.
- ✓ Use ground pin to avoid magnetic coupling between two pins.

8.DEVICE AVAILABLE IN LAYOUT

Resistor: The generic layout of an integrated resistor is shown in figure. We have a resistive layer, which is properly shaped to obtain the required resistance. Contacts to an interconnect layer (metal 1 in the example) are placed at both end of the resistive shape, creating the two resistor terminals. The rectangular shape that is included between the contacts is the resistor section, or resistor body.



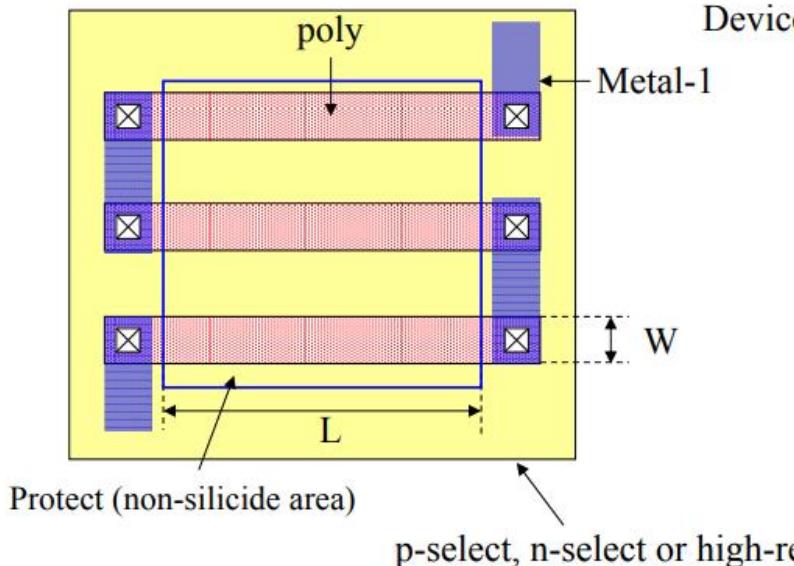
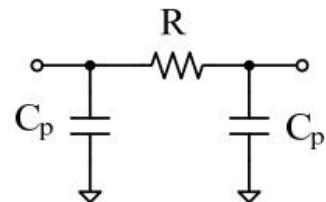
- There are two types of resistor in layout
 1. Poly resistor
 2. Metal resistor

- **Poly Resistor:**

Layout sample of a poly resistor

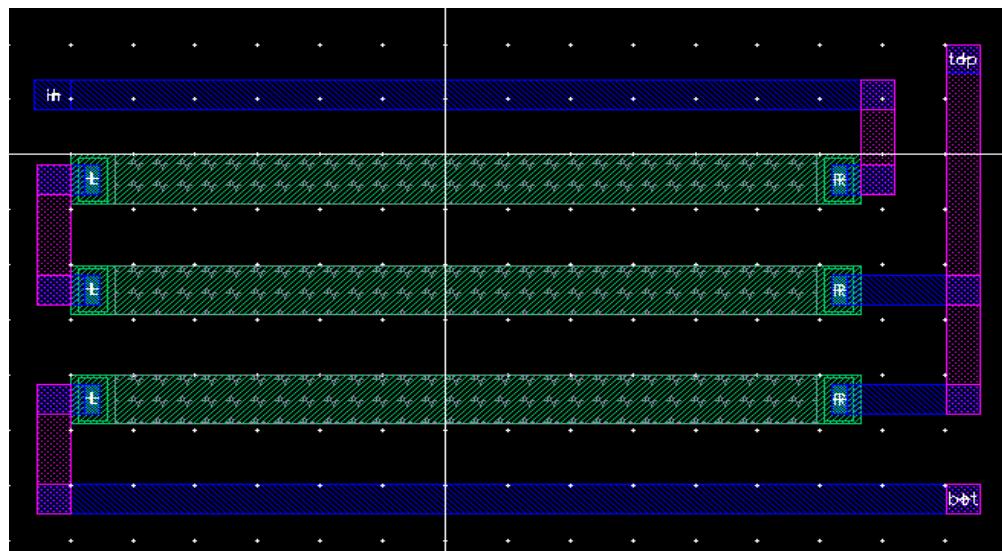
$$R = R_s \frac{L}{W} \quad (\text{recommended } L/W > 5)$$

R_s : Sheet Resistance

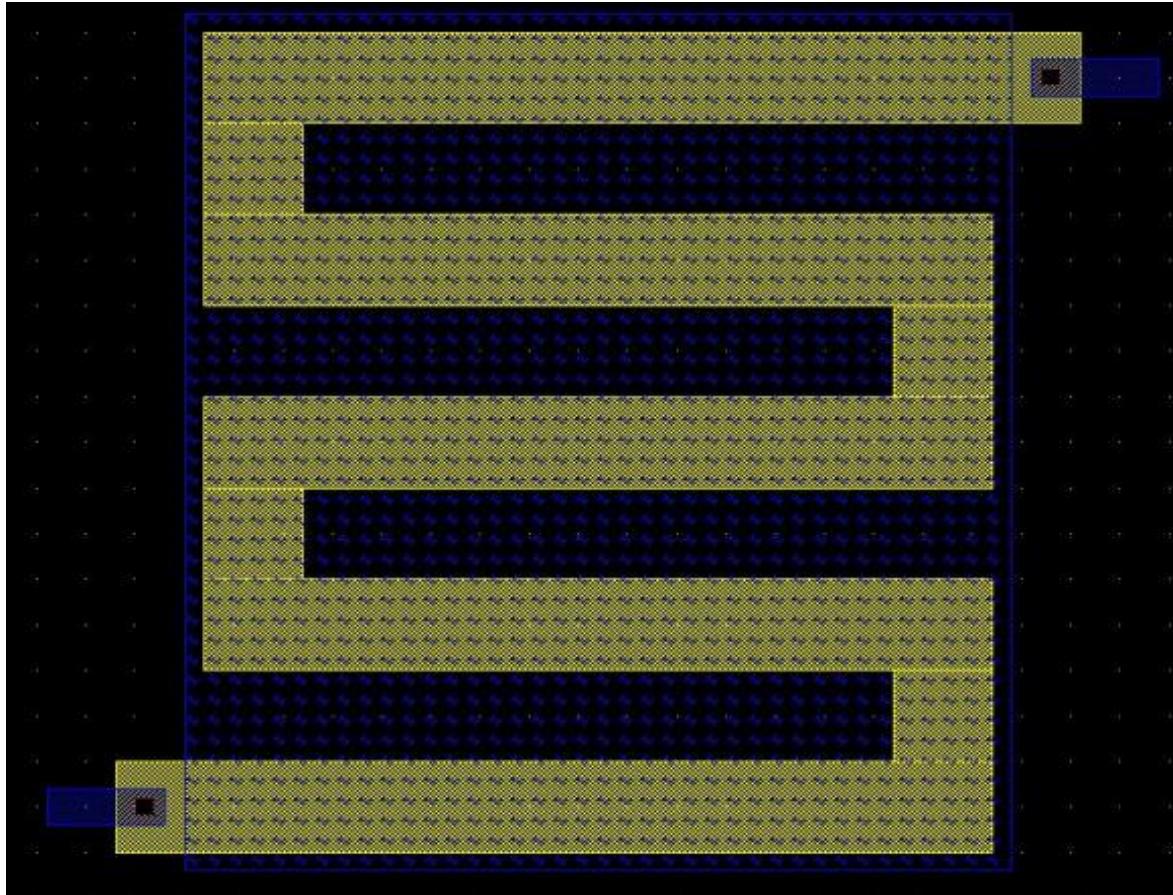


Device model with the parasitic

33



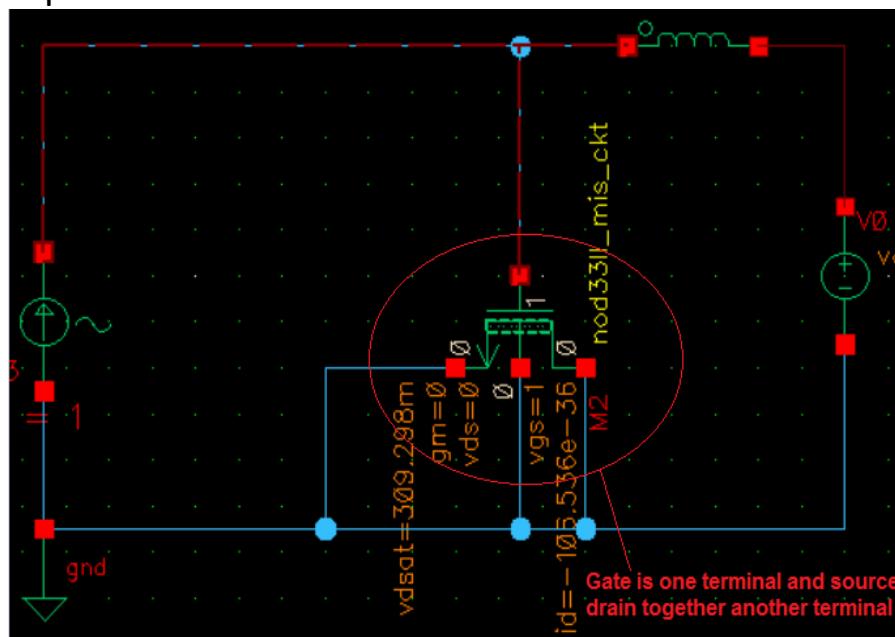
- **Metal Resistor:**



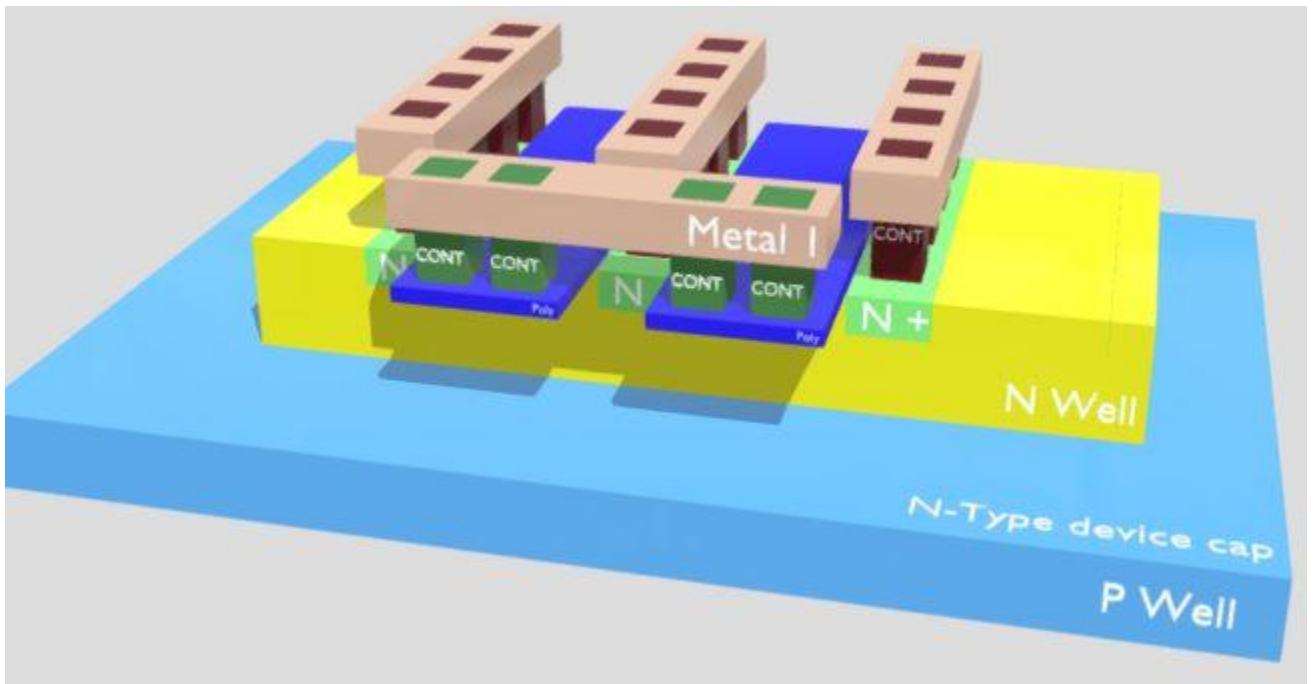
Capacitor:

- **MOS capacitors :**

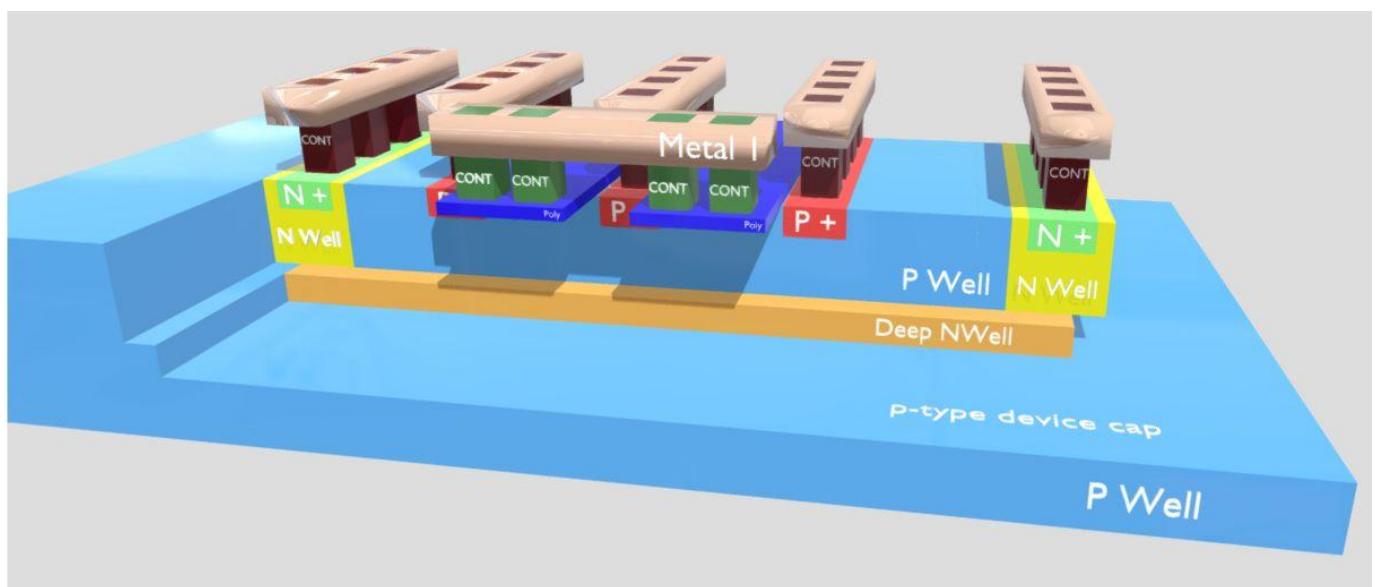
1. Transistors can be used as a cap using the thin oxide layer from the gate as an insulator. The Gate acts as the top plate and the Drain and Source as the bottom plate
2. The drawback is that if the voltage at the gate varies, the depletion area changes at the gate varying the dielectric properties and modifying the capacitance. So, the MOS capacitor's value is very dependent on the applied DC voltage.
3. The capacitance per unit area is higher than the MIM caps, because the insulator of the gate (SiO_2) is much thinner than the insulator between the metal layers. But the variation with the voltage is a big disadvantage.
4. They are useful for local supply decoupling where the DC voltage is constant. Another drawback is the parasitic resistance of the bottom plate.

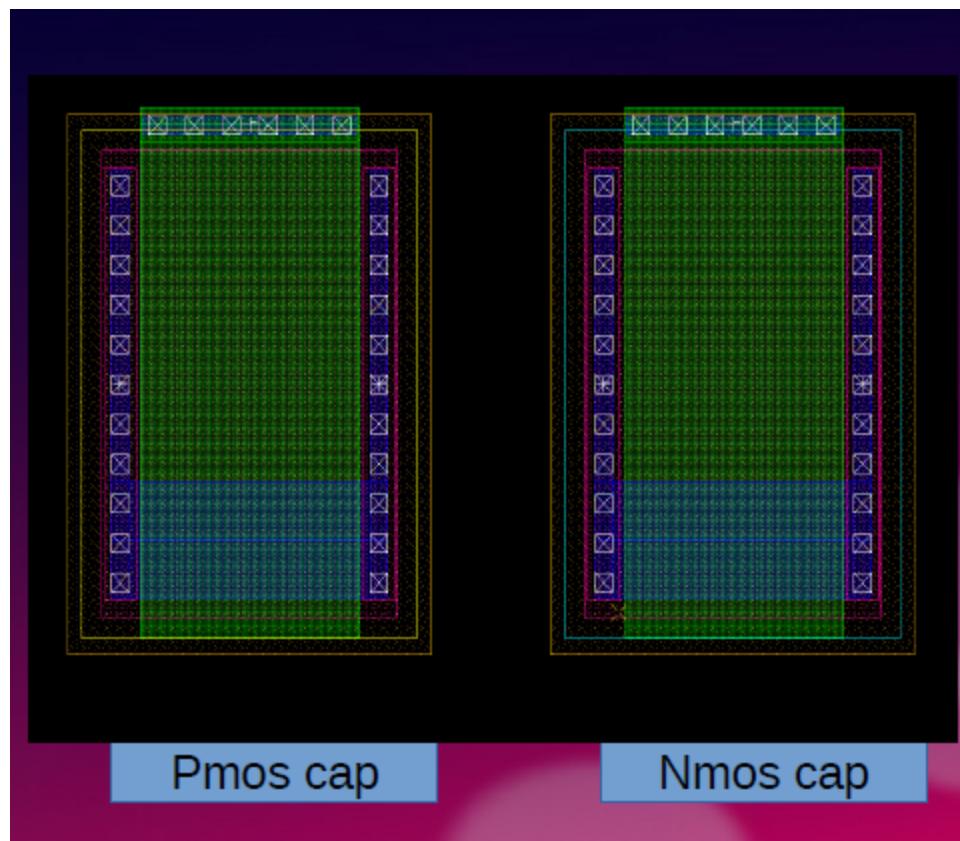


NMOSCAP:



PMOSCAP:



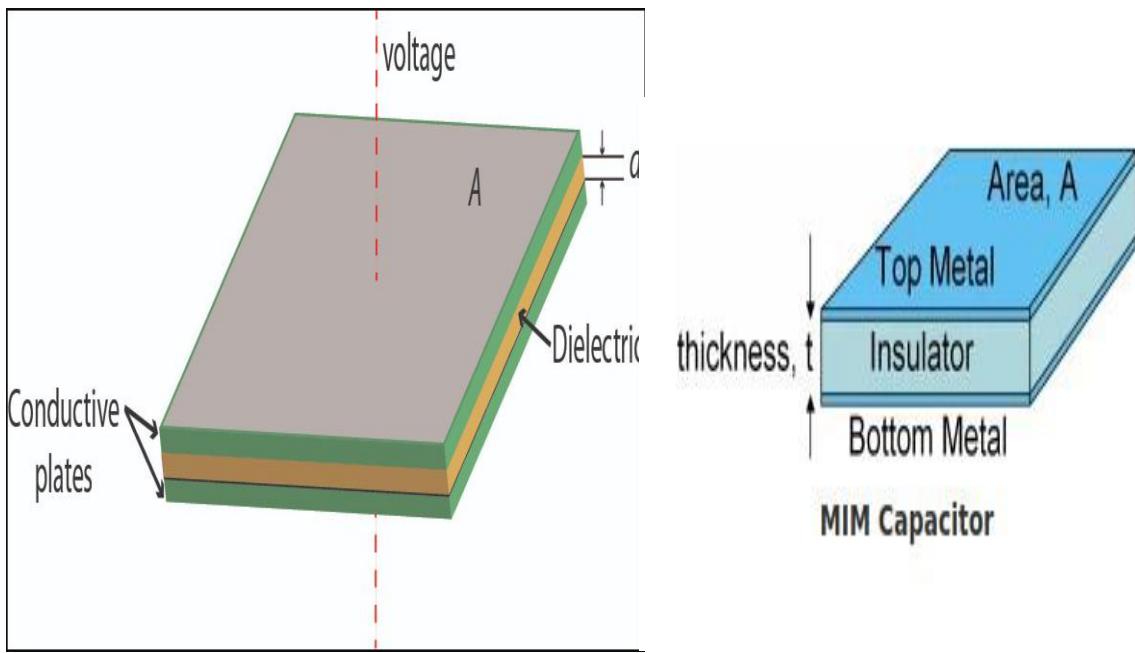


3d view link:

1. N-type device cap: <https://skfb.ly/oGTII>
2. P-type device cap: <https://skfb.ly/oGTIy>

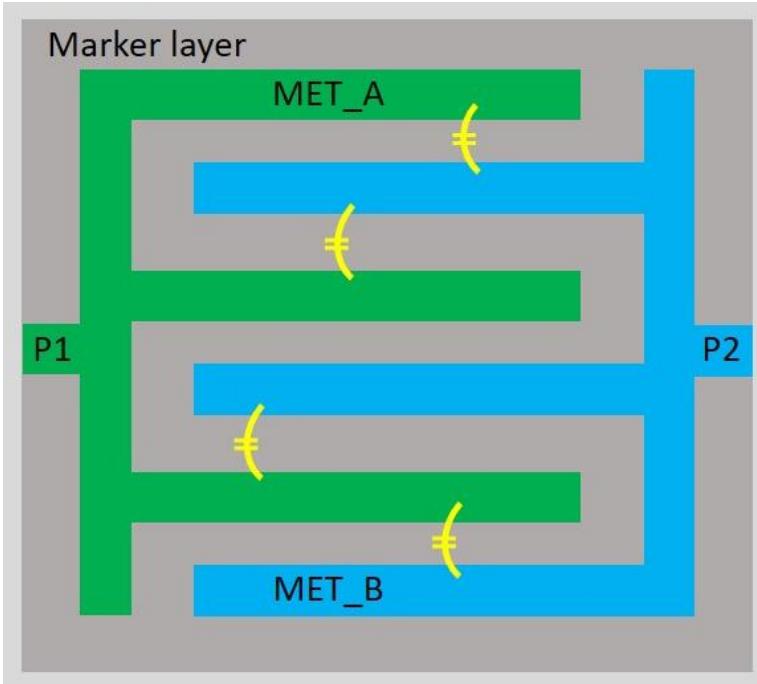
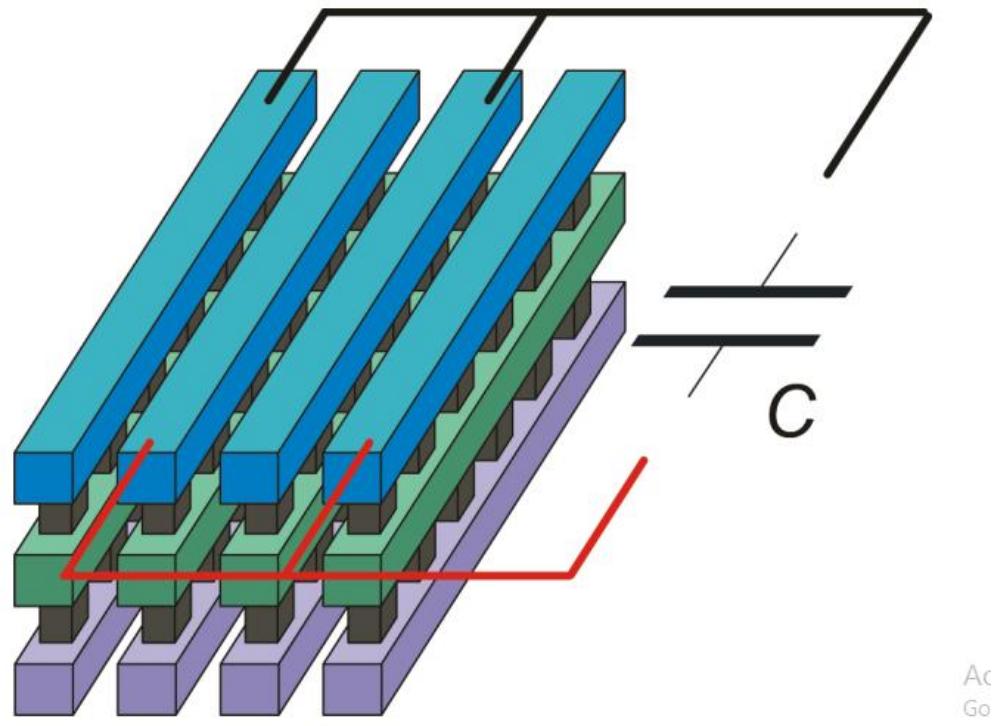
- **MIM capacitors :**

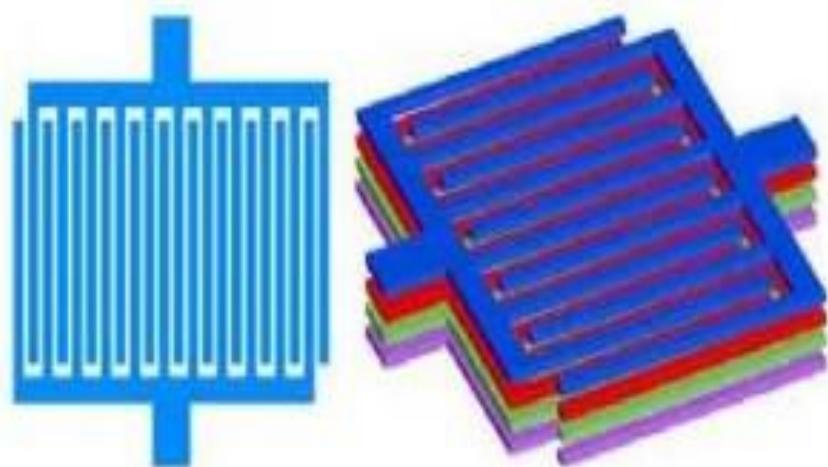
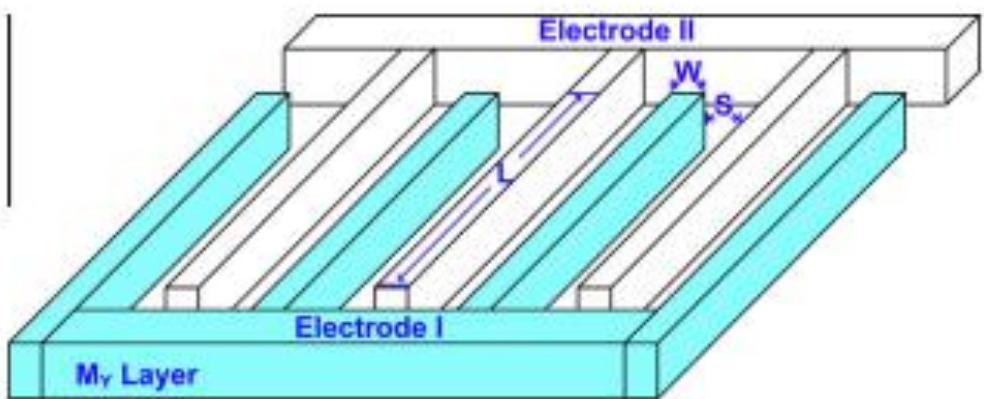
1. It is formed by two parallel metal layers and has a high k-dielectric between them. The bottom layer forms a small parasitic cap with the substrate.
2. This type is the most used, because of the high capacitance per unit area with the lowest parasitic. The drawback is that they require more process steps during the fabrication. Mainly, a new mask and step are added to deposit the insulator between the metal layers. Therefore, a fabrication process with MIM caps is more expensive than others using only MOM-caps.
3. They can be single, double, triple, etc. It depends on the number of layers of metal and insulators. The thickness of the insulator layer is in the range of 25 nm (depending on the fabrication technology).
4. The quality factor is very high because both plates are made out of metal and the voltage dependency is negligible



- **MOM capacitors :**

1. The advantage of the MOM capacitor is that it does not require additional process steps, because it can be obtained using only standard metal layers.
2. MOM capacitors are generally capable of withstanding much higher voltages.
3. The main drawback is that, for the same used area, the capacitances that can be obtained are smaller than allowed by MIM capacitors.
4. This problem can be overcome using multi-layer MOM capacitors, which involve several metal layers connected by means of vias.





MOM caps

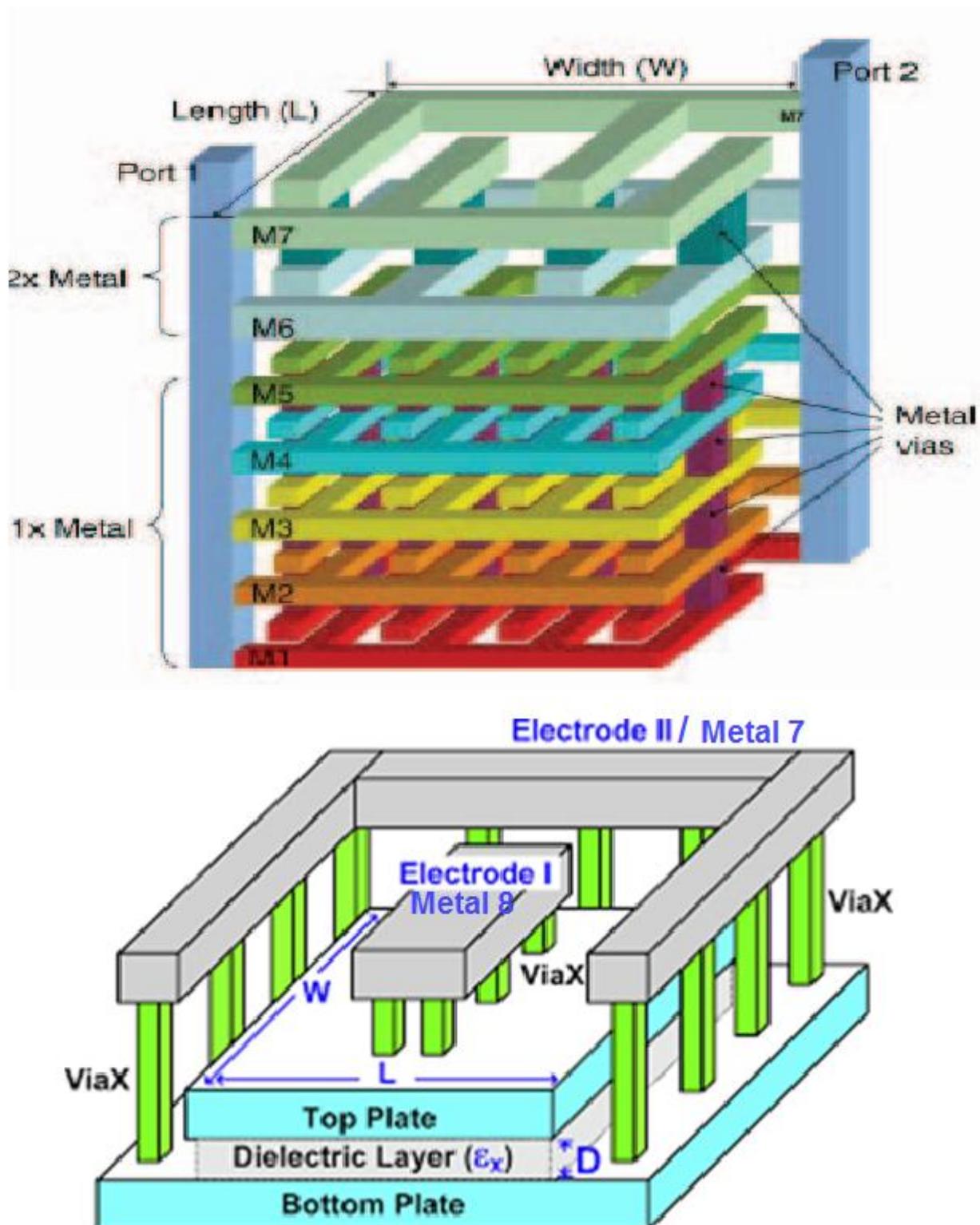


Fig: MIM Capacitor

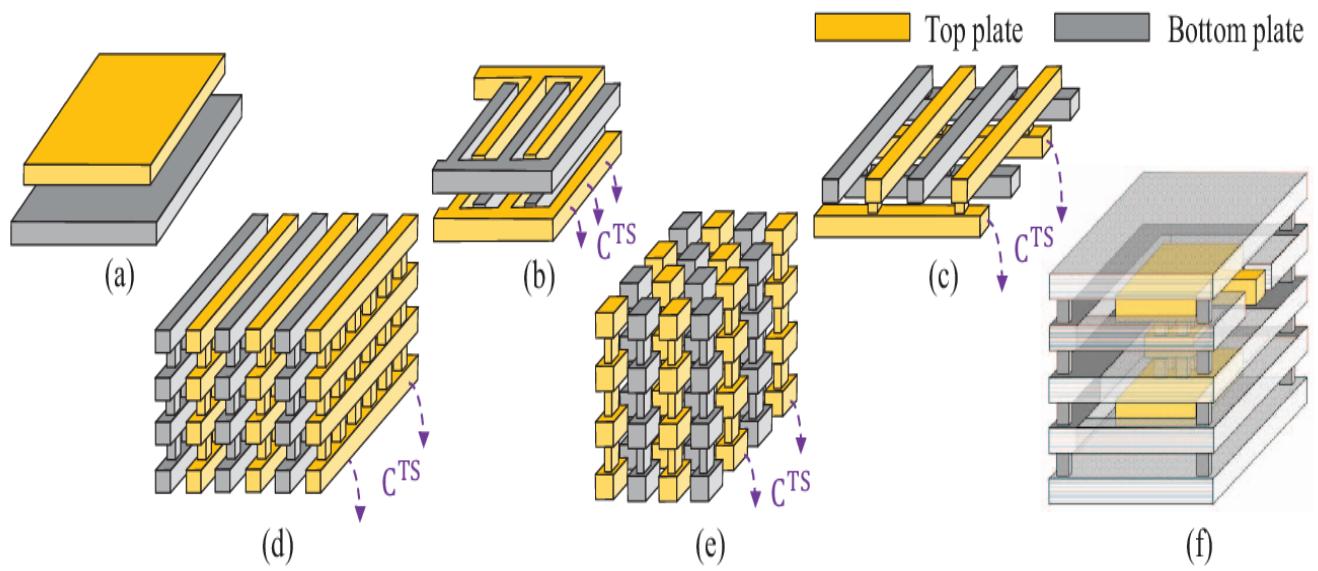
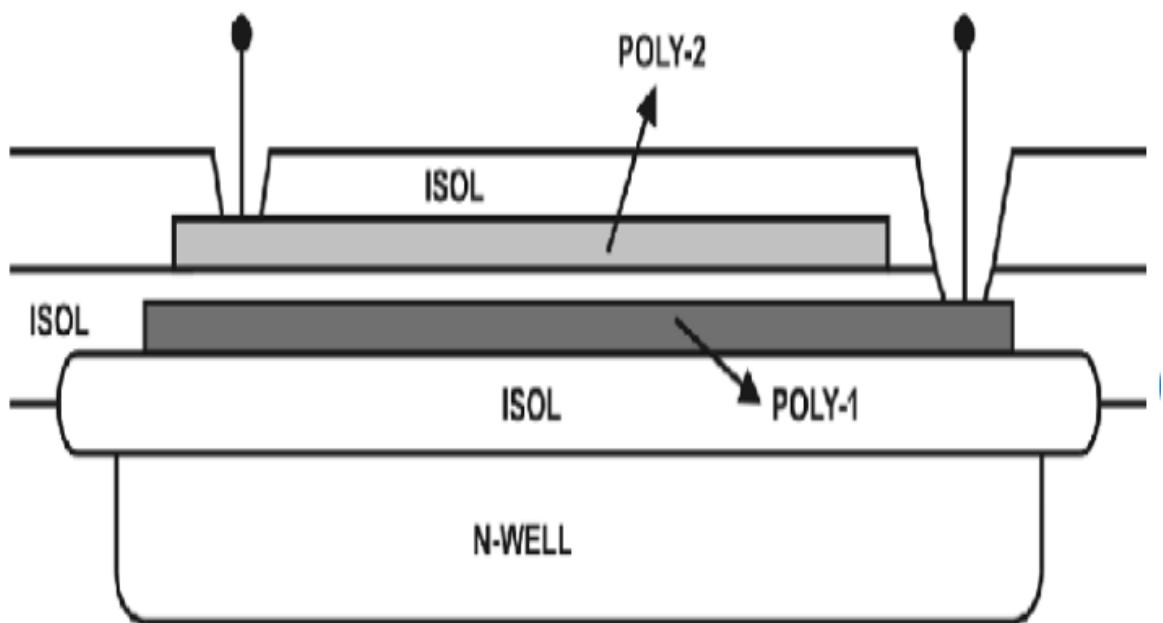
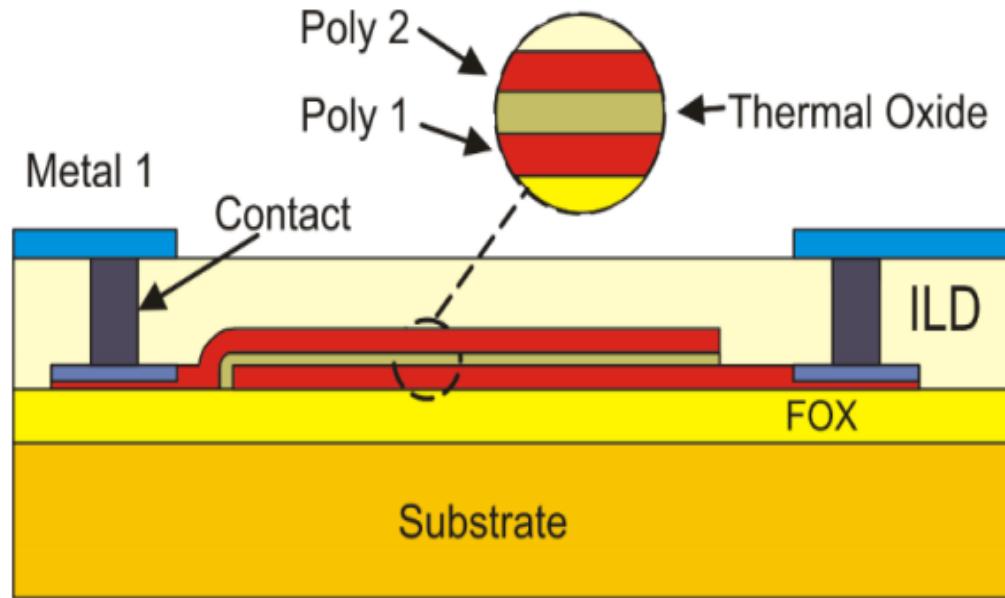


Fig:-

- (a) MIM capacitor. (b) MOM capacitor: interdigitated parallel wires. (c) MOM capacitor: woven.
- (d) MOM capacitor: parallel stacked wires. (e) MOM capacitor: vertical bars.
- (f) MOM capacitor: multi-layer sandwich.

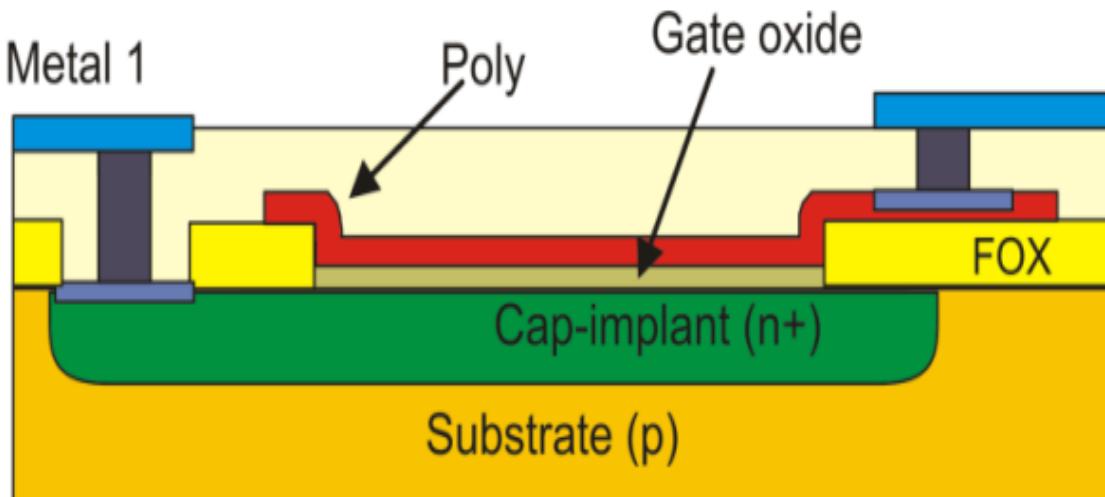
• Poly-Poly capacitors :





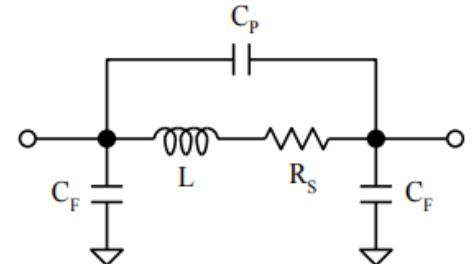
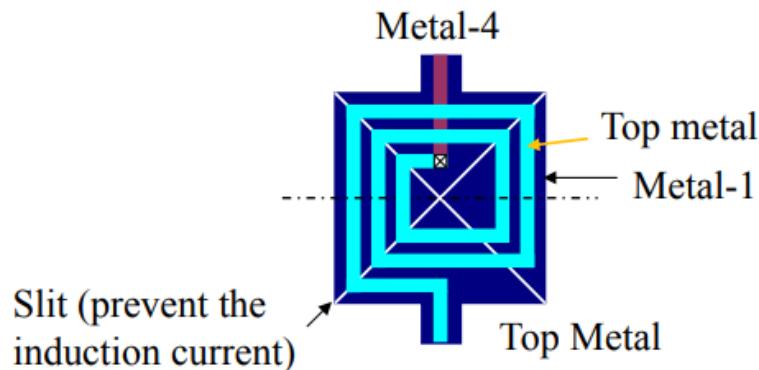
Cross-section of a poly-poly capacitor.

- **Poly-diffusion capacitors:**

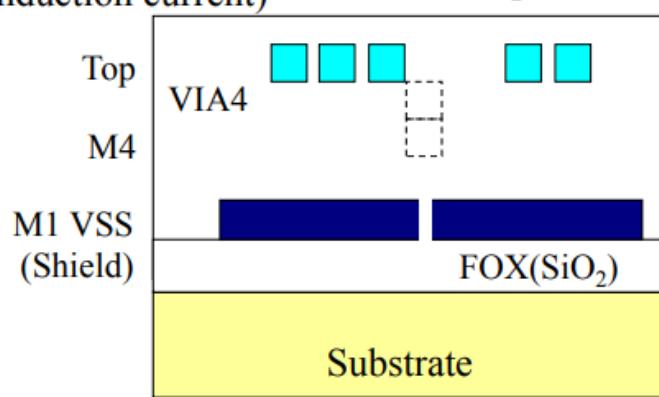


Poly-diffusion capacitor.

Inductor : Structure of spiral inductor



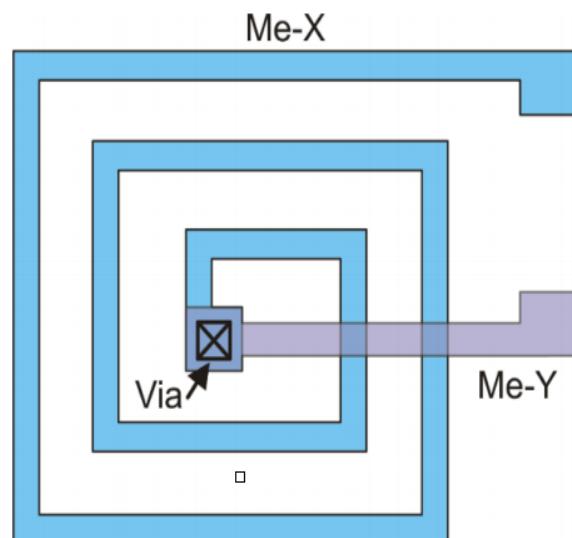
Device model with the parasitic



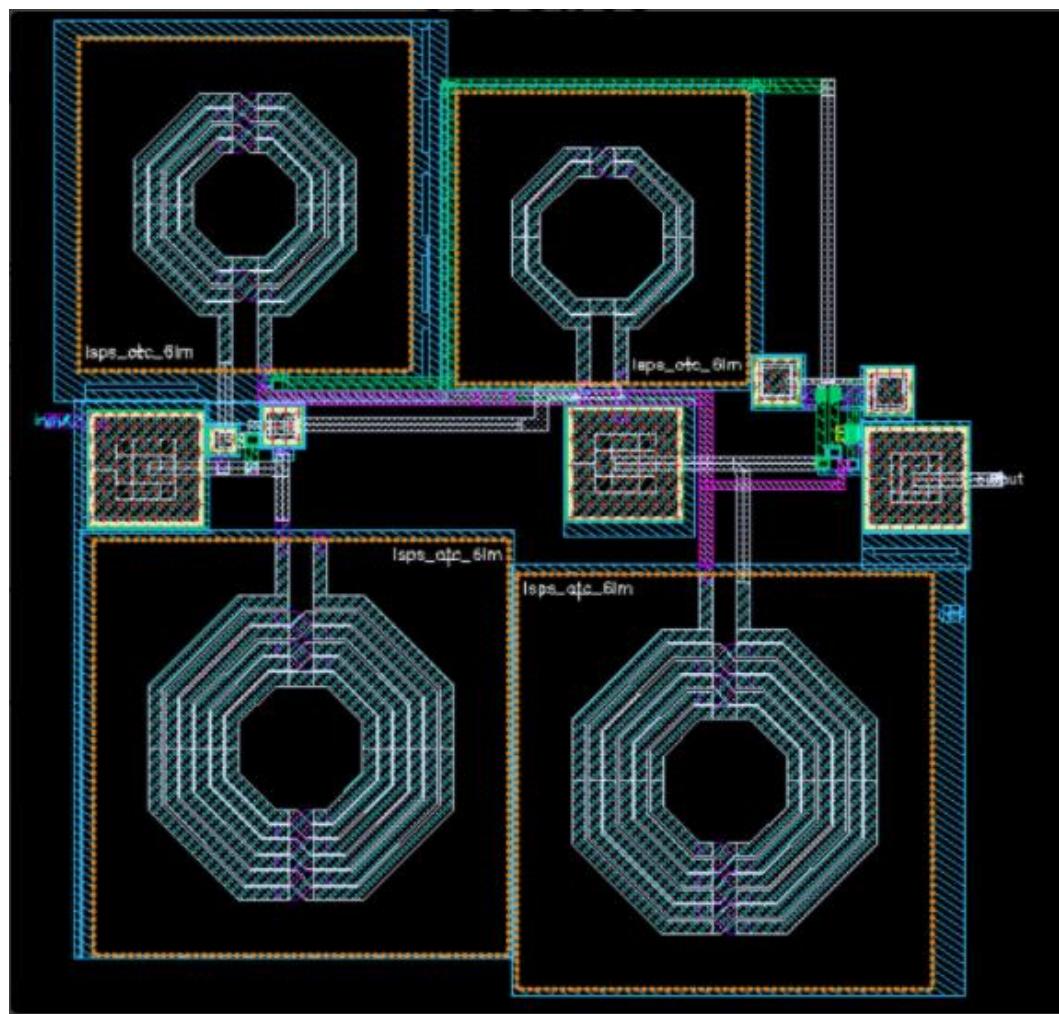
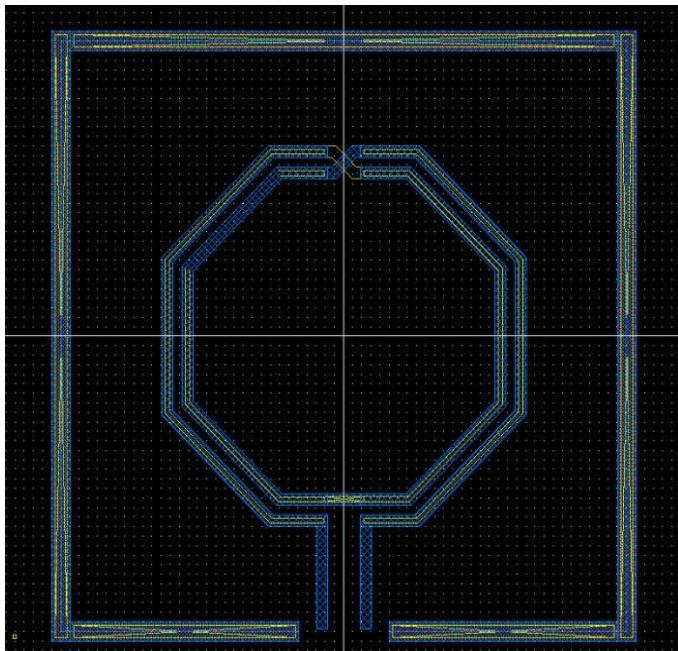
Cross section

Top metal or dedicated layer for inductor is used.

The inductor is dissipative in the chip area.

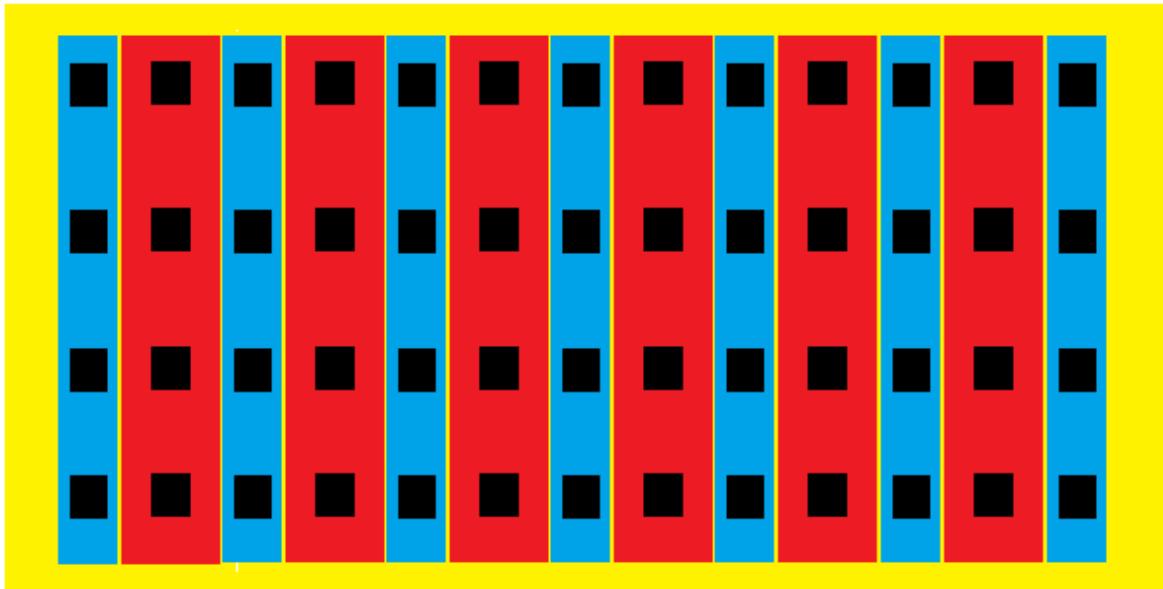


Integrated inductor

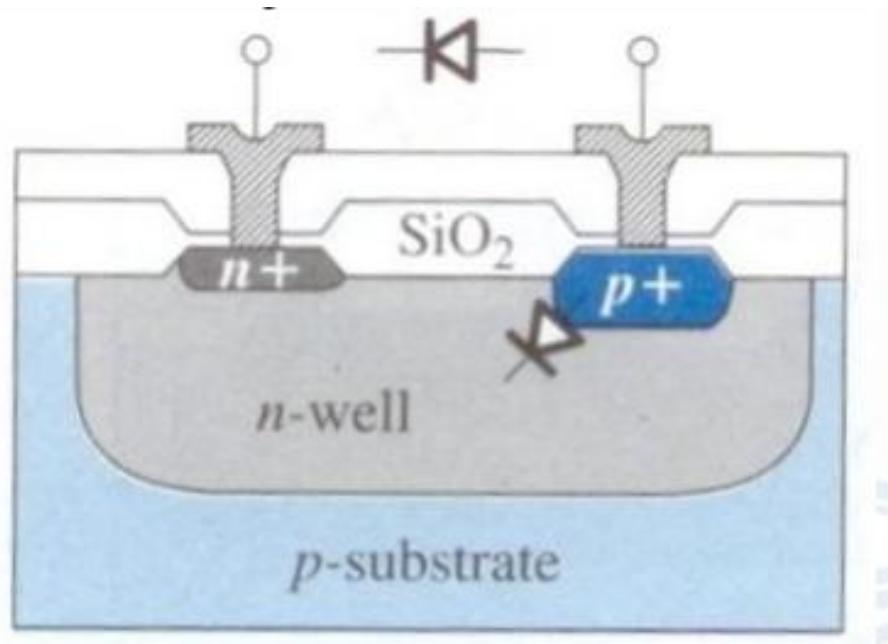


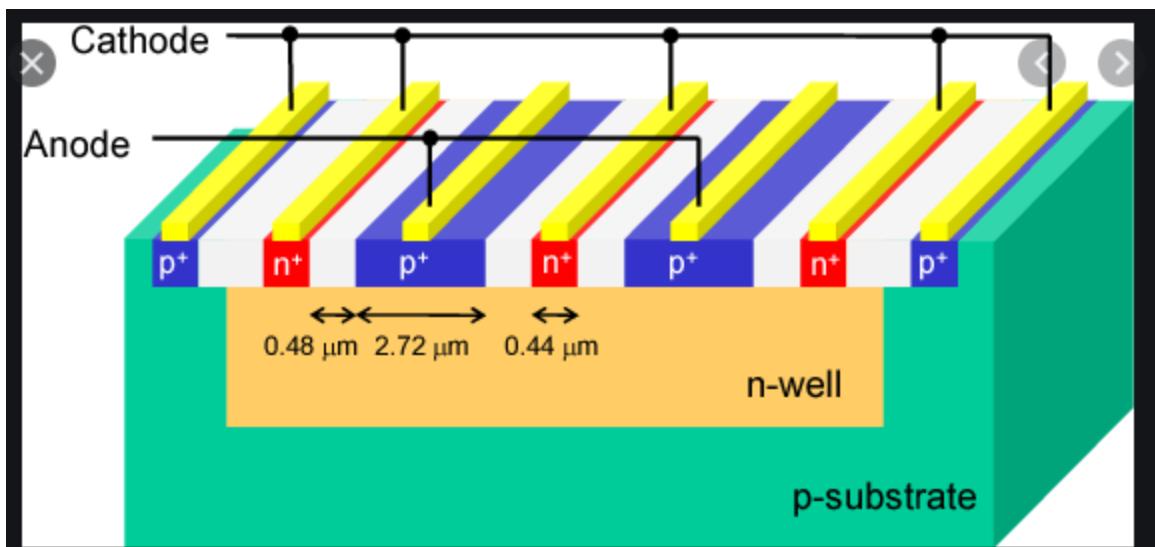
Diode :

Layout view:

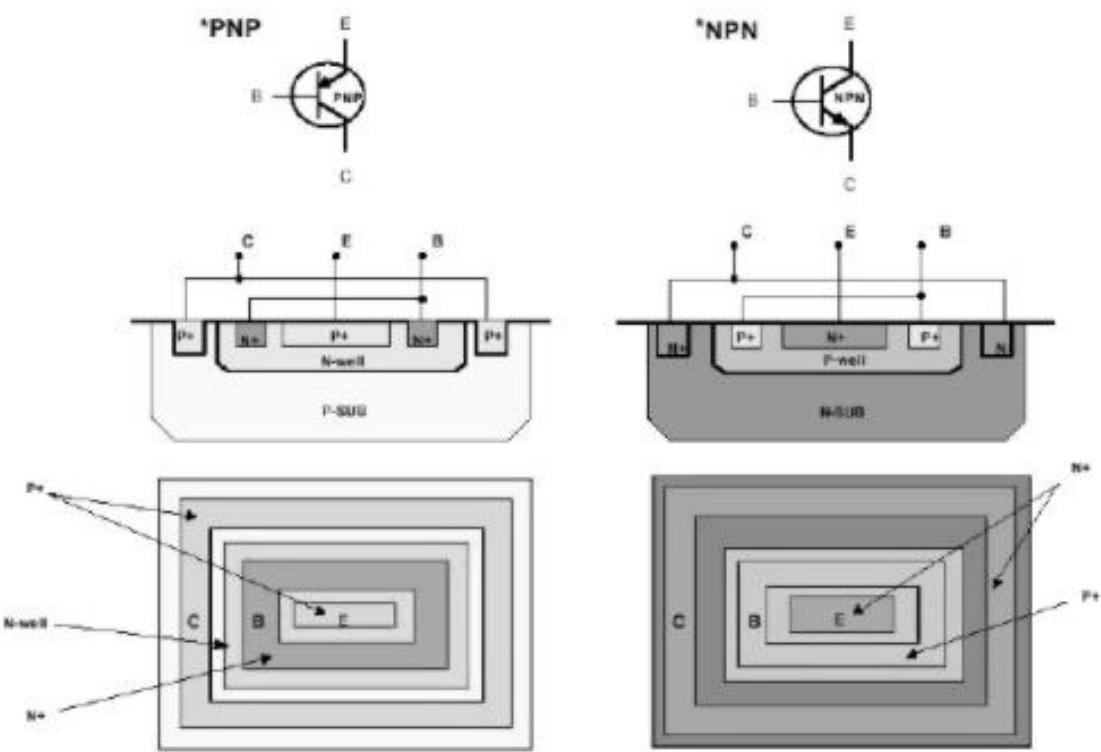
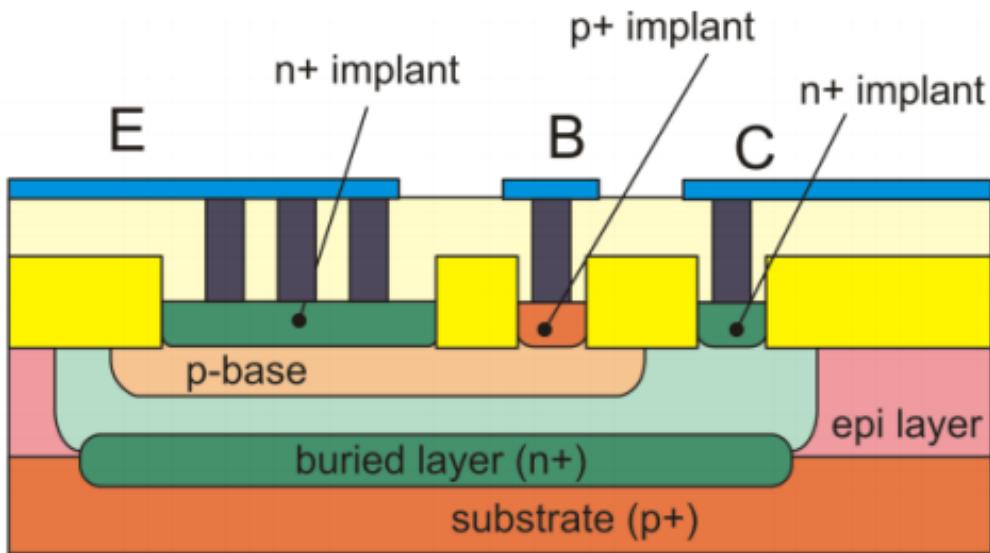


NWell =
P+ =
N+ =
Contact=





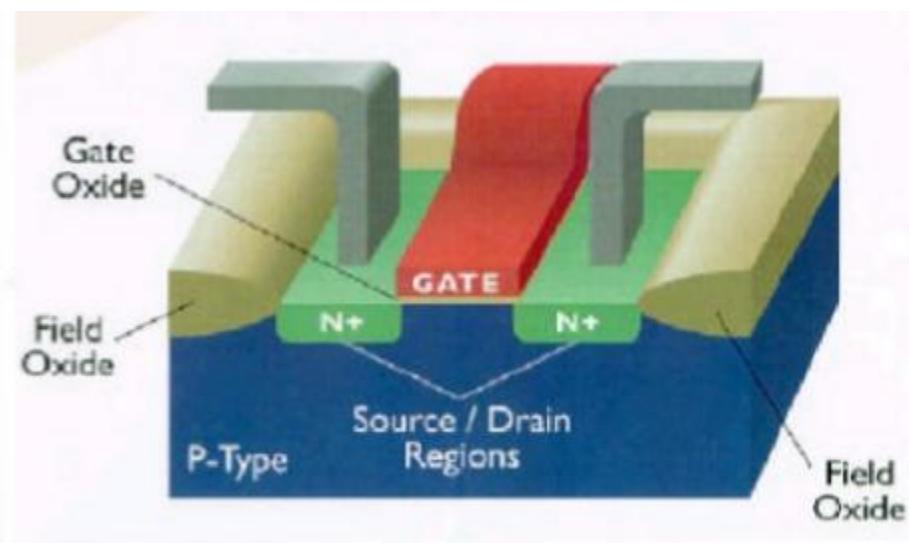
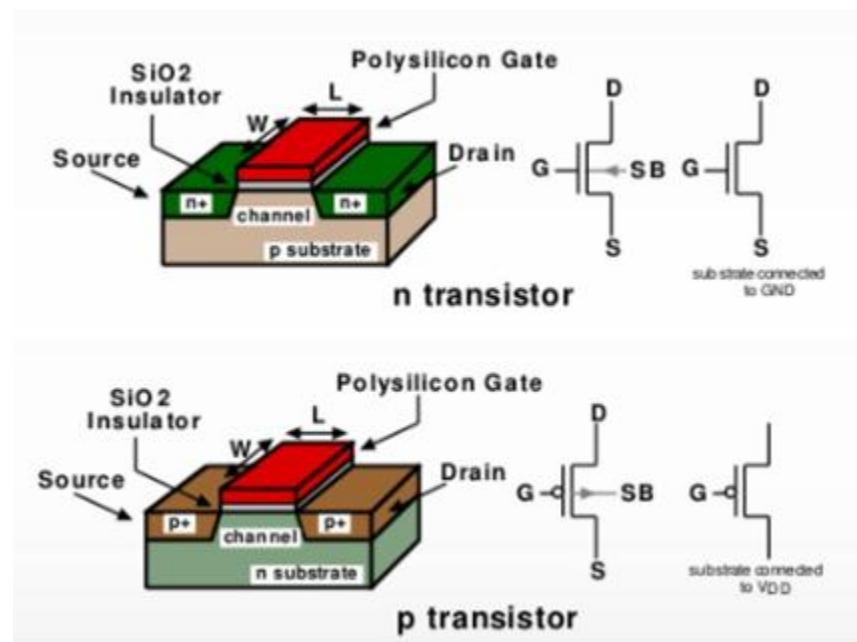
Transistor / BJT :



Video link of BJT : 1. <https://youtu.be/7ukDKVHnac4>

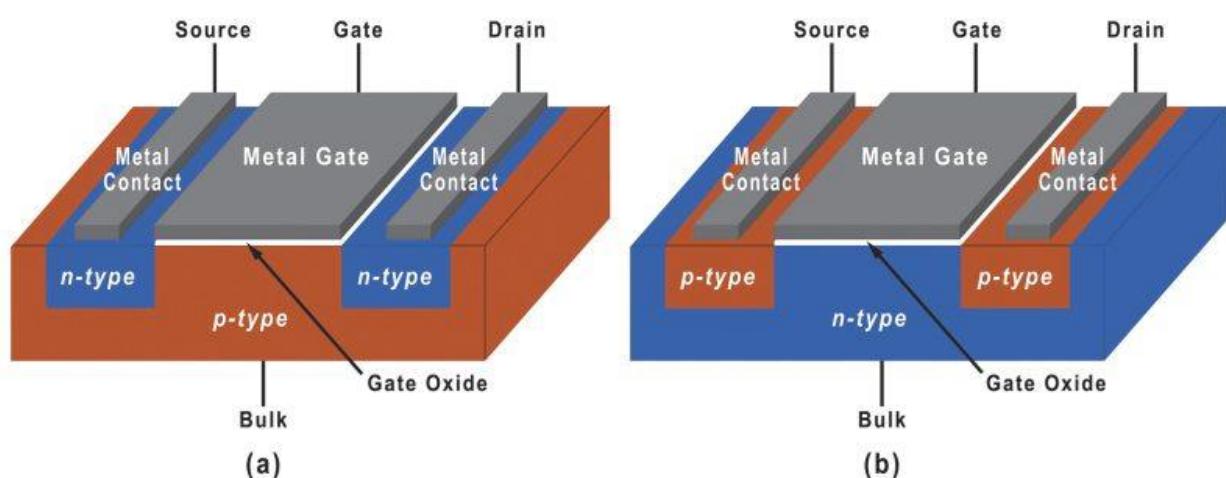
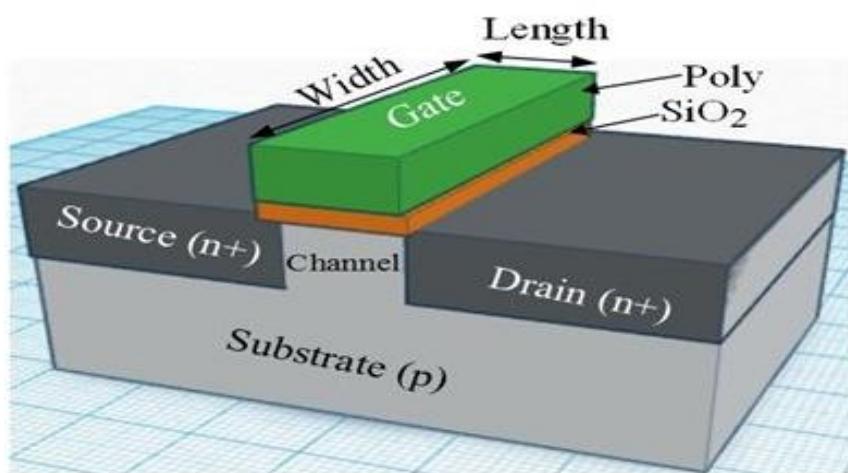
2. <https://youtu.be/zFLmuZa4tDE>

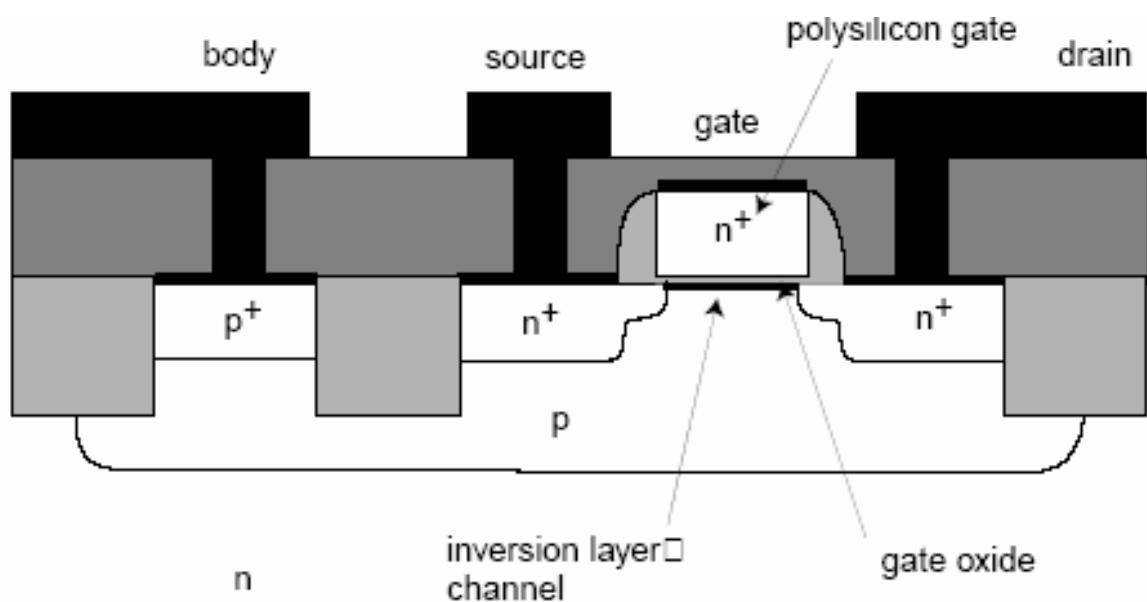
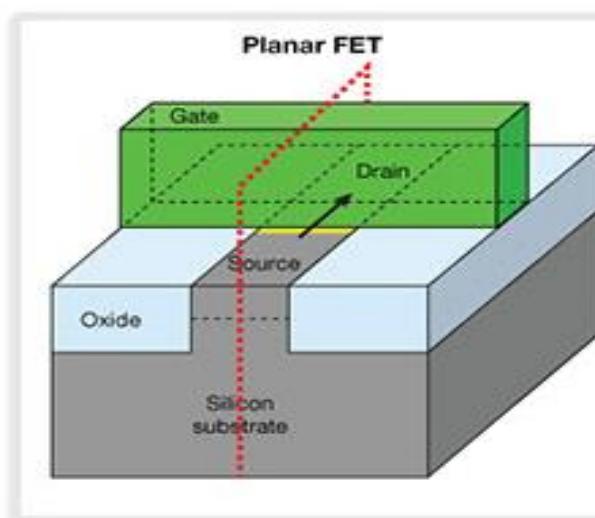
MOSFET :



Planer MOSFET:

The metal–oxide–semiconductor field-effect transistor (MOSFET, MOS-FET, or MOSFET), also known as the metal–oxide–silicon transistor (MOS transistor, or MOS), is a type of insulated-gate field-effect transistor that is fabricated by the controlled oxidation of a semiconductor, typically silicon. The voltage of the covered gate determines the electrical conductivity of the device; this ability to change conductivity with the amount of applied voltage can be used for amplifying or switching electronic signals.

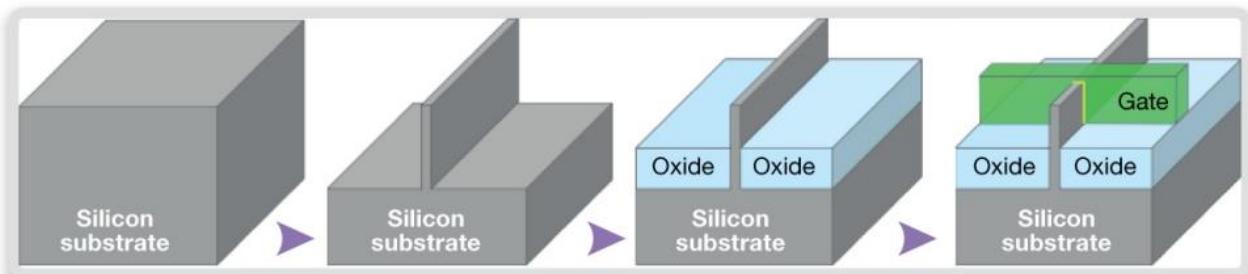




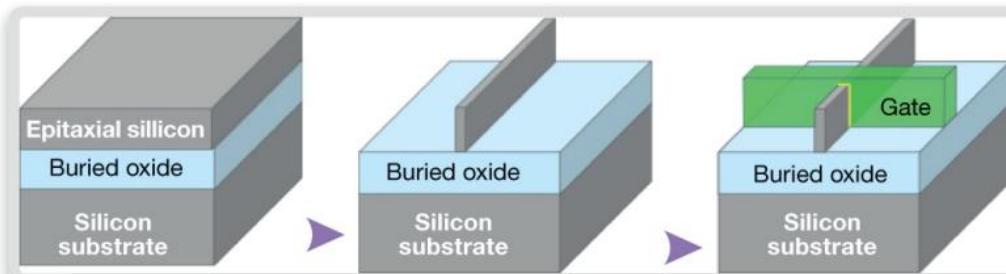
3D view link: <https://skfb.ly/oGOKI>

Finfet:

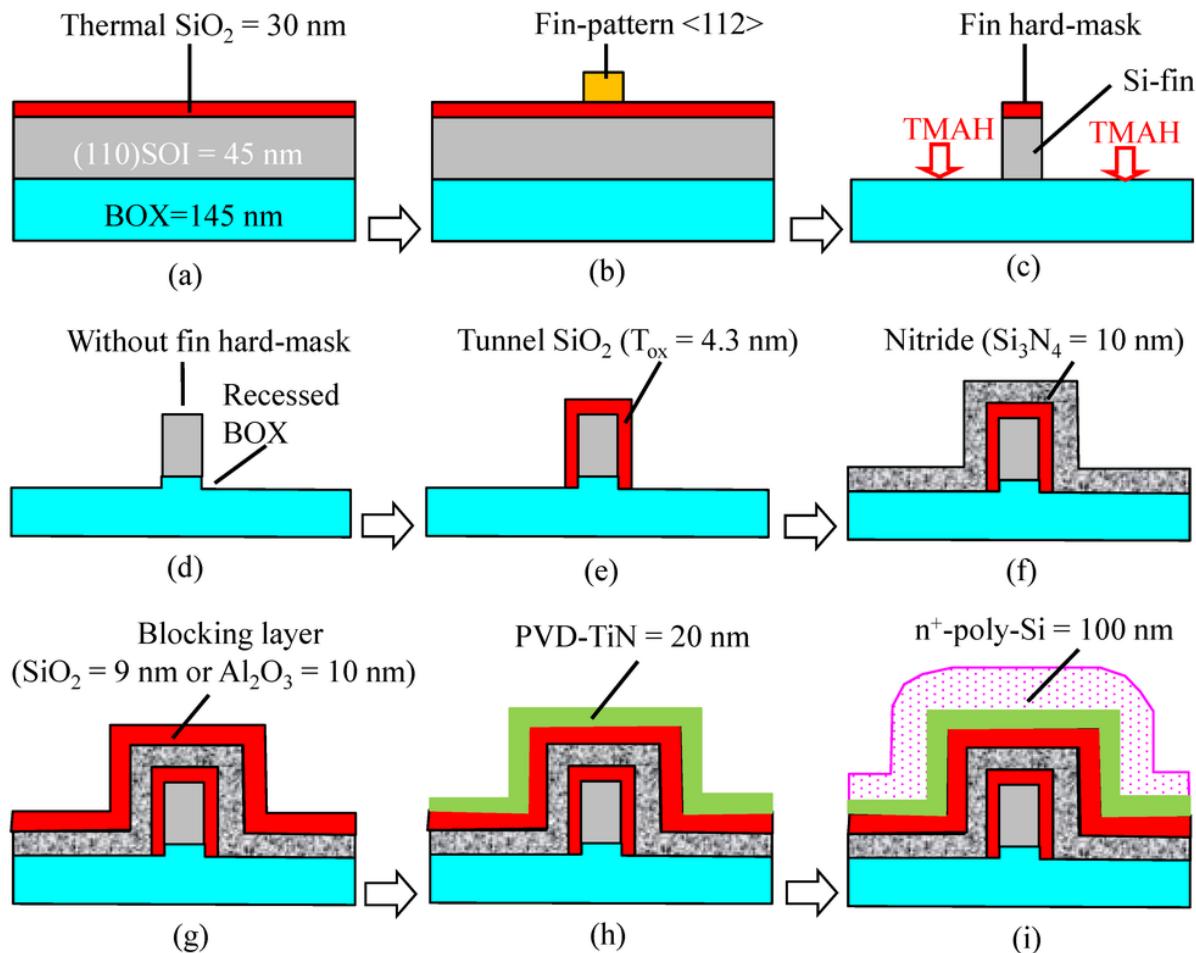
A fin field-effect transistor (FinFET) is a multigate device, a MOSFET (metal-oxide-semiconductor field-effect transistor) built on a substrate where the gate is placed on two, three, or four sides of the channel or wrapped around the channel, forming a double gate structure. These devices have been given the generic name "finfets" because the source/drain region forms fins on the silicon surface. The FinFET devices have significantly faster switching times and higher current density than planar CMOS (complementary metal-oxide-semiconductor) technology.



(a) Normal Wafer: FinFETs on regular wafers rely on a timed etch to form the fins



(b) Silicon-on-Insulator Wafer: FinFETs on SOI wafers rely on the buried oxide layer to stop the fin etch

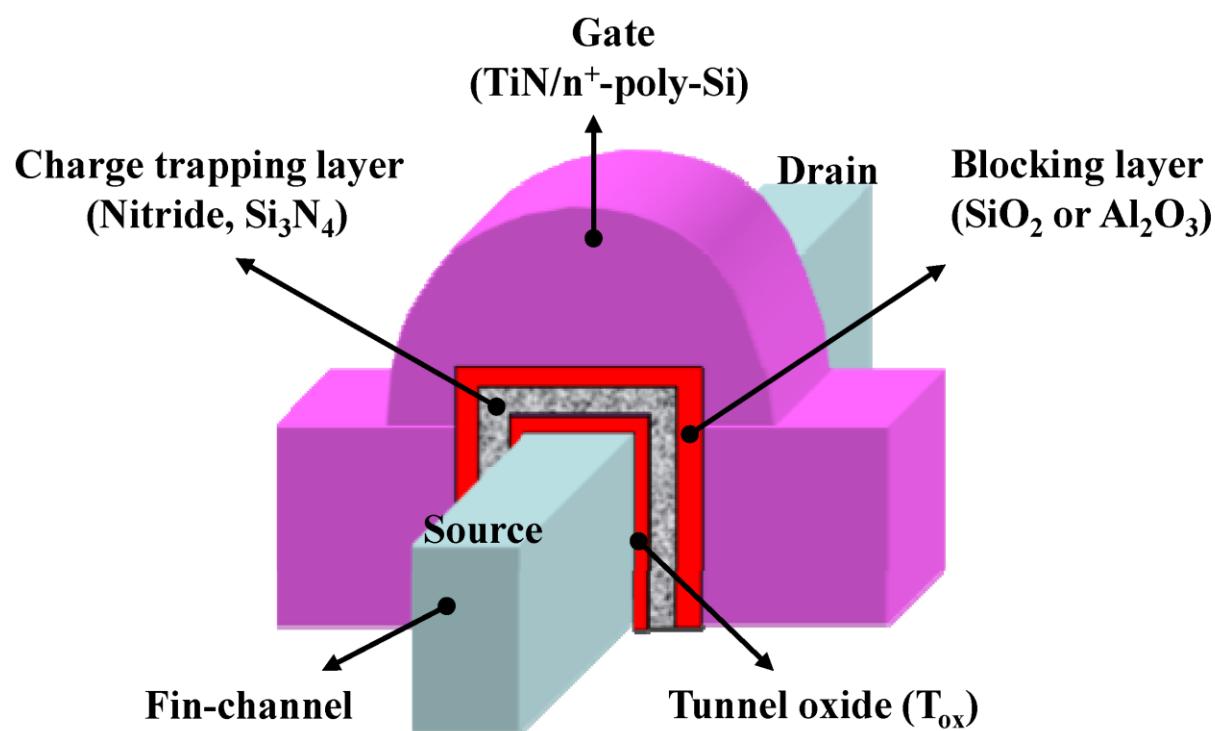


Silicon nitride (Si_3N_4) :-Silicon nitride is a chemical compound of the elements silicon and nitrogen. Si_3N_4 is the most thermodynamically stable of the silicon nitrides. Hence, Si_3N_4 is the most commercially important of the silicon nitrides when referring to the term "silicon nitride". It is a white, high-melting-point solid that is relatively chemically inert, being attacked by dilute HF and hot H_2SO_4 . It is very hard (8.5 on the mohs scale). It has a high thermal stability.

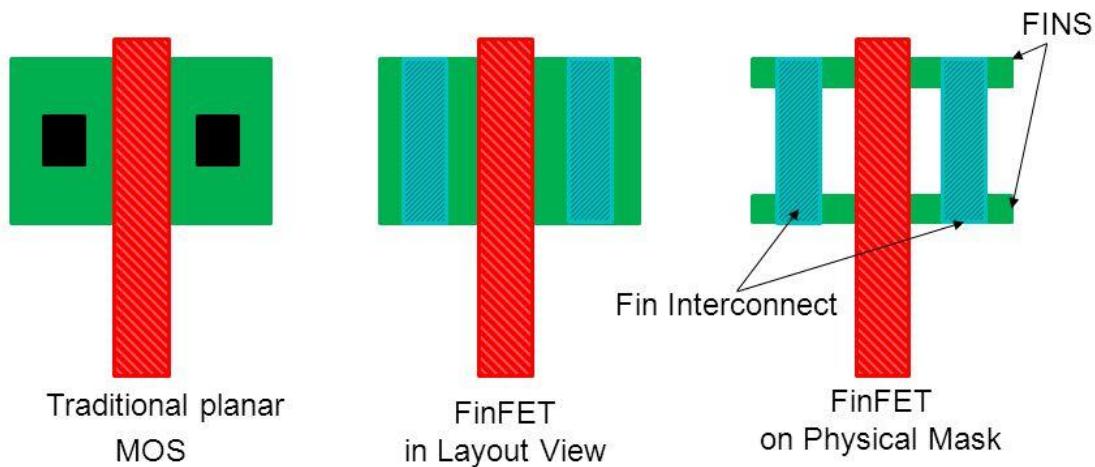
Titanium nitride (TiN, sometimes known as Tinite) :- Titanium nitride is an extremely hard ceramic material, often used as a coating on titanium alloys, steel, carbide, and aluminum components to improve the substrate's surface properties.

Silicon dioxide(SiO_2 , also known as **silica**) :- Silicon dioxide is an oxide of silicon with the chemical formula SiO_2 , most commonly found in nature as quartz and in various living organisms. In many parts of the world, silica is the major constituent of sand. Silica is one of the most complex and most abundant families of materials, existing as a compound of several minerals and as synthetic product. Notable examples include fused quartz, fumed silica, silica gel and aerogels. It is used in structural materials, microelectronics (as an electrical insulator), and as components in the food and pharmaceutical industries.

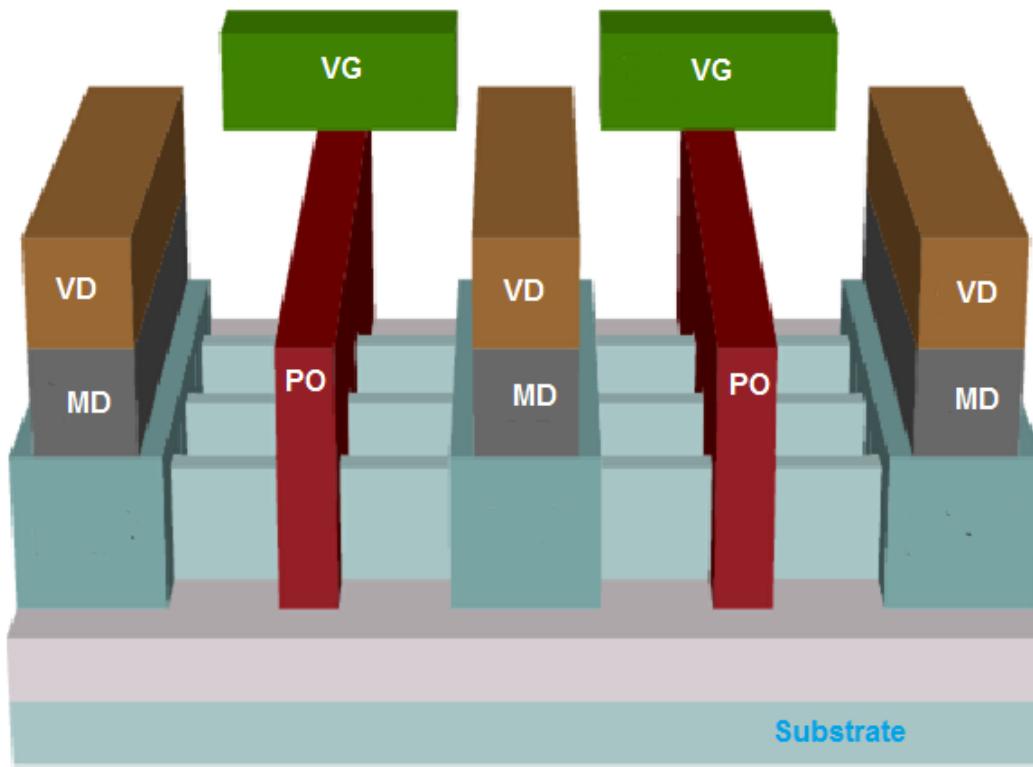
Layers of Fin:

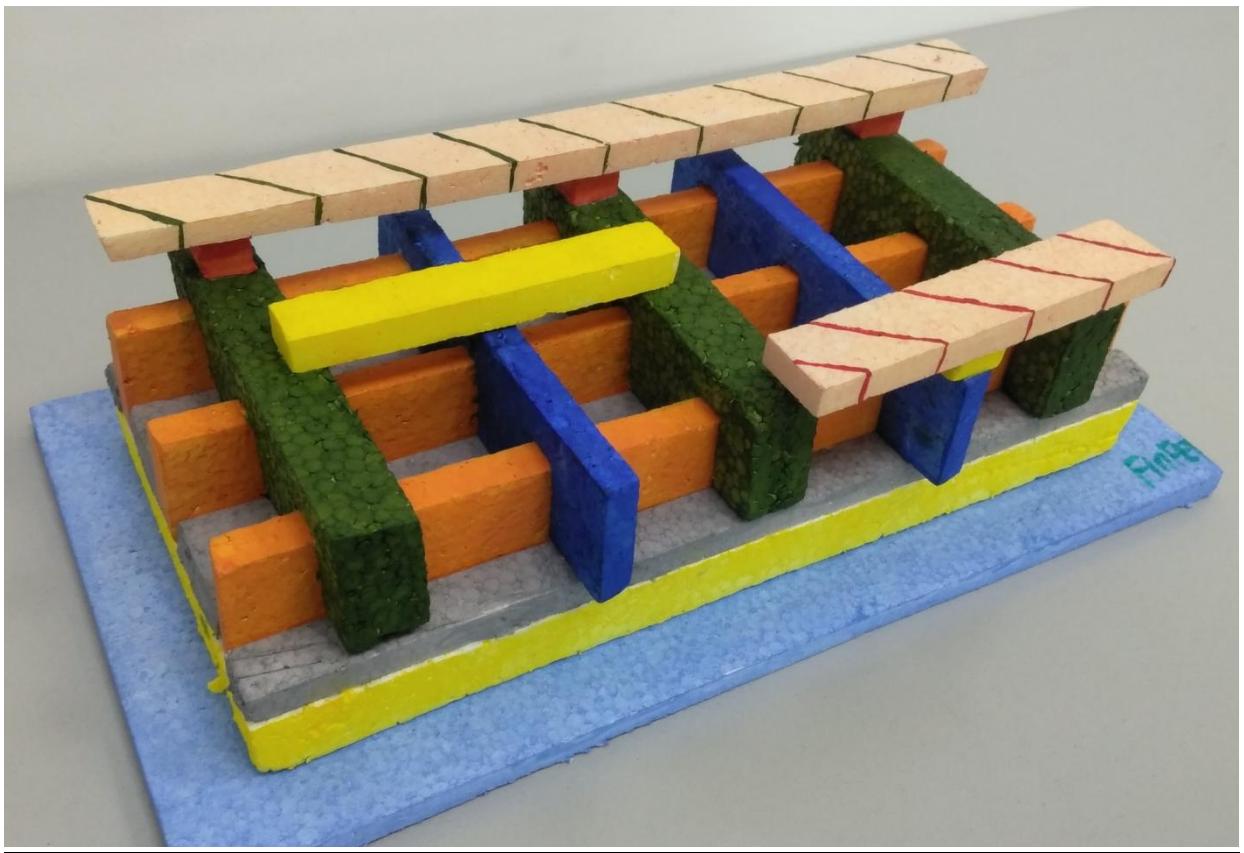


Planar Device vs. FinFET

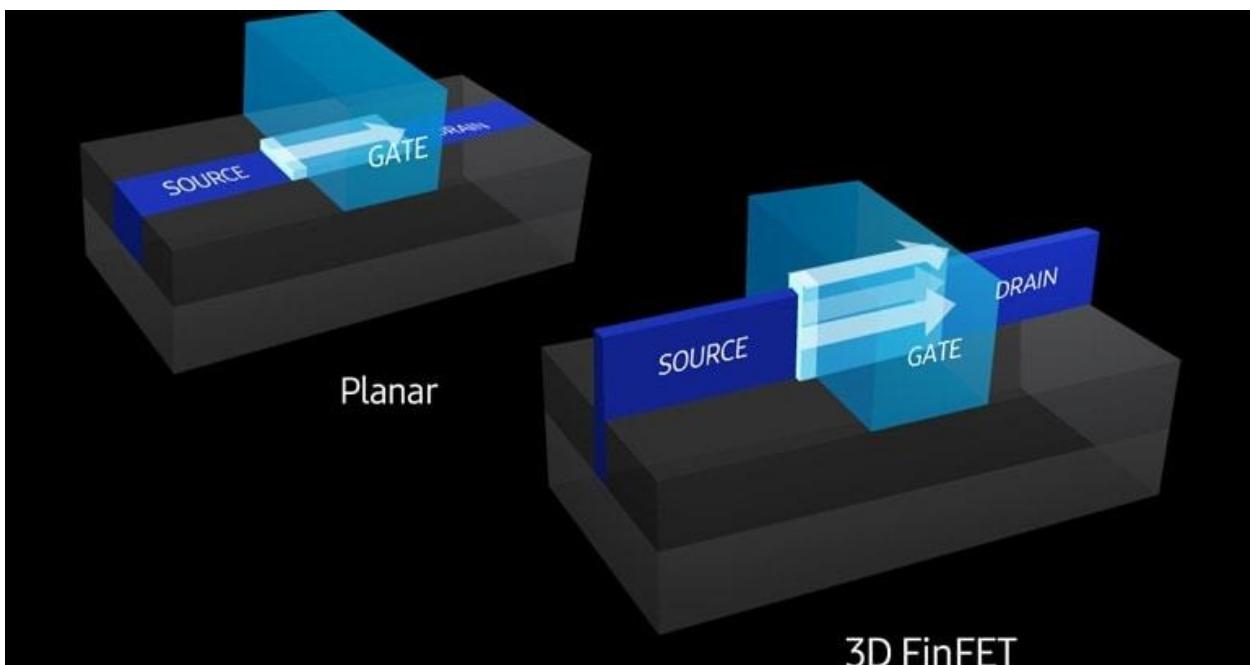


3D view of finfet:

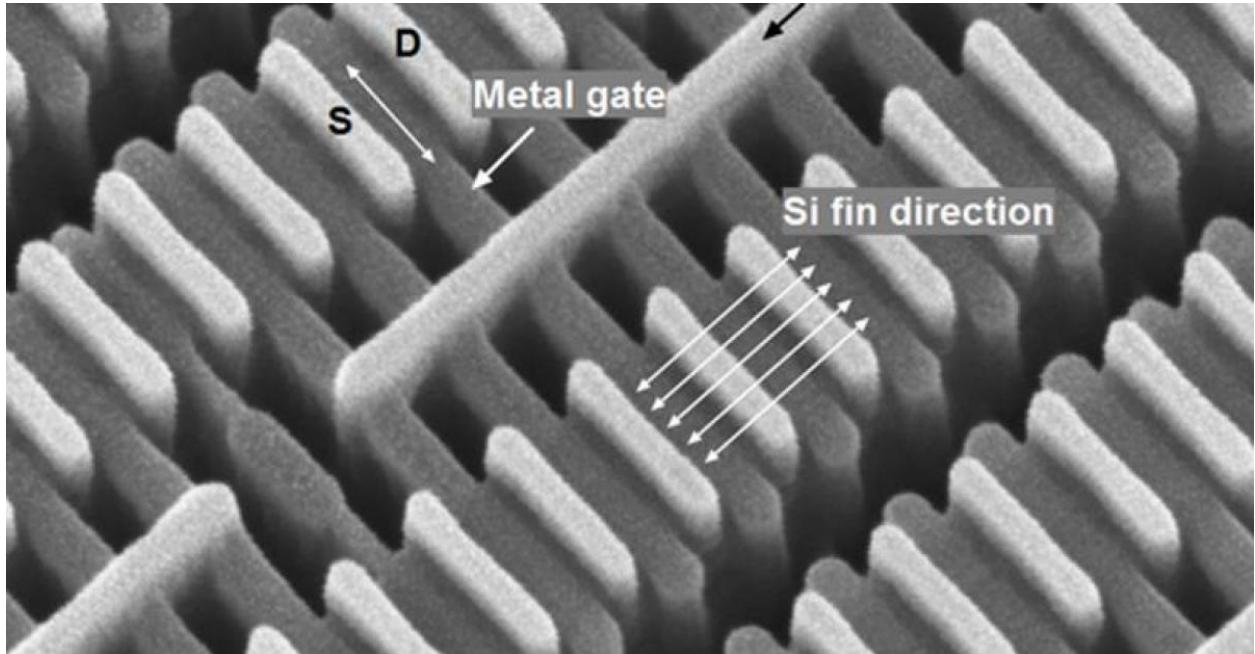




Current flow in planer and finfet:



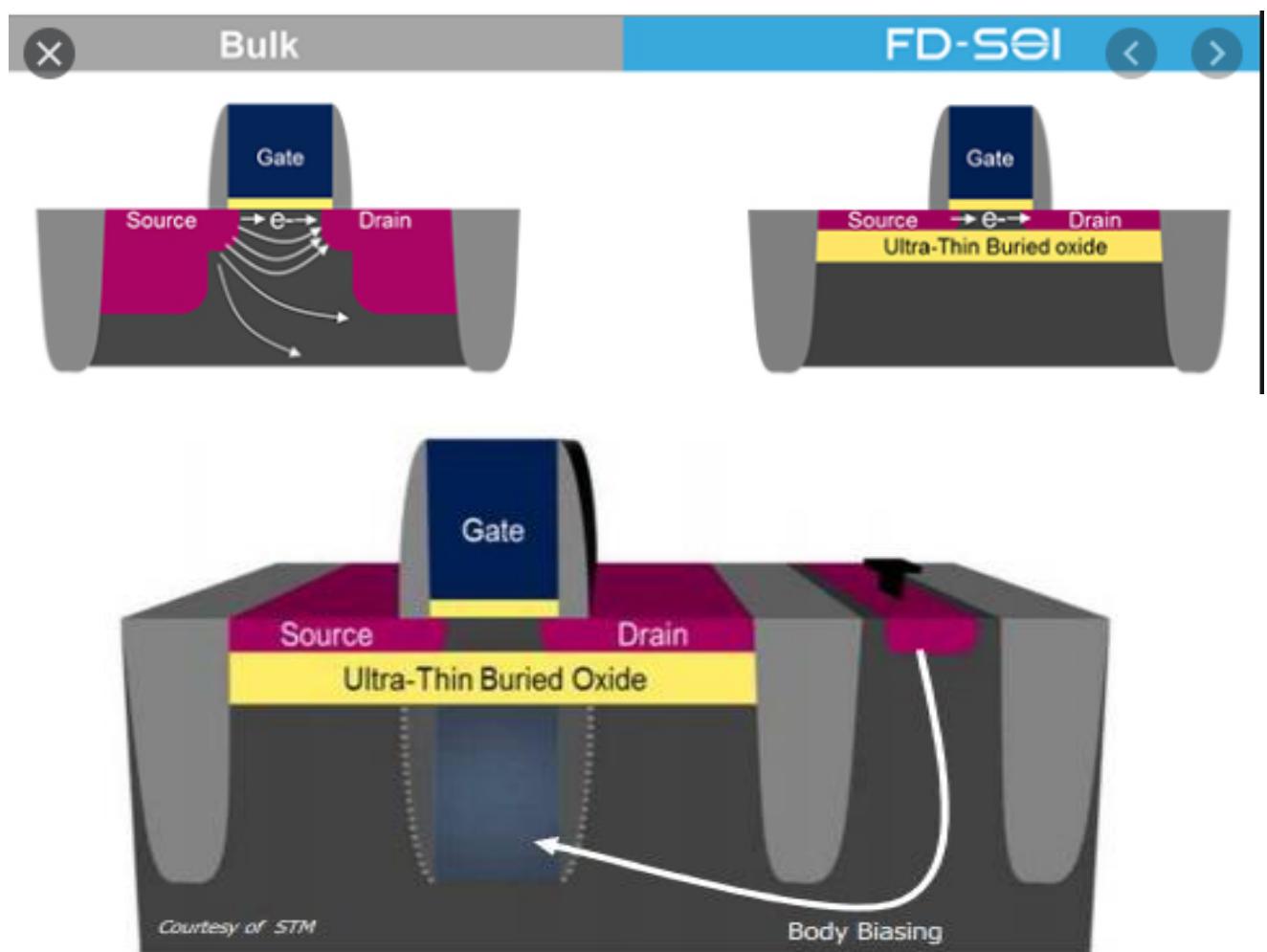
Real image of finfet:



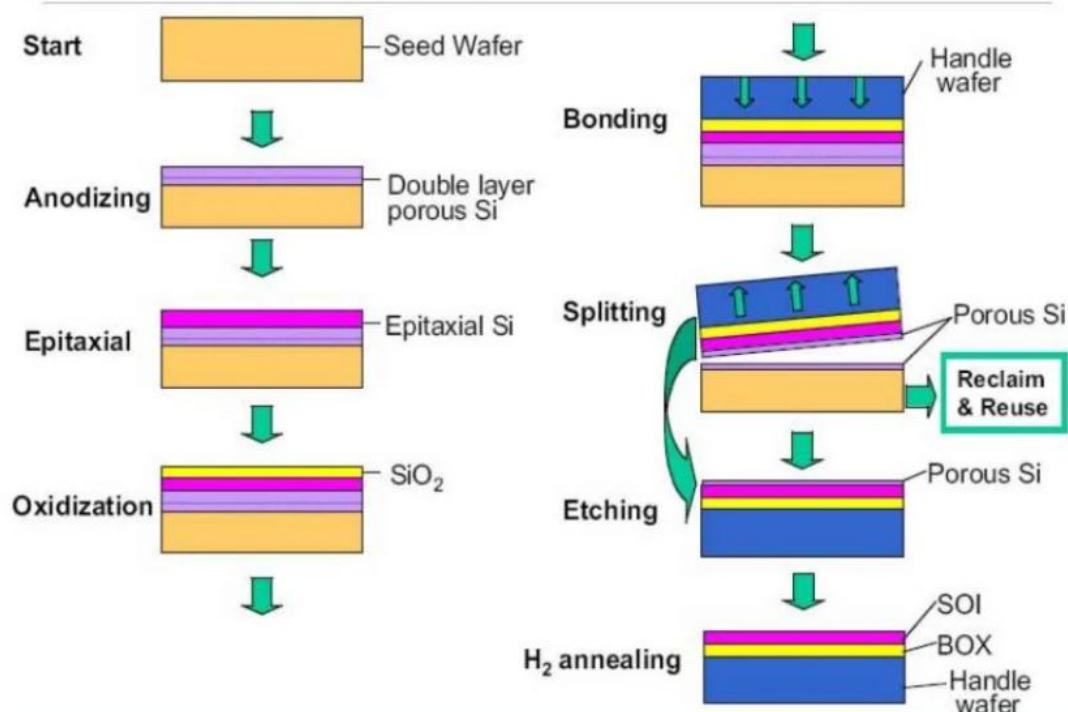
3D view link: <https://skfb.ly/oGOLK>

FDSOI:

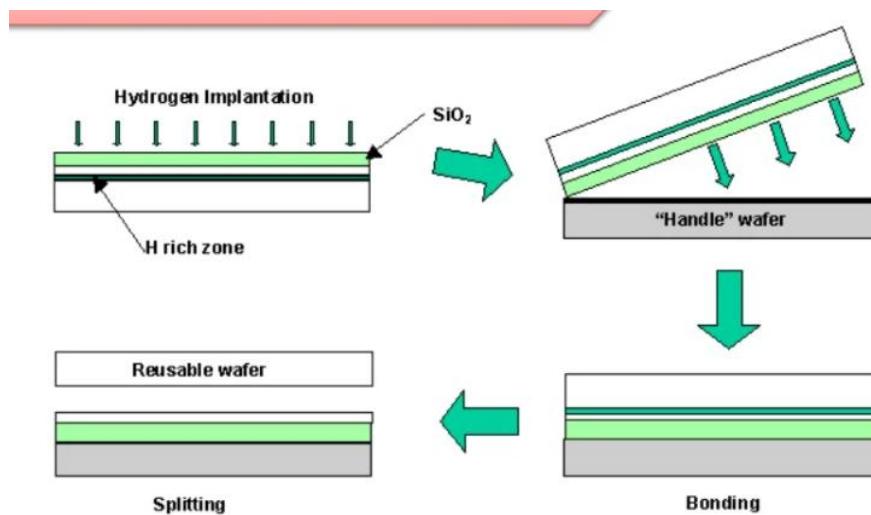
In semiconductor manufacturing, silicon on insulator (SOI) technology is fabrication of silicon semiconductor devices in a layered silicon–insulator–silicon substrate, to reduce parasitic capacitance within the device, thereby improving performance. SOI-based devices differ from conventional silicon-built devices in that the silicon junction is above an electrical insulator, typically silicon dioxide or sapphire (these types of devices are called silicon on sapphire, or SOS). The choice of insulator depends largely on intended application, with sapphire being used for high-performance radio frequency (RF) and radiation-sensitive applications, and silicon dioxide for diminished short-channel effects in other microelectronics devices. The insulating layer and topmost silicon layer also vary widely with application.



ELTRAN (Epitaxial Layer TRAnsfer)

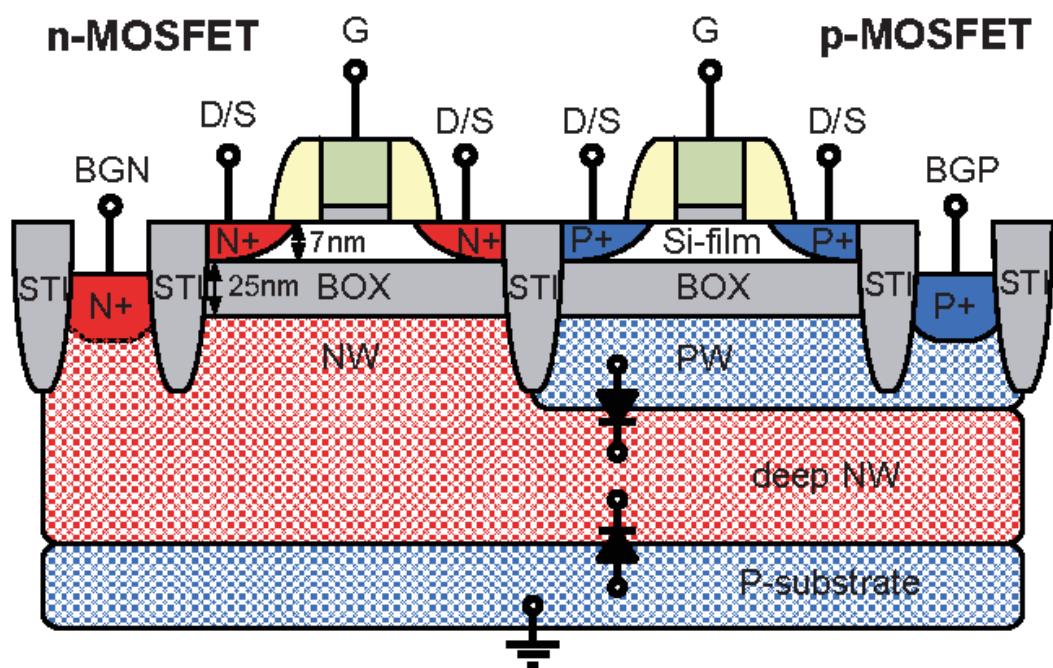


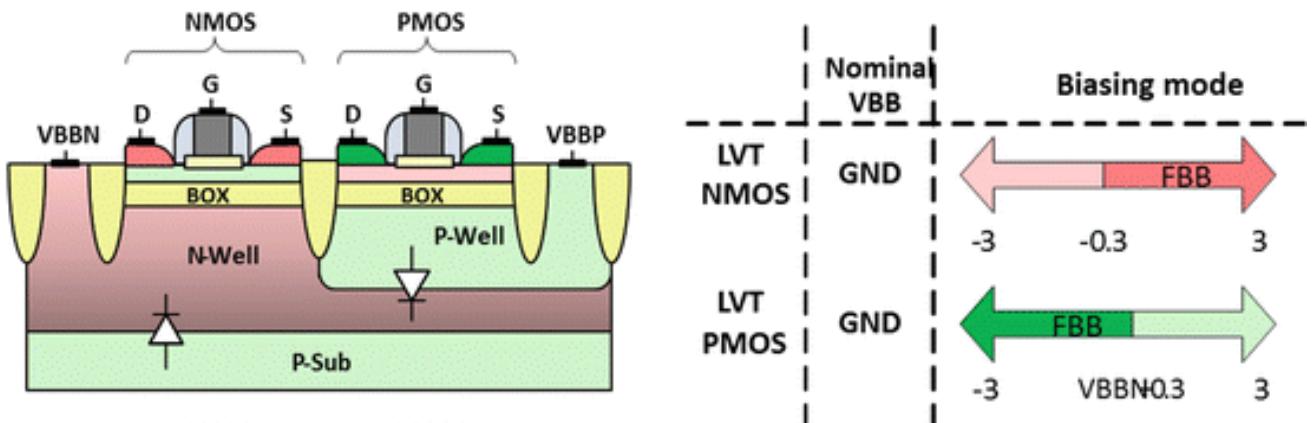
It is a technology developed by Canon



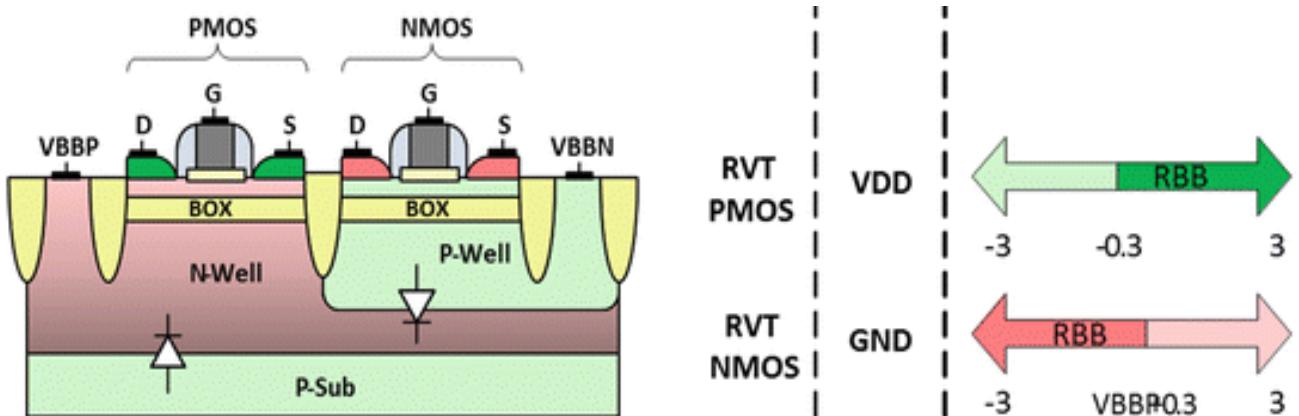


LVT FDSOI





RVT FDSOI



Advantage of FD-Soi:

- Better Electrostatics
 - Faster Operation/ Higher clock speed for mobile chips
 - Low Voltage/ only 0.6V power consumption in mobile device
 - Less power consumption in mobile device
 - Lower production cost
- Total dielectric isolation
 - Lower S/D leakage
 - Lower S/D Capacitance
 - Latch-up immunity
- Lower leakage current
 - Less sensitive to temperature
- No channel doping
 - Improved V_t Variability
 - Improved mismatch (SRAM & analog)
 - Better analog gain
- Less Self heating effect

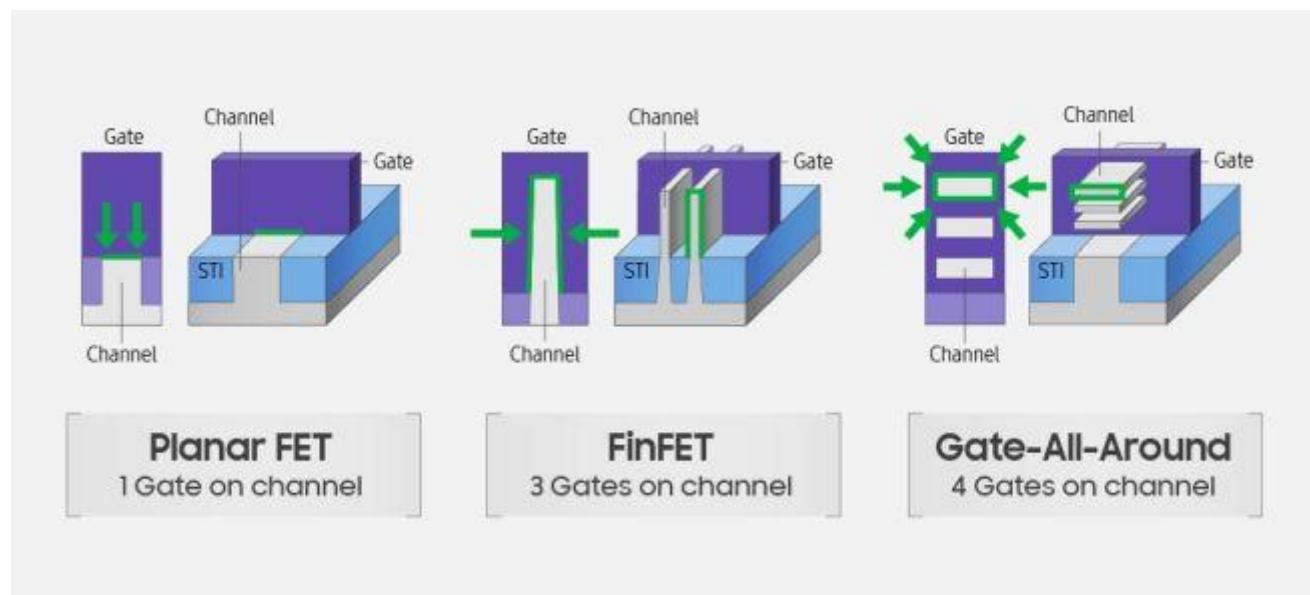
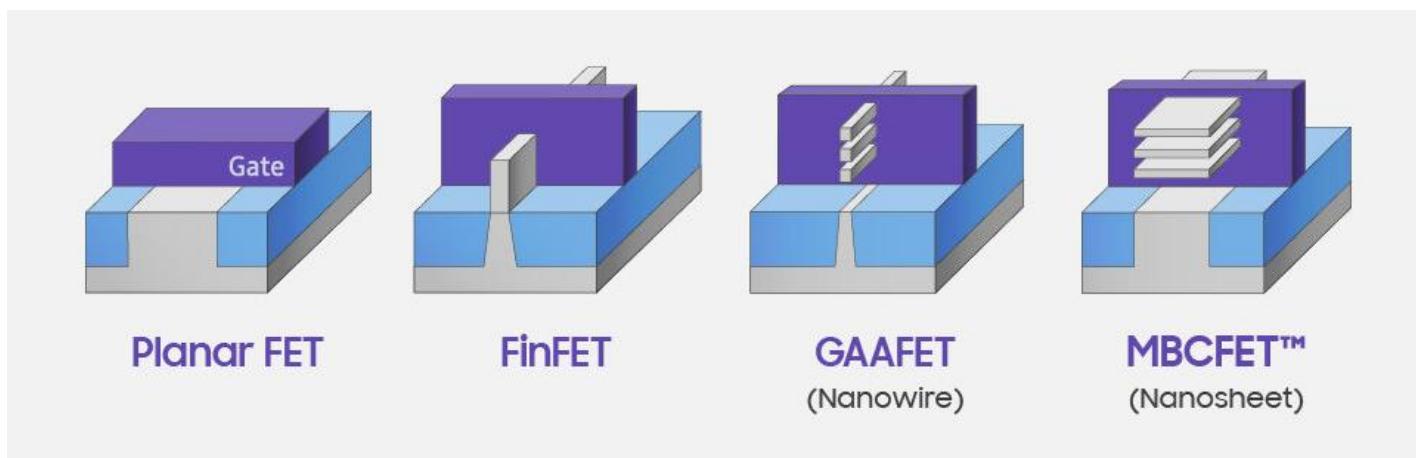
3D view link: <https://skfb.ly/oGOMr>

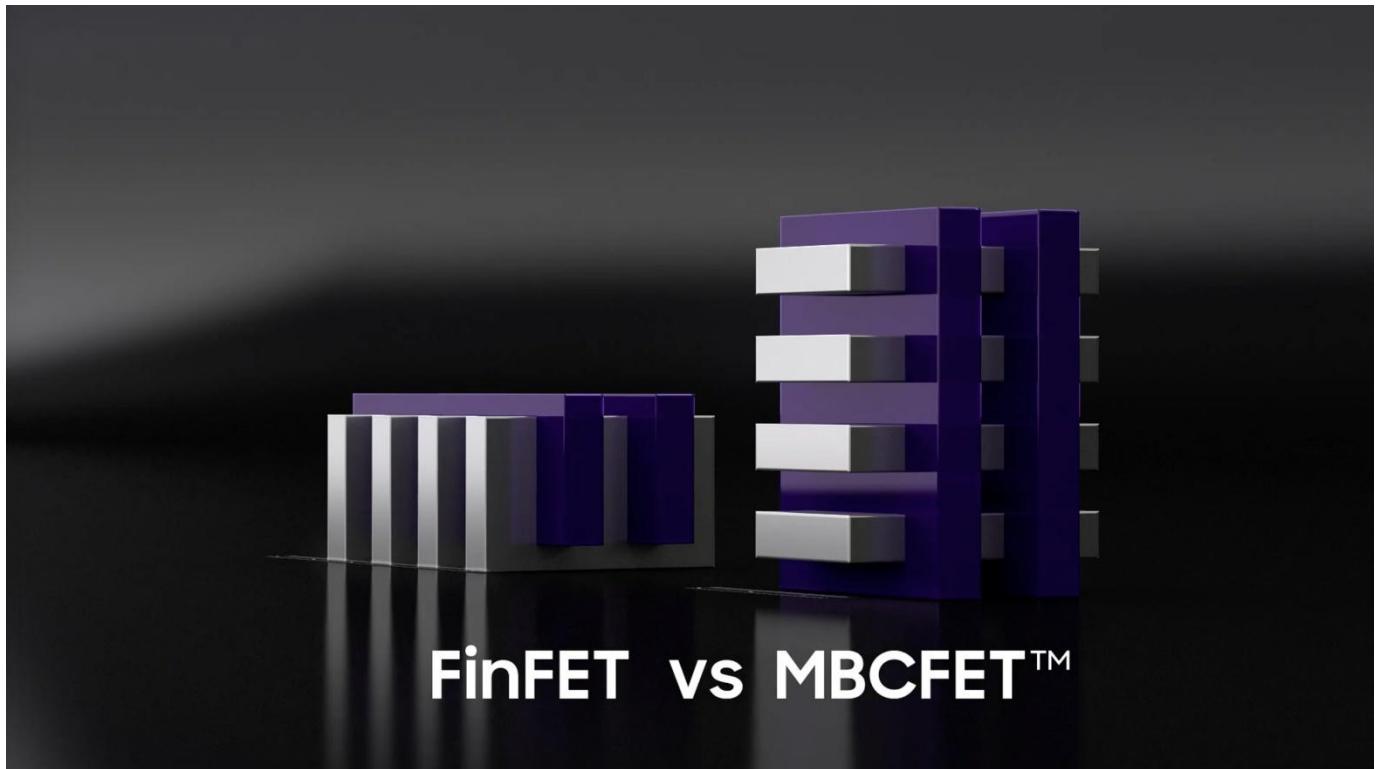
Generation of MOSFET:

FinFET: Fin Field-Effect Transistor

GAAFET: Gate-All-Around FET

MBCFET: Multi-Bridge Channel FET





Voltage rating of different technology node:

Node	Voltage domains
350nm HV	3.3V 15V
250nm BCD, gen. I and II	12V 24V 40V 60V
180nm BCD, gen. I and II	18V 24V 32V 40V 60V
180nm CMOS	1.8V 3.3V 5V
130nm CMOS	1.0V 1.2V 3.3V 5V 7V
90nm CMOS	1.2V 1.8V 3.3V
65nm CMOS	1.0V 1.2V 1.8V 2.5V 3.3V 5V
40nm CMOS	0.9V 1.2V 1.8V 3.3V 5V
28nm CMOS	0.85V 0.9V 1.8V 3.3V 5V 5.5V 12V
16nm FF+	0.8V 1.2V 1.5V 1.8V 2.5V 3.3V 5V

Video link for MOSFET :

1. Planer : <https://youtu.be/stM8dgcY1CA>

2. FinFet: https://youtu.be/TXxw1kdF5_Q

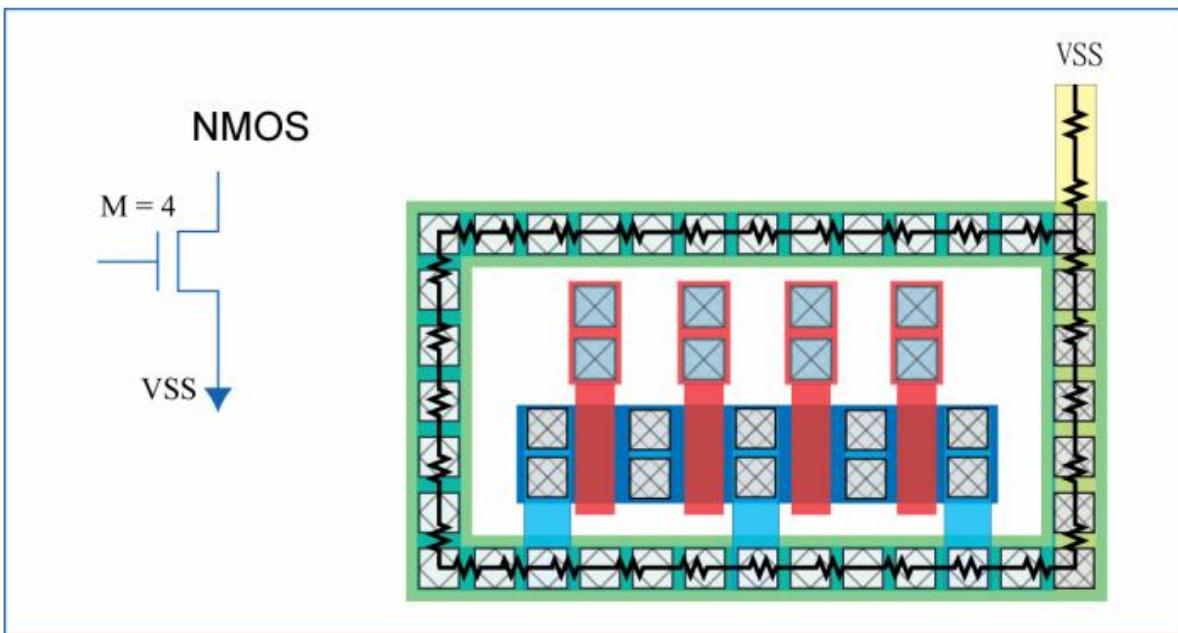
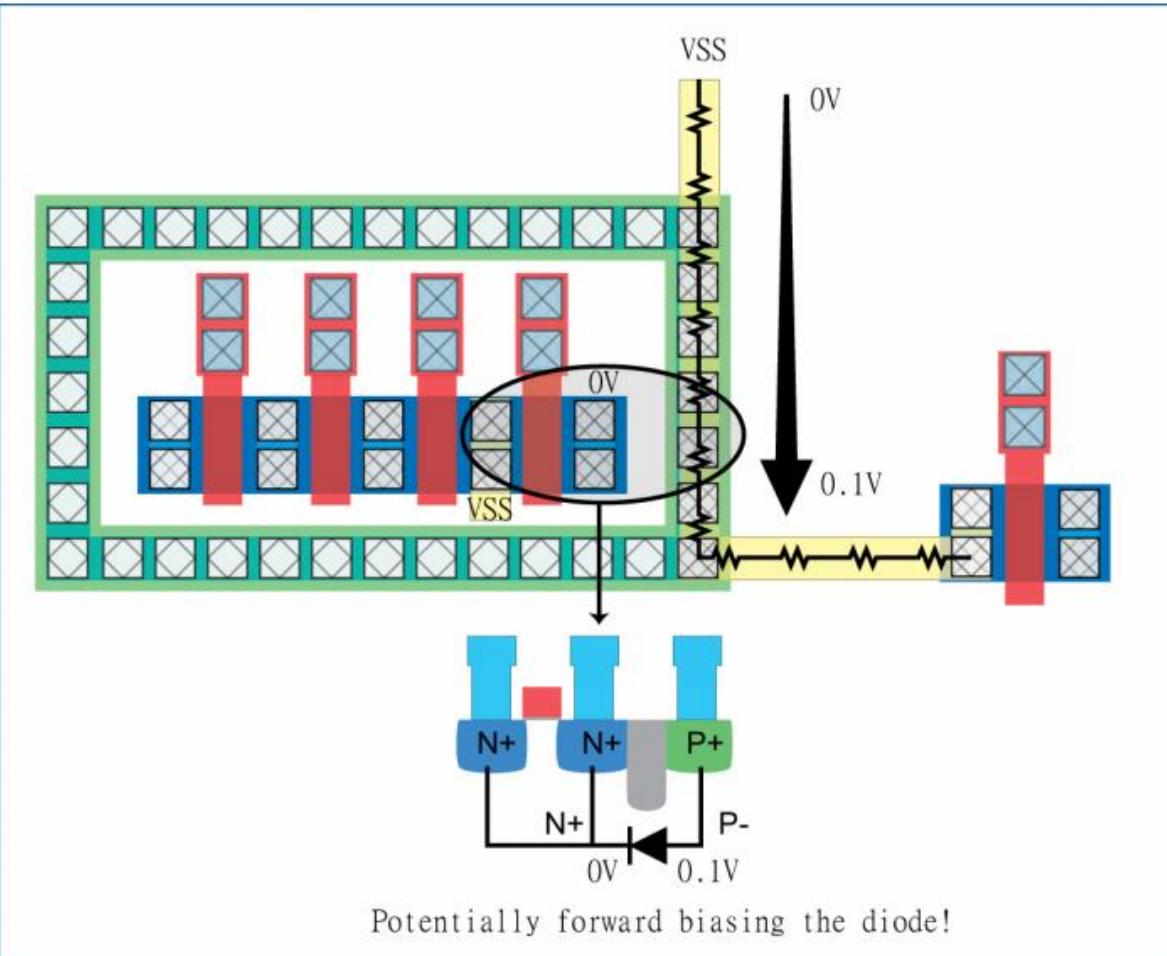
3. Fdsoi: <https://youtu.be/uvV7jcpQ7UY>

4. Fdsoi: <https://youtu.be/qLlwGXhEcl4>

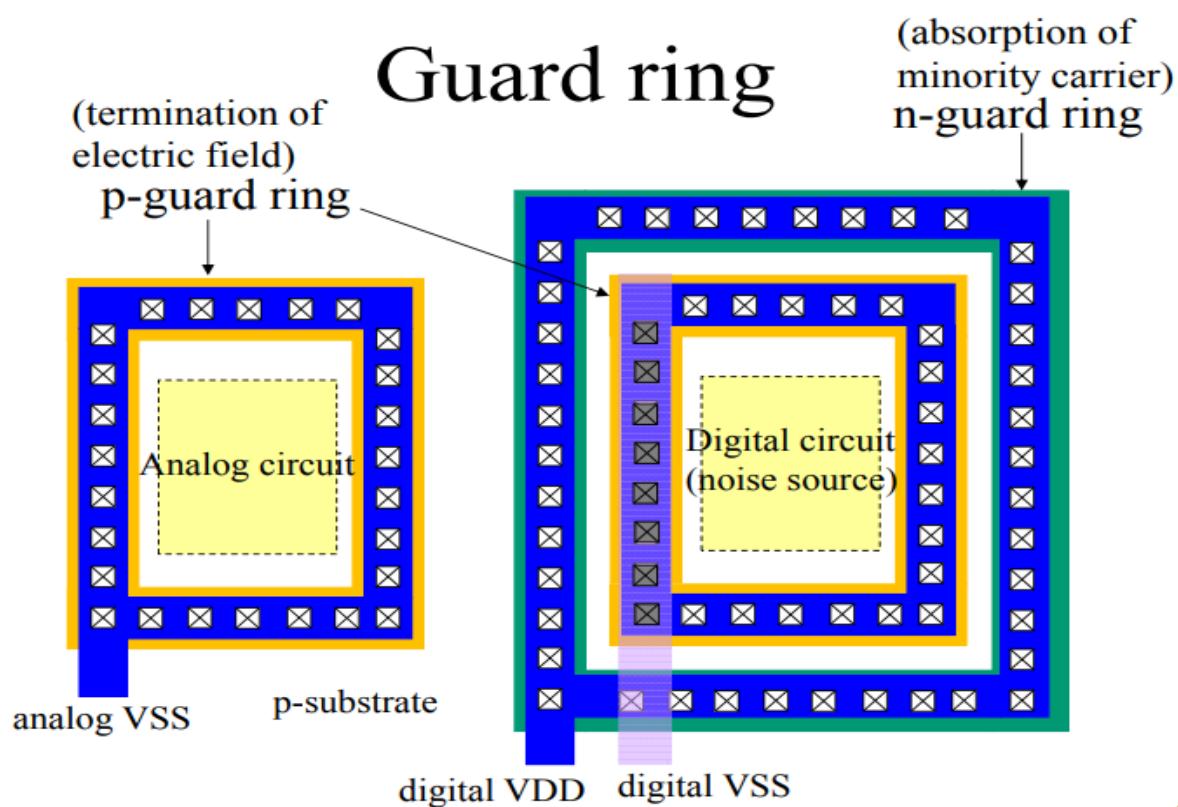
5. GAAFET: <https://youtu.be/3otqUu-7WUQ>

3D view link of GAAFET: <https://skfb.ly/oGOMz>

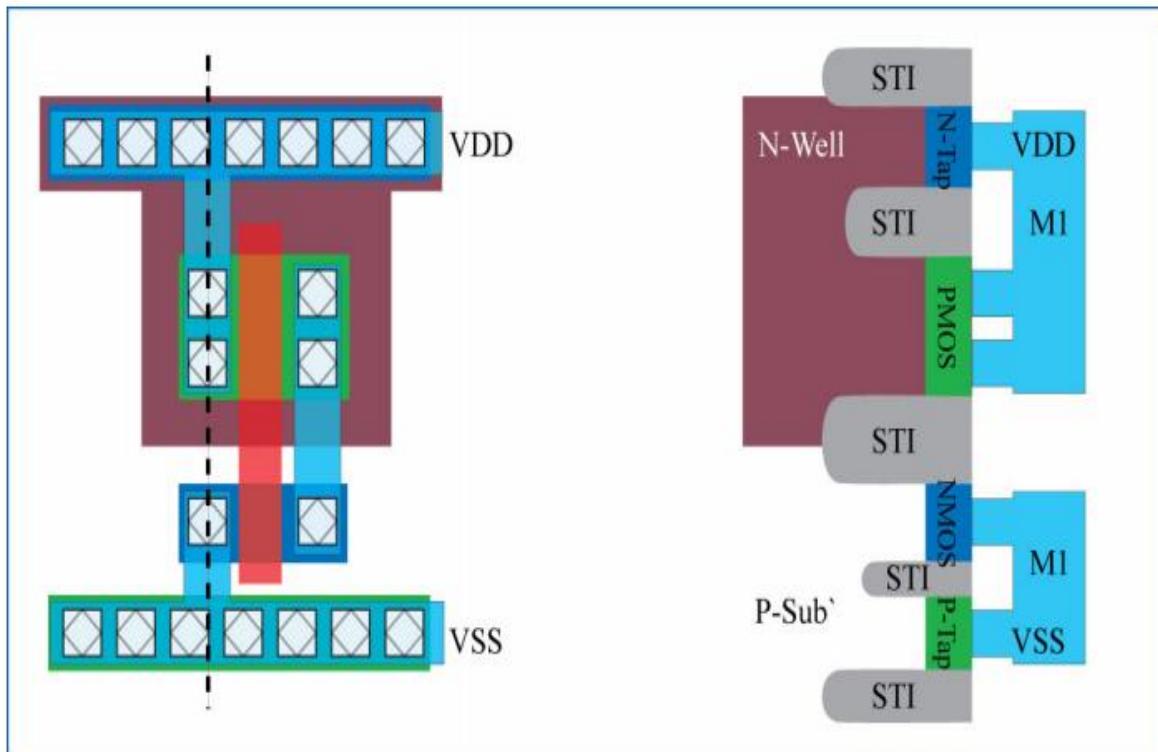
Guardring:



Guard ring

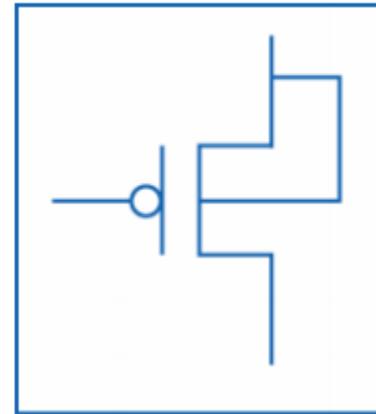


41



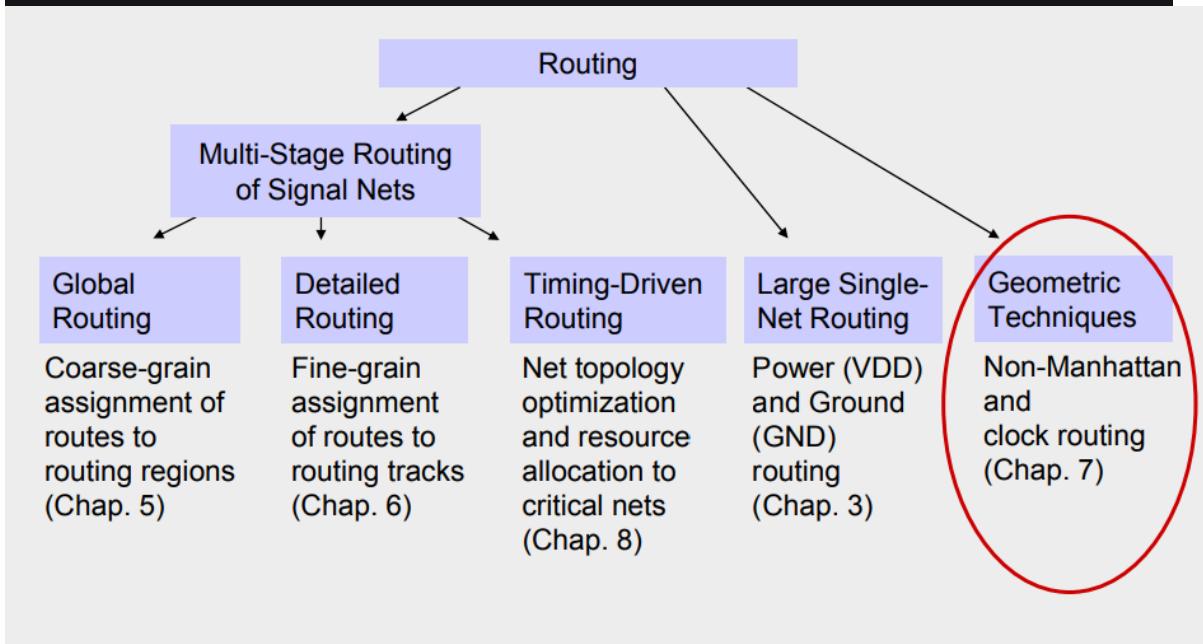
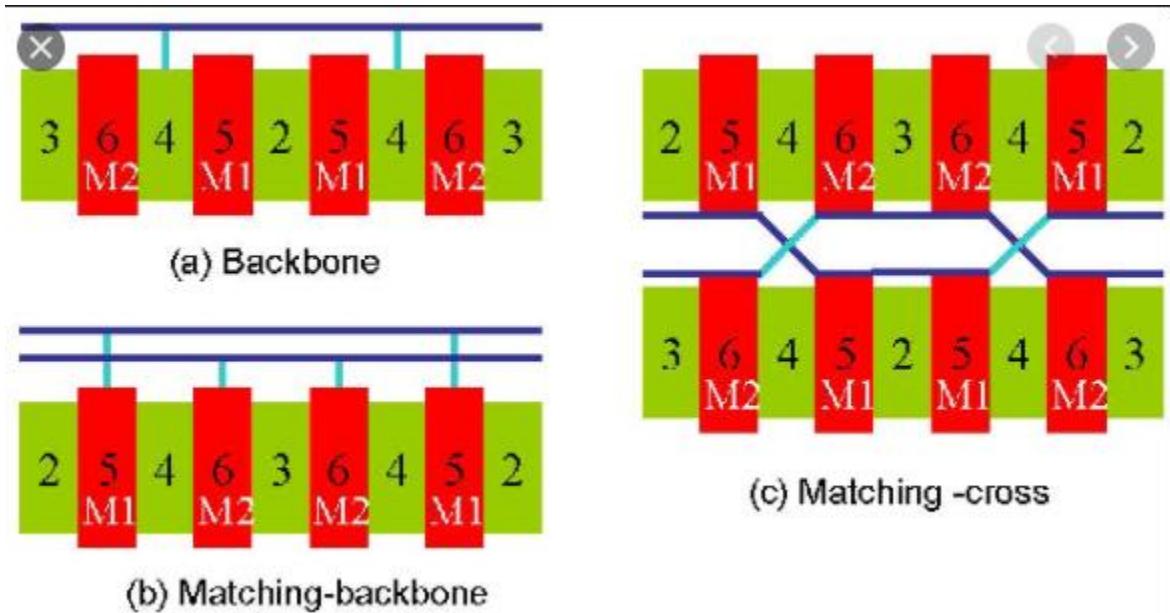
Floating Well:

Wells that are tied to and varied with the source terminal of the transistor are called floating well, as illustrated on the right. Floating well is quite commonly used in analog design to reduce the “body effect” of the transistor. The PMOS with this type of bulk connection needs an N-well on its own, and hence the layout has a huge area overhead. It is important to identify all the PMOS that sit on a floating well before floor plan the layout.

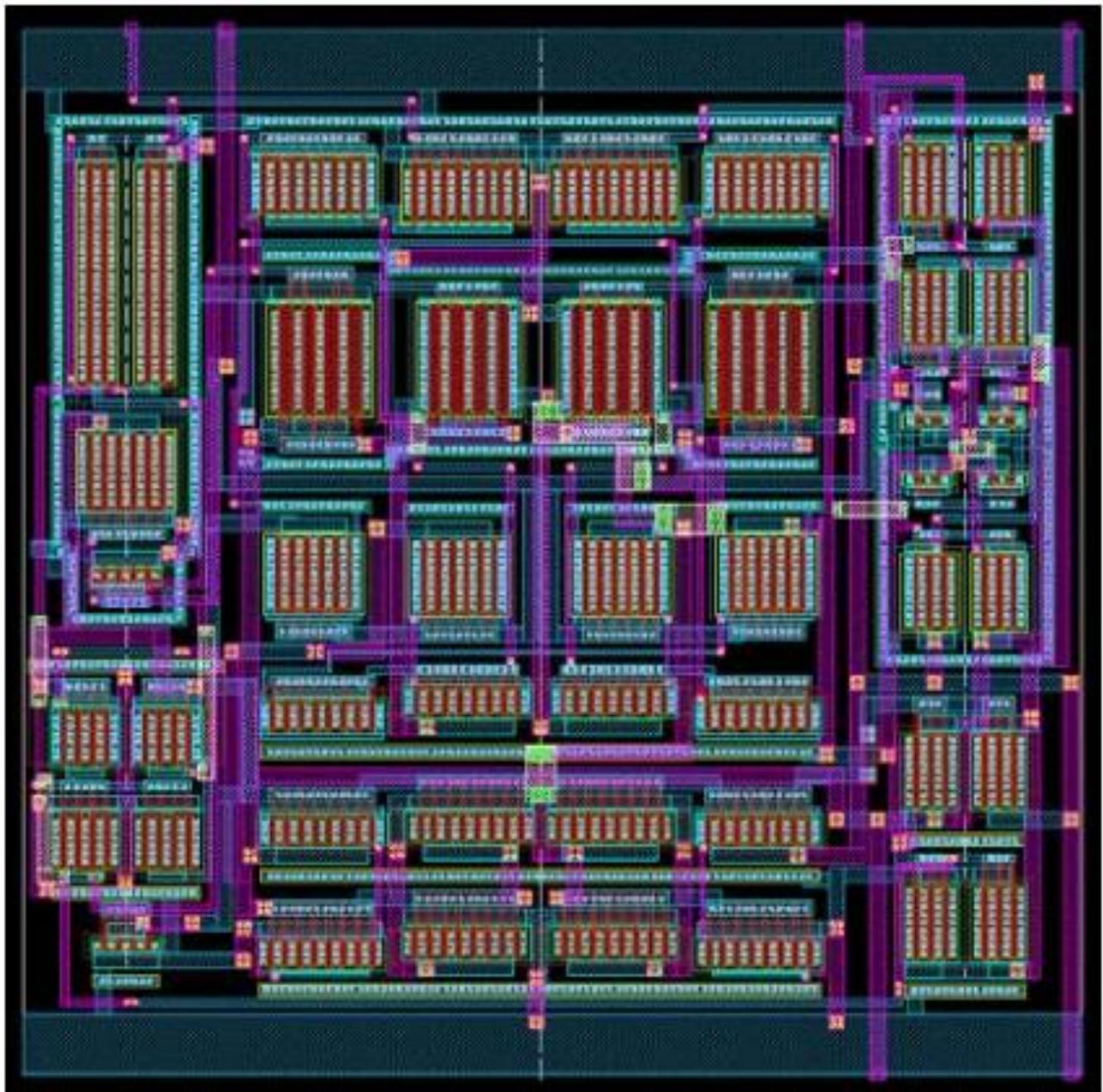


9.ROUTING IMPORTANCE

Signal Routing Matching :

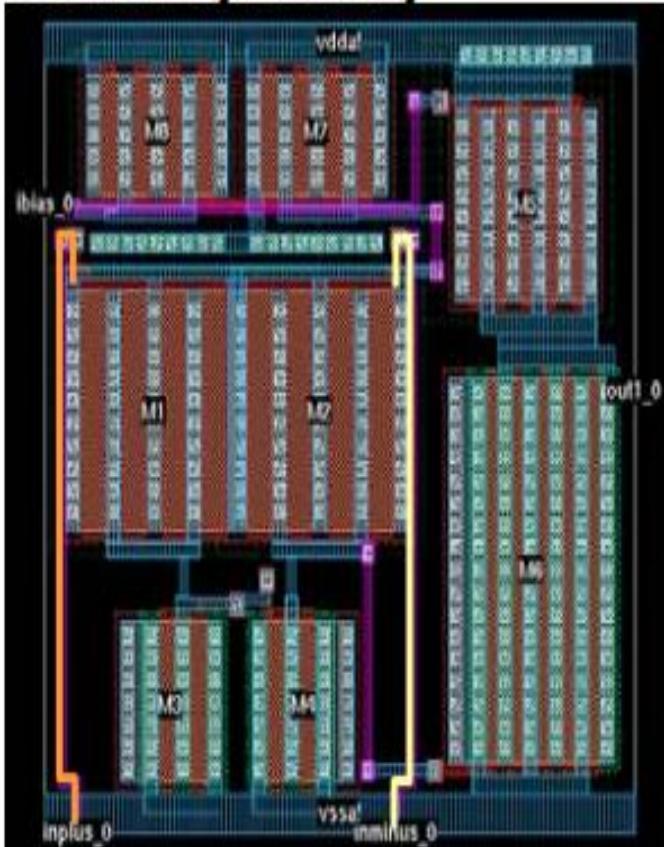


➤ Example of routing in Layout :

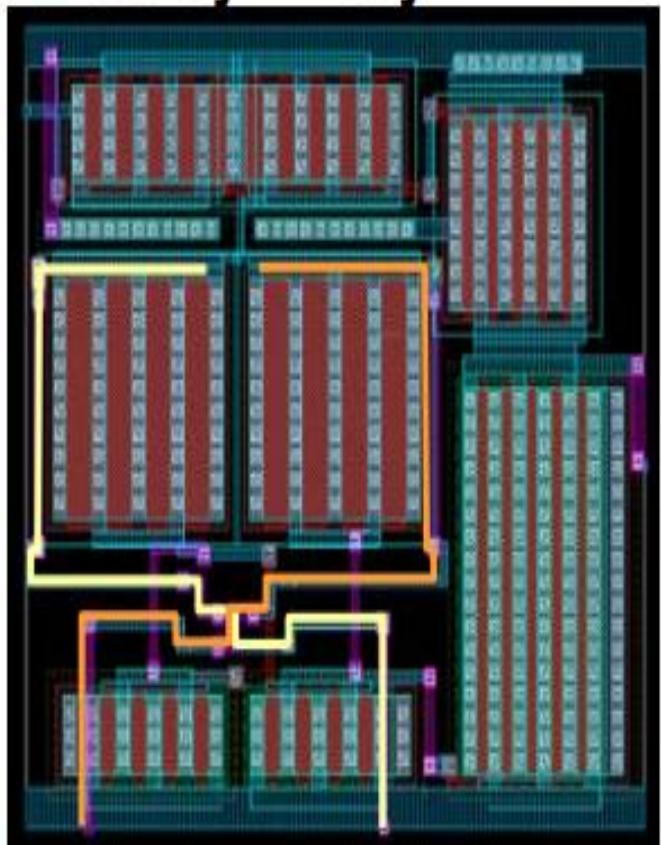


➤ A Few of the options for symmetric Nets :

Mirror symmetry



Cross symmetry

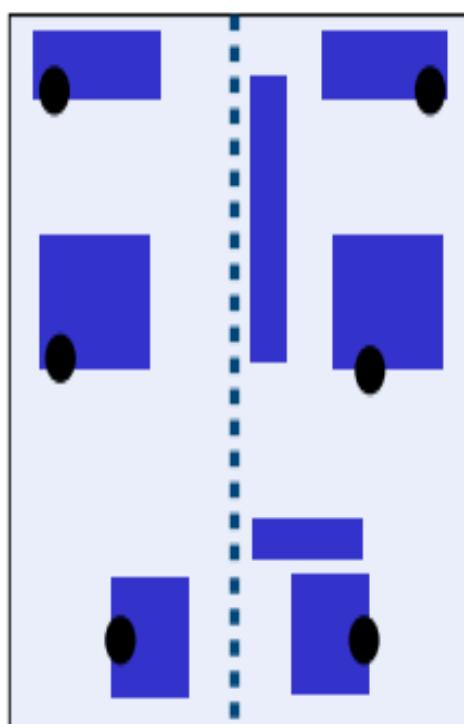


■ Complications

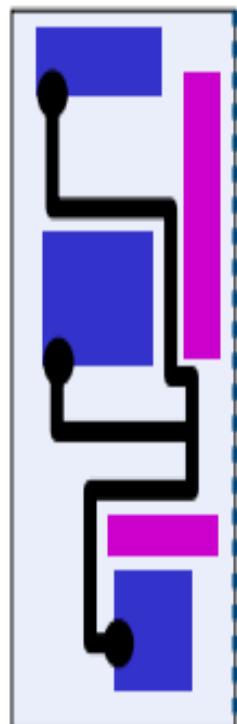
- There are lots of forms of symmetries, letting designers specify them easily is tough
- Sometimes, the pins are “not quite symmetric” or there are a few extra non-symmetric pins on the net. Still need to route “most” of the net as symmetrically as possible

➤ **Symmetric Routing: Basic Trick**

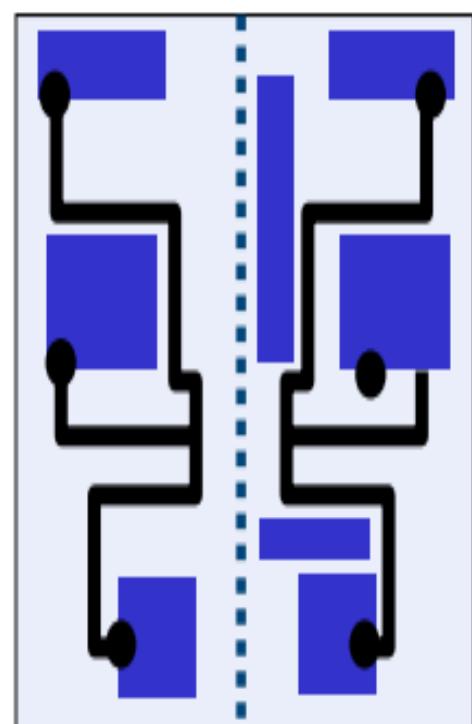
- Only route **one** wire, but **reflect obstacles** from other side across symmetry line, into one **shared** left-right model of space



Symmetry
line

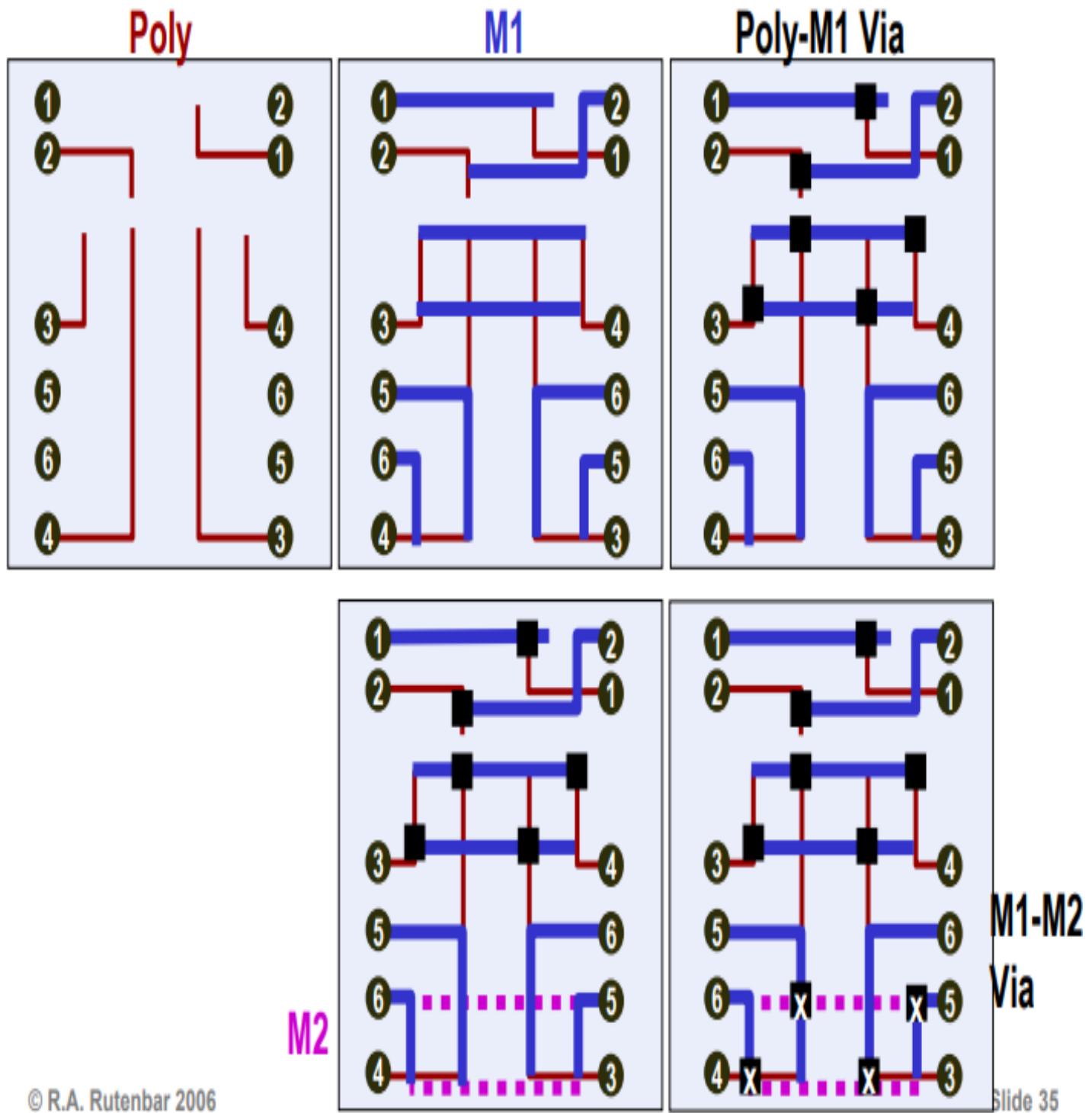


Shared LR
model, route
1 wire here



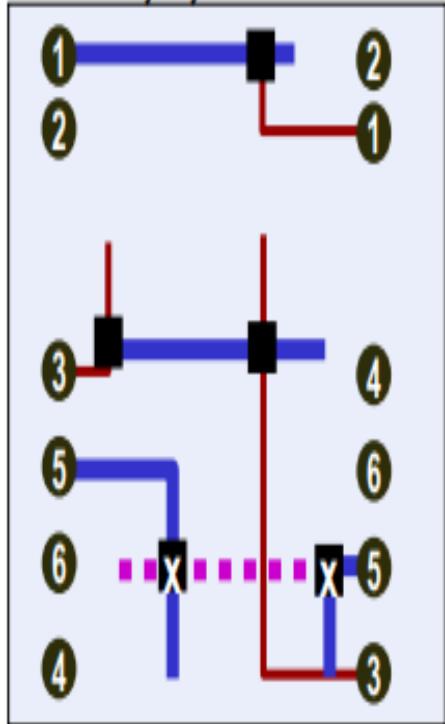
Reflect single routed wire
back across sym line

➤ Detailed Solution to Balanced Route example

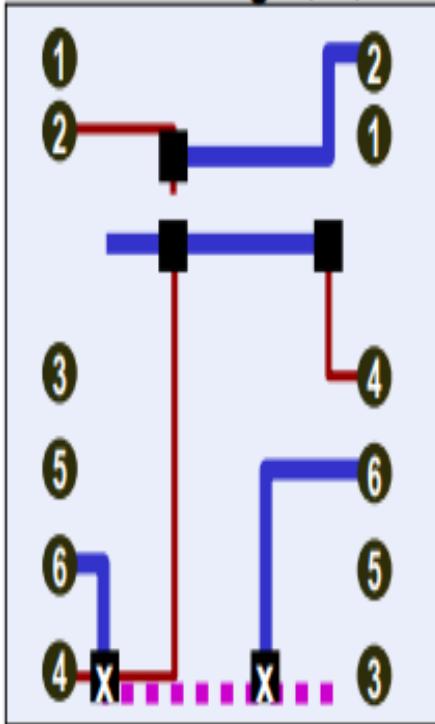


➤ Detailed solution to balanced route example :

Nets 1, 3, 5...



...matching 2, 4, 6



Each net pair has ~same length on each layer, same num and type of vias

■ Observations

- Not every dense arrangement of pins (with obstacles) can be routed
- Much of this problem is getting the placement right, with space reserved
- Routing here much more like channel-ed problems, with more constraints
- Can attack these as routing problems, or as “wire placement” problems

Physical Design Routing Process :

- **Metal routing:**

- a. It involves generating metal wires to connect the pins of same signal while obeying manufacturing design rules.
- b. Before routing is performed on the design, cell placement has to be carried out wherein the cells used in the design are placed.
- c. The connections between the pins of the cells pertaining to same signal need to be made. At the time of placement, there are only logical connections between these pins.
- d. The physical connections are made by routing. More generally speaking, routing is to locate a set of wires in routing space so as to connect all the nets in the netlist taking into consideration routing channels' capacities, wire widths and crossings etc.
- e. The objective of routing is to minimize total wire length and number of vias and that each net meets its timing budget. The tools that perform routing are termed as routers.
- f. You typically provide them with a placed netlist along with list of timing critical nets. These tools, in turn, provide you with the geometry of all the nets in the design.

- **Stages of routing :**

- I. GLOBAL ROUTING
- II. TRACK ASSIGNMENT
- III. DETAIL ROUTING
- IV. SEARCH & REPAIR

1. Global Routing

- In global routing, the region to be routed are divided into sectors (tiles/rectangles) called global routing cells.
- Then it decides tile to tile path for the nets and simultaneously trying to optimize the length, without actually making any physical connection.
- The routing capacity of each cell depends on the blockages, routing tracks, pin density inside it.
- This rough routing is done on the basis of available tracks in the region.
- If the required routing resources are greater than the available routing resources, then it will lead to congestion.
- So it is called coarse grain routing assignment.

Objectives of global routing

- Minimize total overflow
- Minimize total wire length
- Minimize total run time for carrying out routing process

2. Track Assignment

- After the cell estimation, tracks are assigned to each global routes.
- The tracks are assigned in vertical and horizontal direction for each partition
- The direction of routing is dependent on the metal used, which has preferred routing direction.
For eg. If metal1 has routing direction Horizontal, then Metal2 has direction Vertical
- In this stage, the global routes are replaced with metal layer, which has many DRC violations, Signal Integrity (SI) and timing violations

3. Detailed Routing

- In detailed routing, the router uses the scheme made in the global routing and track assignment phases to lay metals to connect the nets to the pins
- The violations that were created in the previous stage, will be fixed by multiple iterations of routing, so that no connections will be left short, open or spacing violations.
- First, the block is divided into specific areas called the sboxes (switch boxes) which comprises of multiple cells
- These boxes are in alignment with the cell boundary.

4. Search and Repair

- It is done along with detailed routing, specifically after the primary iteration
- The shorts and spacing violations are sorted and is fixed

Fill Stage

Fill stage comes after routing optimization, where filler cells and metal fills are added to meet the DRC rules. Two steps are mainly performed in this stage.

1. Adding filler Cells
2. Adding Metal Fills

Adding Filler Cells

- Filler cells are used for rail continuity and to fill up gaps between standard cells in the row, and thereby reducing the DRC violations created by the base layers
- Filler cells are physical-only cells designed in such a way that they contain only n-well, p-well and power rails
- It is also possible to reduce the IR drop by inserting de-cap filler cells, but this comes at a cost of higher leakage current

Adding Metal Fills

- The metal fills also known as dummy metal layer, are small, floating metal nets, inserted in empty spaces in the design after post-route optimization in order to maintain uniformity in metal layer density
- These are added to meet to metal density DRC rules (density violations) which are mandatory by most manufacturing processes

Noise and Shielding:

Interconnect shielding is used in VLSI designs to avoid noise interference from the cross-coupling capacitance between adjacent signals. This paper takes advantage of the shields already present in the design, and uses them to tune the propagation delay of the clock signals, thus eliminating expensive dedicated delay buffers. The problem of obtaining the desired delay at a minimum shielding cost (silicon area) is formulated as a calculus of variations problem. An analytical solution shows that a square root shield profile is optimal. The shield should connect to a quiet low-impedance node such as an analog ground line.

Noise in Power and Ground:

- For the low noise circuit, a p-well ring, which is tied to VSS is recommended to surround all PMOS device in each analog circuit block
- For the low noise circuit, a N-well ring, which is tied to VDD, is recommended to surround all NMOS device in each analog circuit block
- Put NMOS in RW (NW hole in DNW) is a good practice of isolating critical circuit from substrate noise. Make sure every NW connected to DNW must be the same potential
- Use separate power supplies and ground buses for the noisy and sensitive circuit and also for the analog and digital circuits
- Keep enough distance b/w the noisy and sensitive area use guard ring to stabilize substrate and well potential
- For low noise OTA circuit, recommend input pair to be surrounded by identical dummy device tied to ground or power

Signal name that need to be shield:

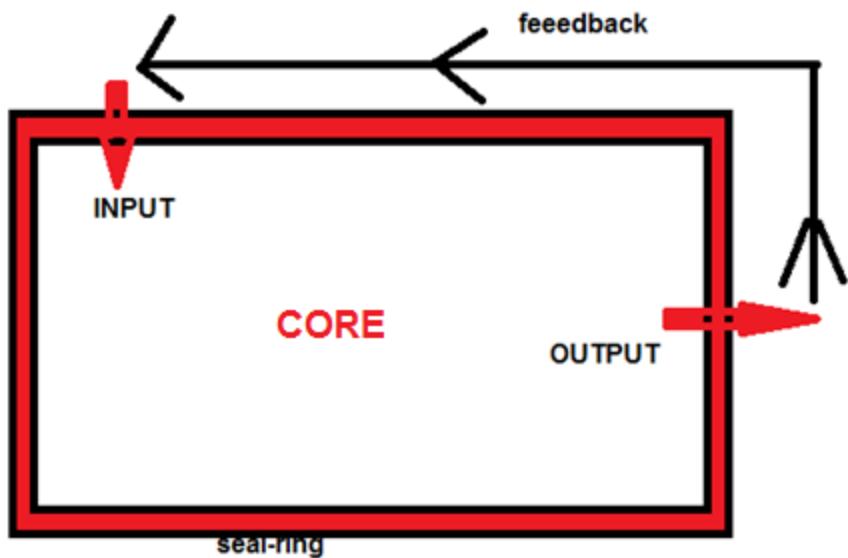
- CLK line (minimum cross-talk)
- DATA Line (minimum cross-talk)
- Reference Signal
- Bias / Current signal

Type of noise:

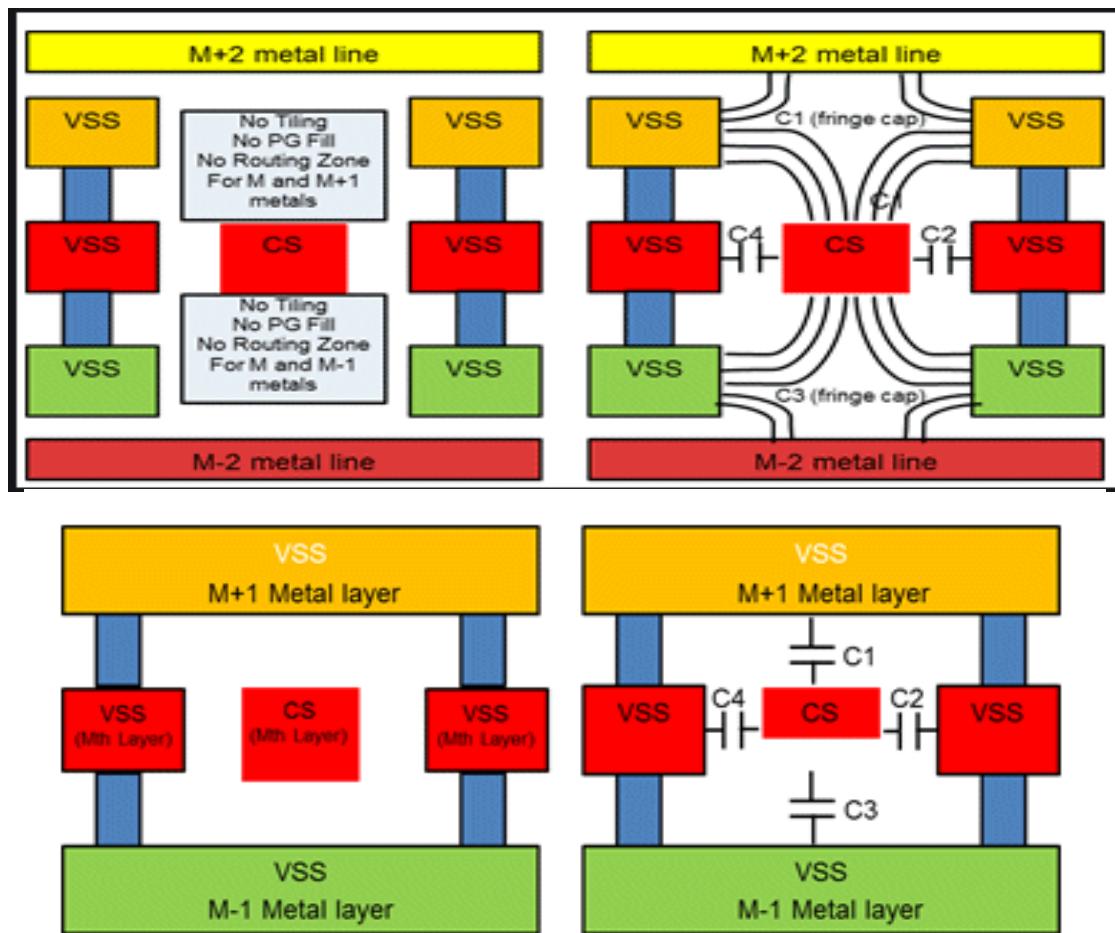
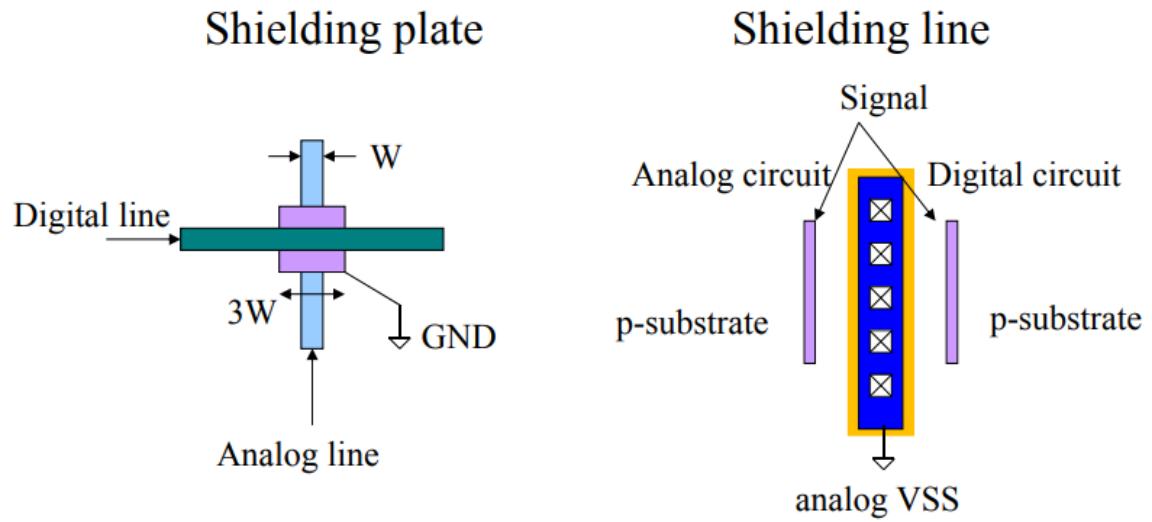
- **Inherent noise**
 - Noise resulting from the discrete and random movement of charge in a device
 - Thermal noise, Flicker noise, shot noise
 - The noise floor depends on the circuit design quality
- **Quantization noise**
 - Noise resulting from the finite digital word size
 - The SNR (signal-to-noise ratio) depends on the accuracy of ADC and DAC
- **Coupled noise (Crosstalk)**
 - Noise resulting from the signals adjacent circuits bleeding into each other
 - The noise immunity depends on a layout.

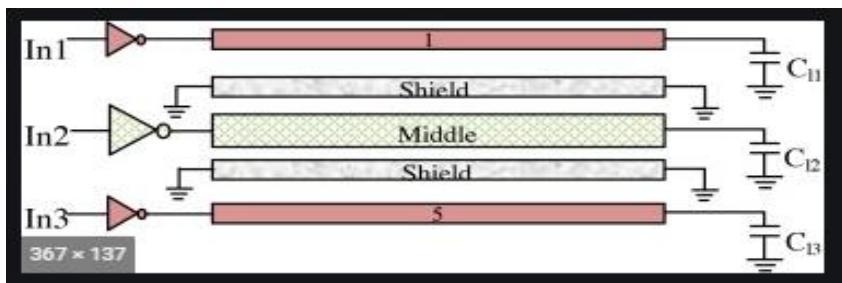
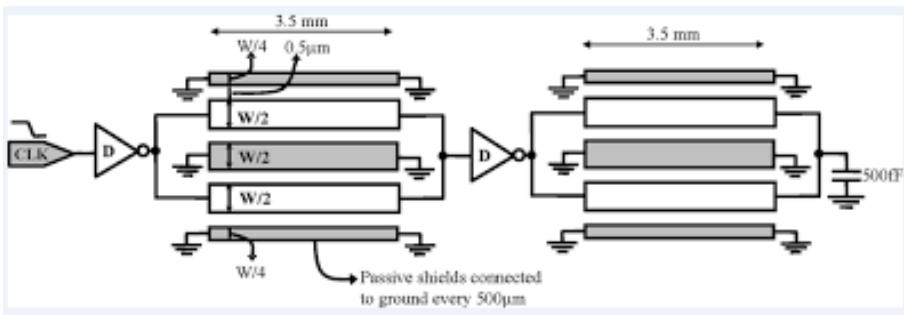
Noise in Signal:

- Keep high frequency signal in higher metal layer
- Use metal shield for victim line that is noise sensitive
- Use metal and poly shield for attacker line that travels through long distance
- Prevent feedback path through chip seal-ring between critical input and output.
Use additional guard ring to isolate the coupling .



Shielding of interconnects:





Bylateral Shielding:



Coaxial Shielding:

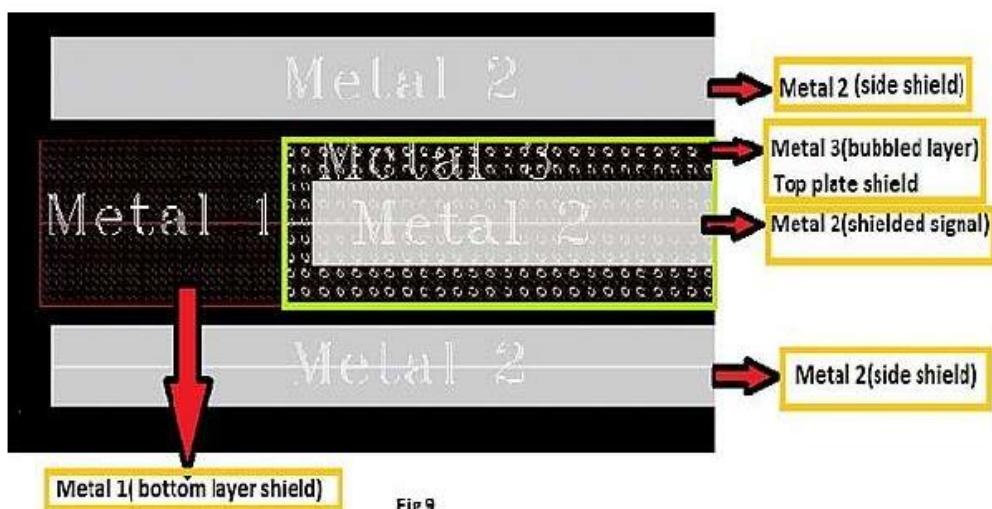
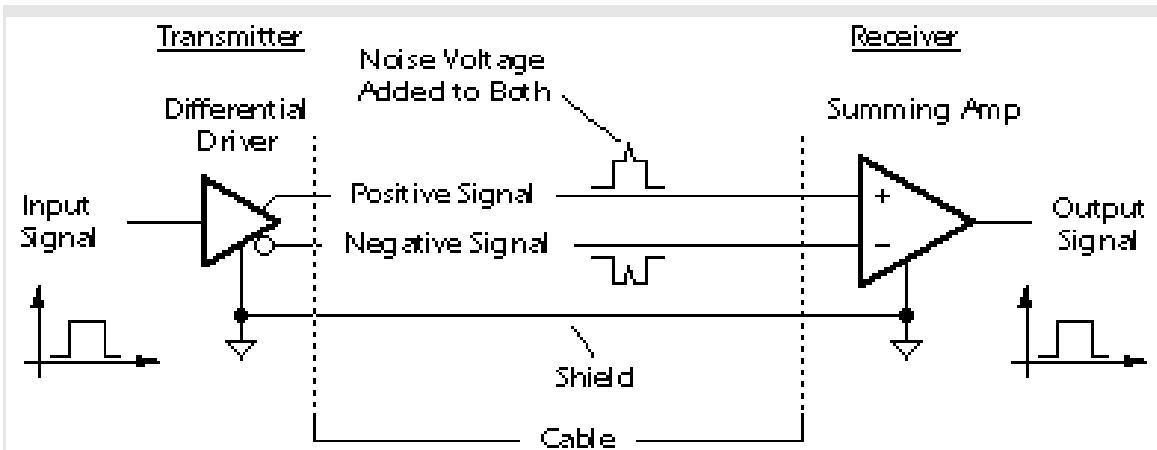
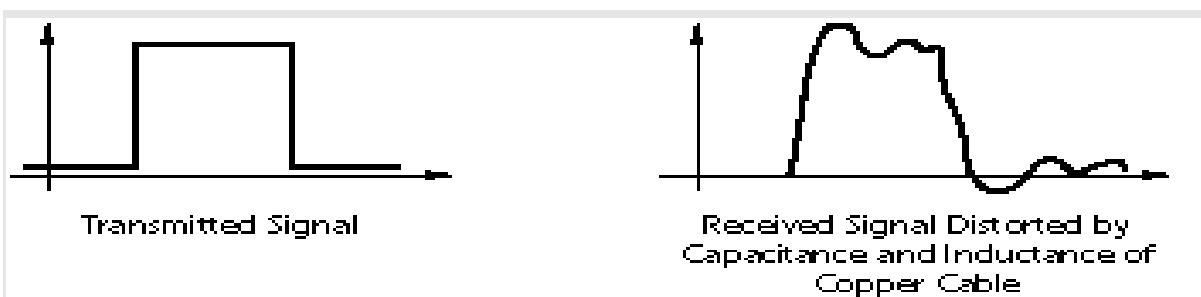
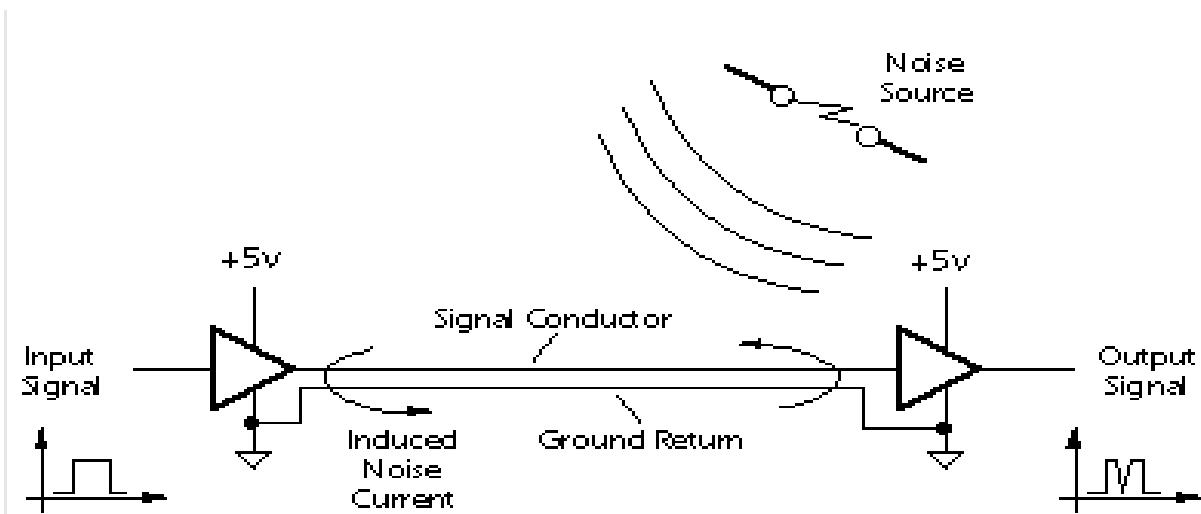


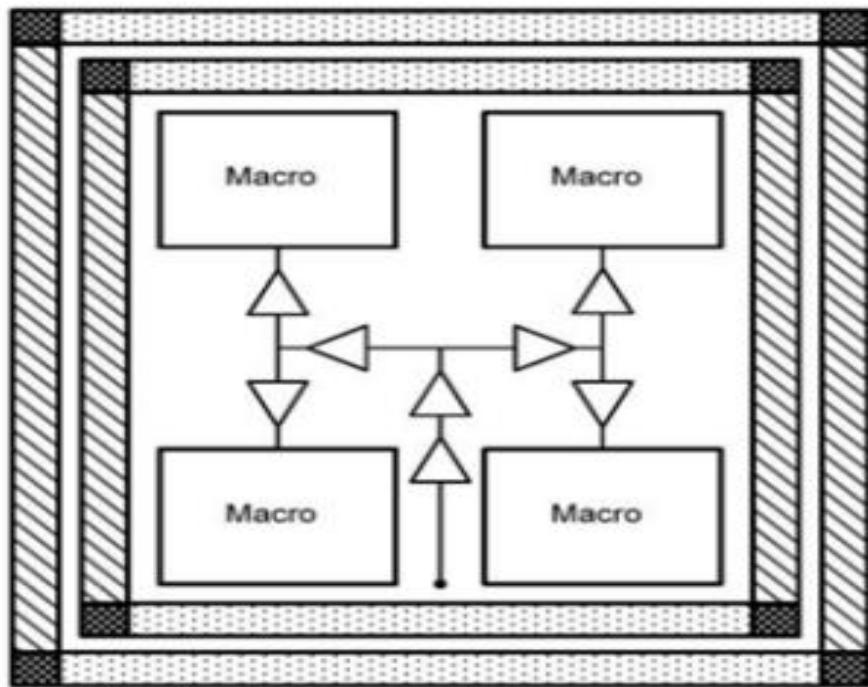
Fig 9

Buffers

- It Increases the fan-out capacity
- Inserting buffer decreases the transition time, which decreases the wire delay
- Due to buffer, Area and power consumption become increase.



Clock Tree connection by using buffers:



Double patterning (metal coloring)

Technology- advance node

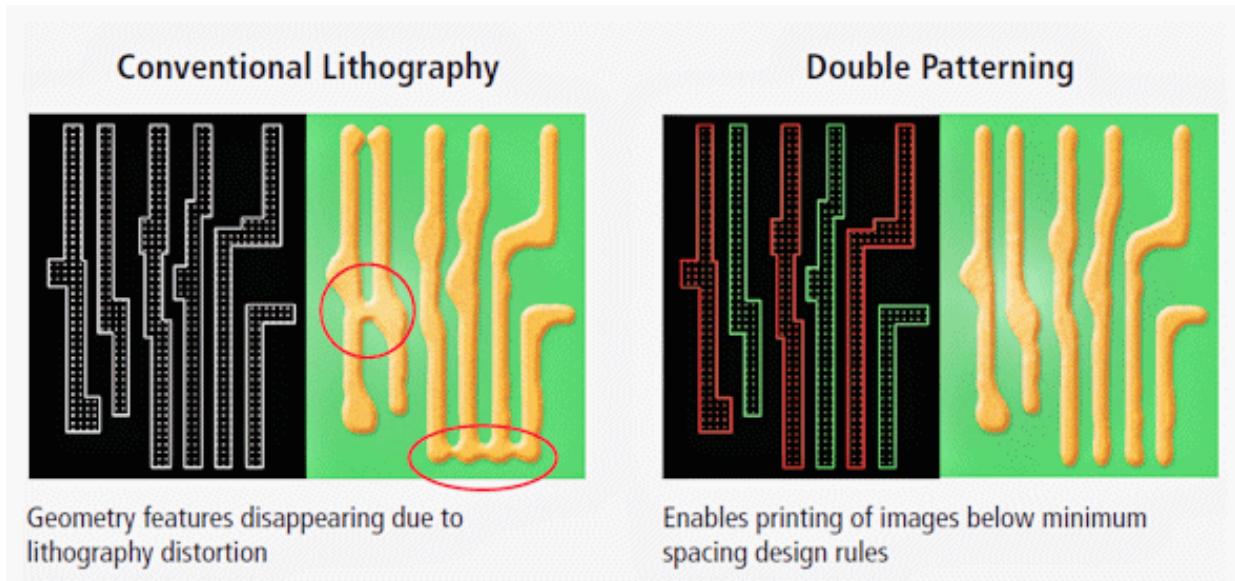
Multiple patterning (or **multi-patterning**) is a class of technologies for manufacturing integrated circuits (ICs), developed for photolithography to enhance the feature density. It is expected to be necessary for the 10 nm and 7 nm node semiconductor processes and beyond. The premise is that a single lithographic exposure may not be enough to provide sufficient resolution. Hence additional exposures would be needed, or else positioning patterns using etched feature sidewalls (using spacers) would be necessary.

Even with single exposure having sufficient resolution, extra masks have been implemented for better patterning quality such as TSMC at its 22nm node. Even for electron-beam lithography, single exposure appears insufficient at ~10 nm half-pitch, hence requiring double patterning.

How double patterning works:

In single-pattern lithography, a wafer is covered with a light-sensitive material, known as a photoresist. Light is then streamed through a patterned photomask (a template of the chip, essentially). The light strikes the photoresist and changes the chemical properties of the material. The wafer is then bathed in a chemical solution, which washes away the areas the light touched. This process is repeated multiple times, and the end result is (hopefully) a microprocessor.

When the silicon features become too small relative to the wavelength of light being used to etch them, however, the defect density skyrockets. Double patterning — using two photomasks, each with half of a pattern — can correct this, as shown below.



Double patterning :

- M1 metal require a set of 2 masks
- Masks are referred as E1/E2 (Exposure1, Exposure2)
- The process of generating the two masks is call decomposing or coloring

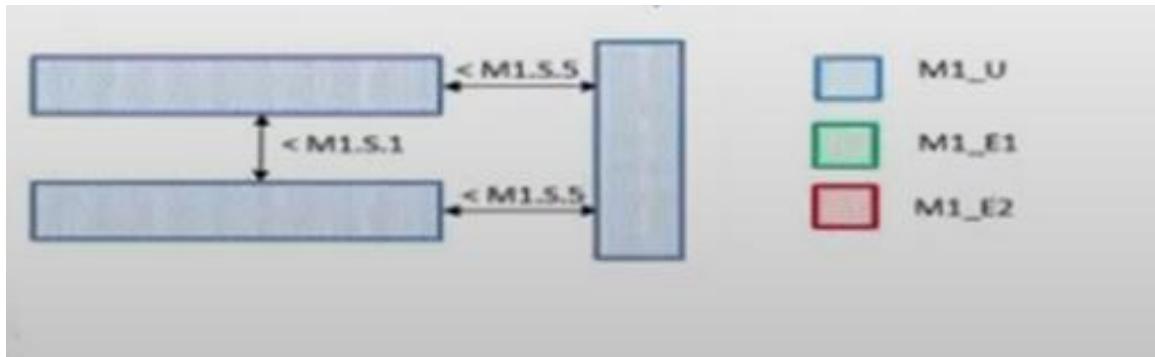


- The decomposing can be done manually or with the DRC deck
- Decomposition can be done at the cell or chip level
- The basic requirement to allow decomposition is to avoid odd cycles

What is an odd cycle?

- Is a set of 3 or more metal patterns which can't be decomposed because coloring would lead to spacing violation on one of the masks(E1 or E2).
- The solution is to increase separation in at least one of the space between the patterns

Conclusion: Layout should be DRC checked to be odd cycle clean.



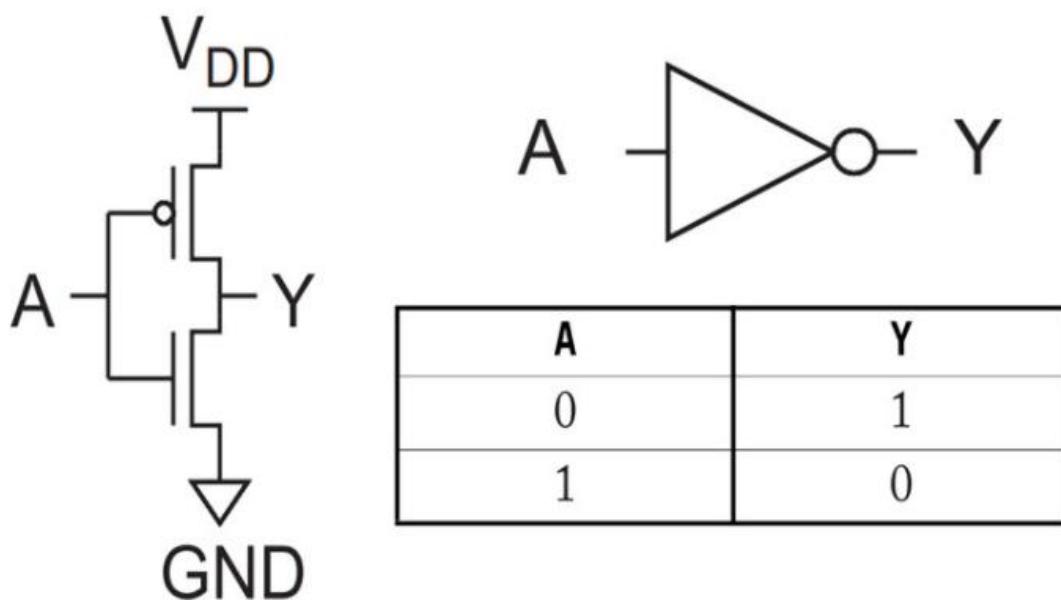
Video link :

1. <https://youtu.be/CYvPs3tyu3Y>
2. https://youtu.be/YDnIYmQ_Ll4

10.Logic cells

INVERTER

- The simplest of the logic gates is the Inverter. It is an essential part of digital design and understanding the operation and properties of an Inverter will make it significantly easier to study NAND Gates, Adders, Multiplexers and even Microprocessors.
- Following is the circuit of a CMOS Inverter Gate along with its symbols.



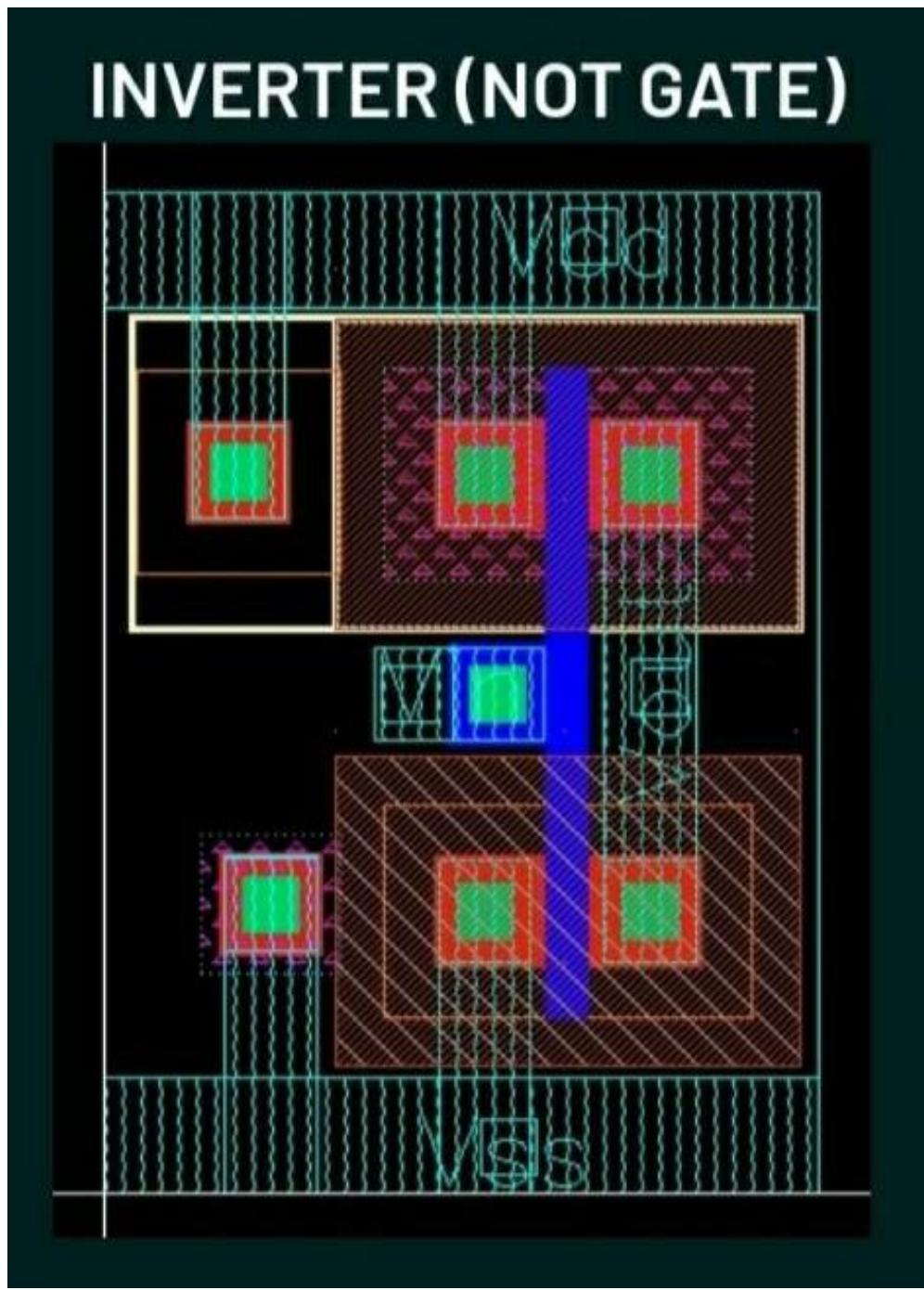
- CMOS Inverter is also known as the NOT Gate. From the above circuit, you can see that a CMOS Inverter consists of an N-channel MOSFET (NMOS) and a P-channel MOSFET (PMOS).
- When the input A is LOW i.e. Logic 0, the NMOS Transistor is OFF and the PMOS Transistor is ON. The P-channel MOSFET provides a path for the V_{DD} to appear at the Output. Hence, the output is HIGH i.e. Logic 1.
- Similarly, when the input is HIGH, NMOS is ON and PMOS is OFF. The output is connected to GND and the output is LOW.

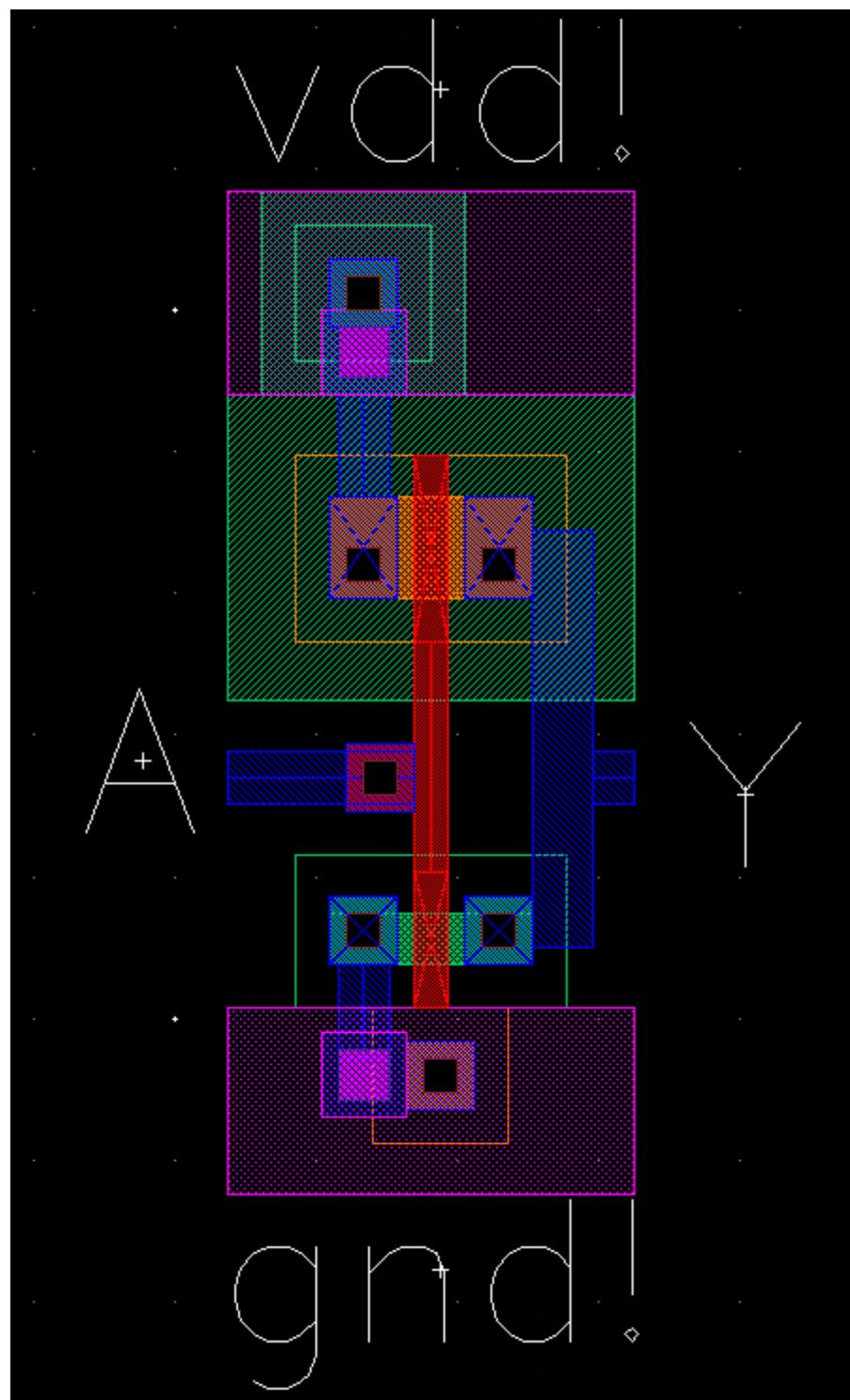
Layout view

Video link :<https://youtu.be/oSrUsM0hoPs>

3D view link: 1. Finfet inverter: <https://skfb.ly/oGOLU>

2 cmos inverter: <https://skfb.ly/oGOLC>

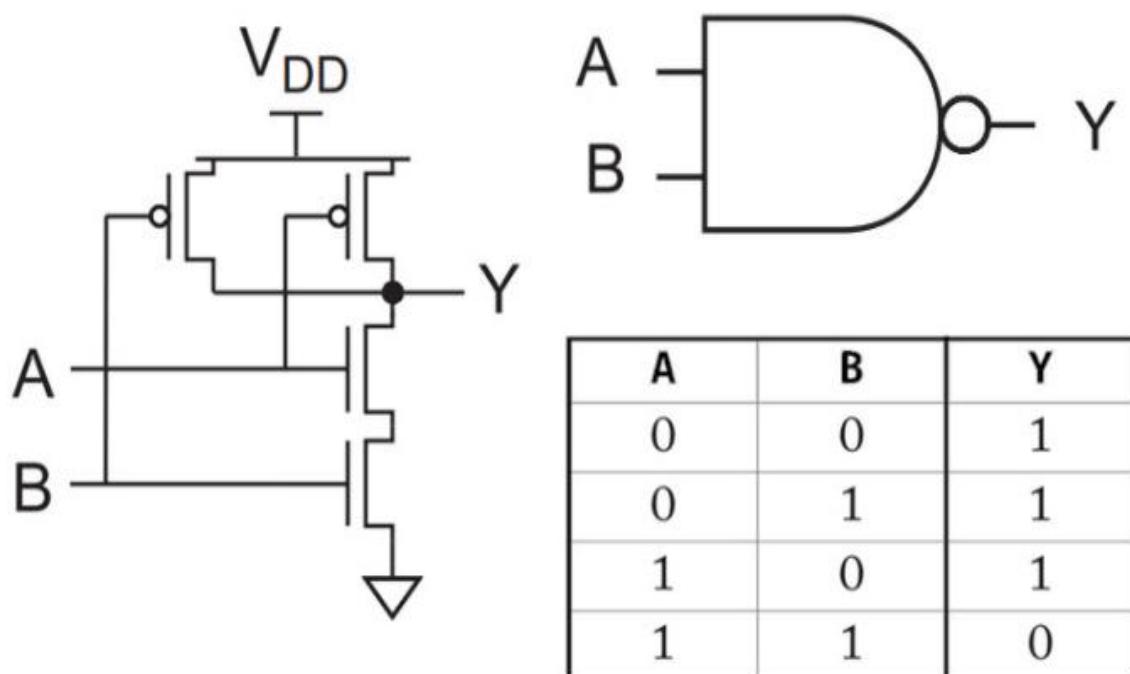




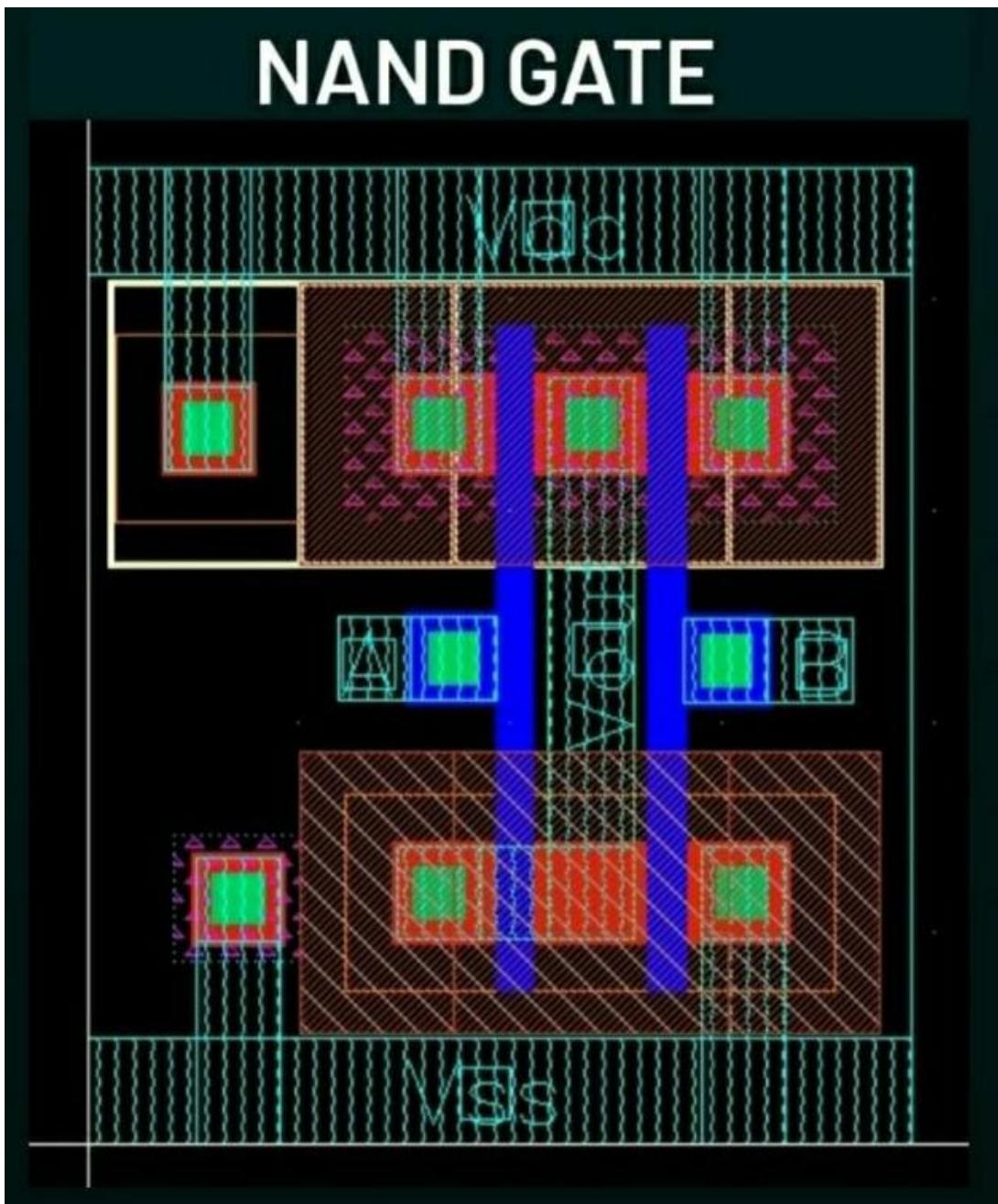
NAND

The following circuit shows a 2-input CMOS NAND Gate. As seen in the image, a 2-input NAND gate consists of two N-channel MOSFETs connected in series between output and GND and two P-channel MOSFETs connect in parallel between V_{DD} and output.

When any of the inputs A or B is LOW, at least one of the NMOS Transistors will be OFF. For the output to be LOW, both the inputs should be HIGH. For all the other combinations of the inputs, the output will be HIGH.



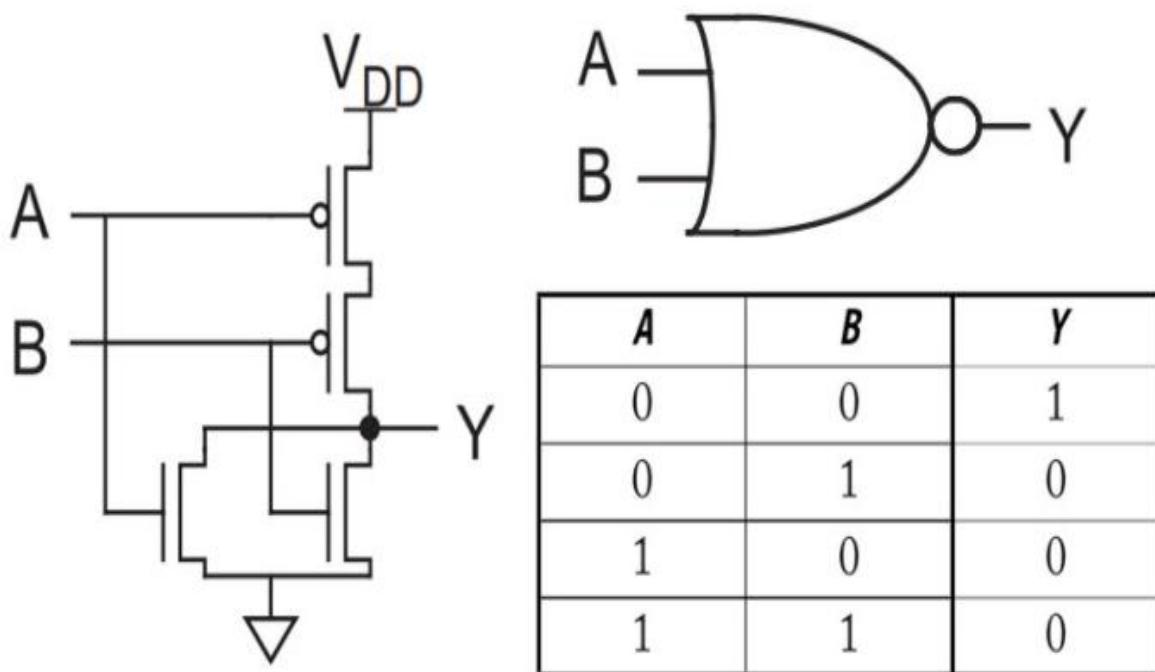
Layout view



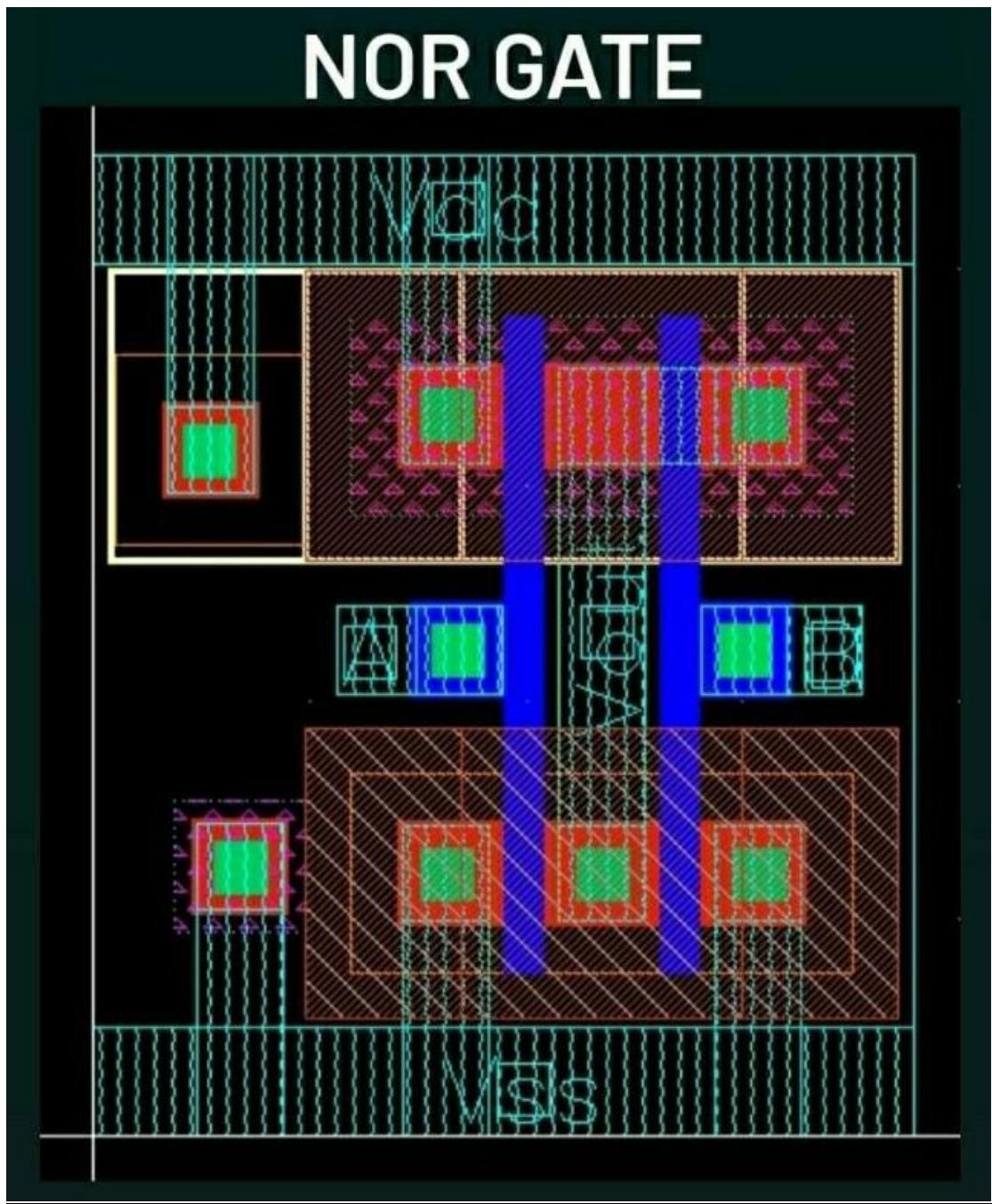
NOR

Circuit of a 2-input CMOS NOR Gate is shown below. It consists of two P-channel MOSFETs connected in series between V_{DD} and Output and two N-channel MOSFETs connected in parallel between Output and GND.

When either of the inputs A or B is HIGH, the output is LOW as at least one NMOS Transistor is ON. For the output to be HIGH, both the inputs must be LOW.



Layout view

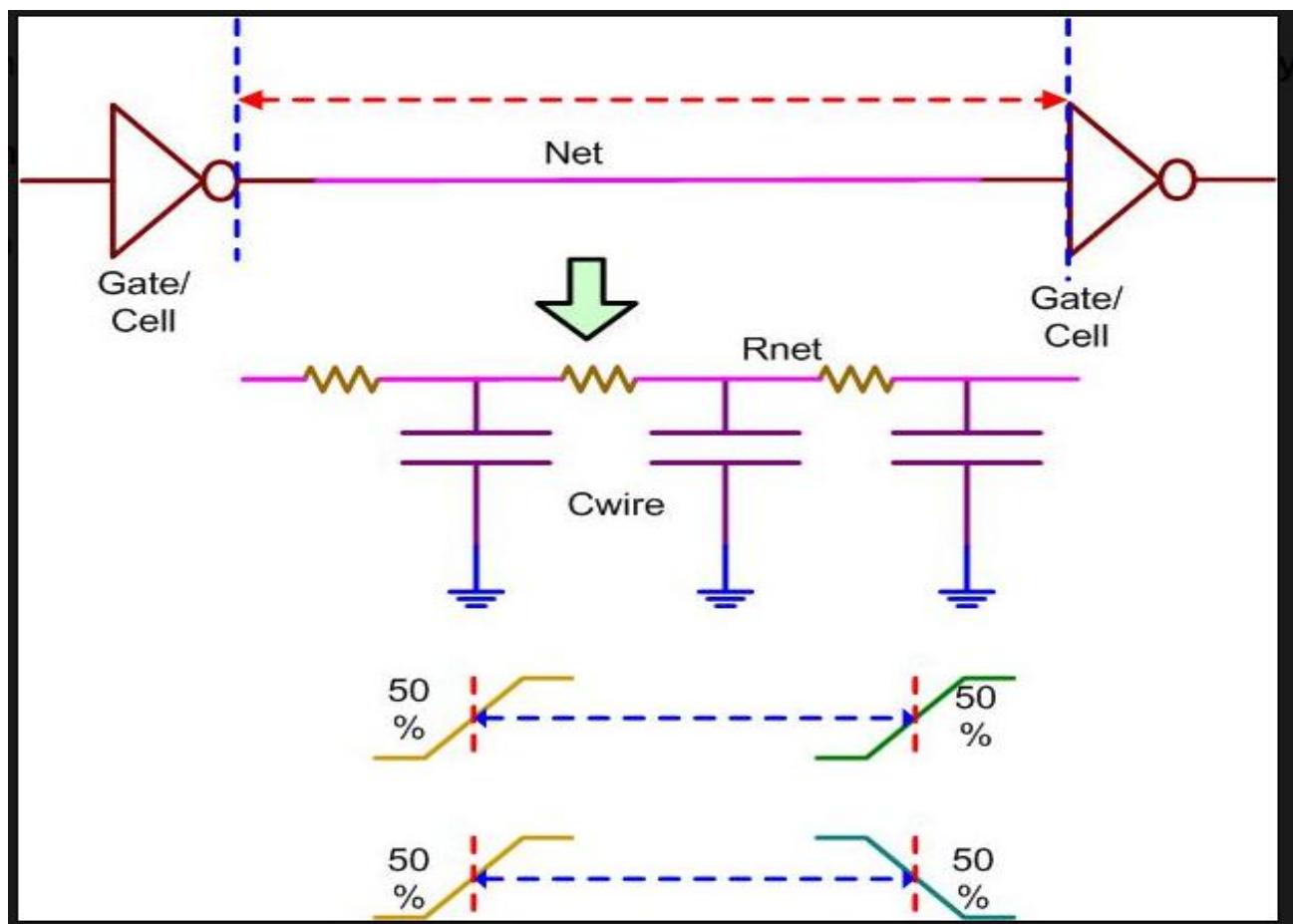
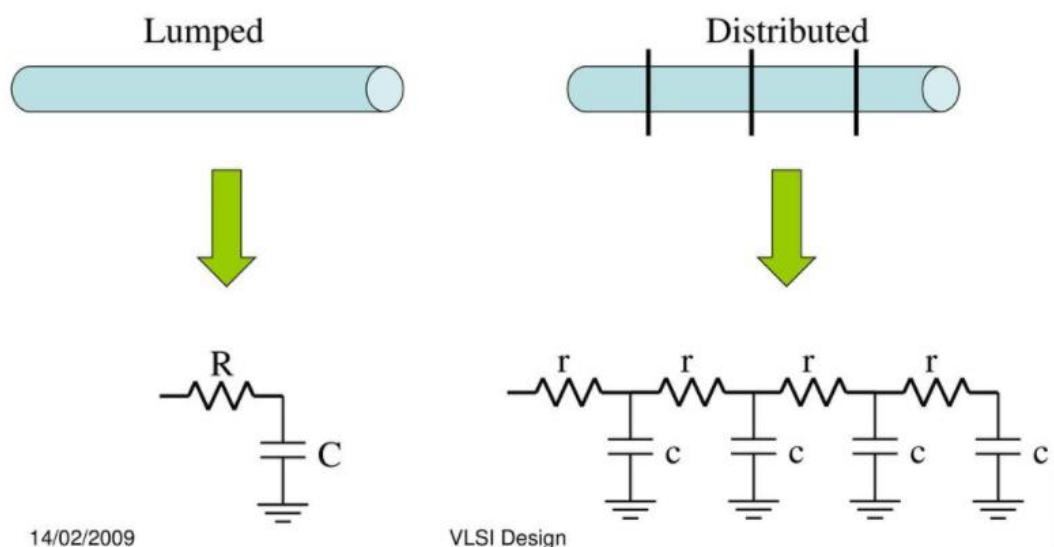


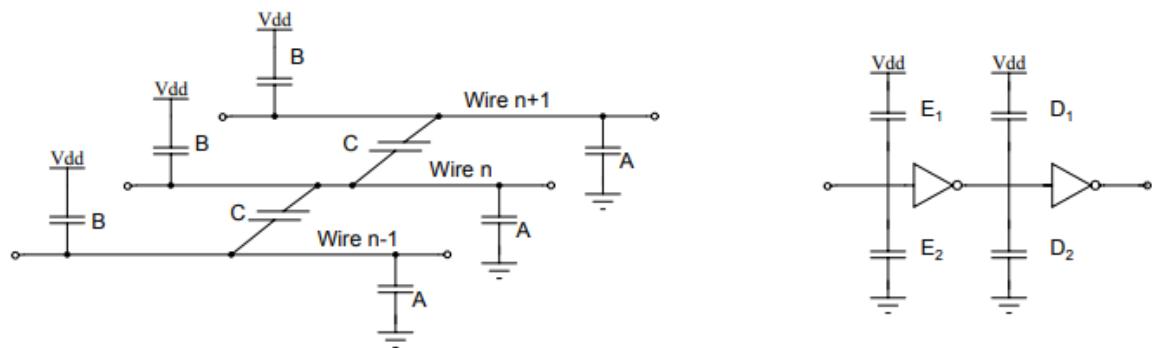
11.Parasitic capacitance and Resistance

- As the Semiconductor industry is growing so does the density of devices on chip. With the increasing density and decreasing spacing rules, the most significant effect that takes birth is parasitic. Parasitics can be of resistance or capacitance types, both have to be handled carefully.
- In VLSI applications the parasitic capacitance between signal lines can deplete our whole design. At low frequencies parasitic capacitance can usually be ignored, but in high frequency circuits it can be a major problem. For example, in amplifier circuits with extended frequency response, parasitic capacitance between the output and the input can act as a feedback path, causing the circuit to oscillate at high frequency. These unwanted oscillations are called *parasitic oscillations*.
- The parasitic capacitance arises from an electrical coupling between one signal line and another signal line or a signal line and the substrate. In some designs it becomes mandatory for us to reduce the parasitic capacitance of a particular net with respect to other signal

Solution to decrease parasitic capacitance

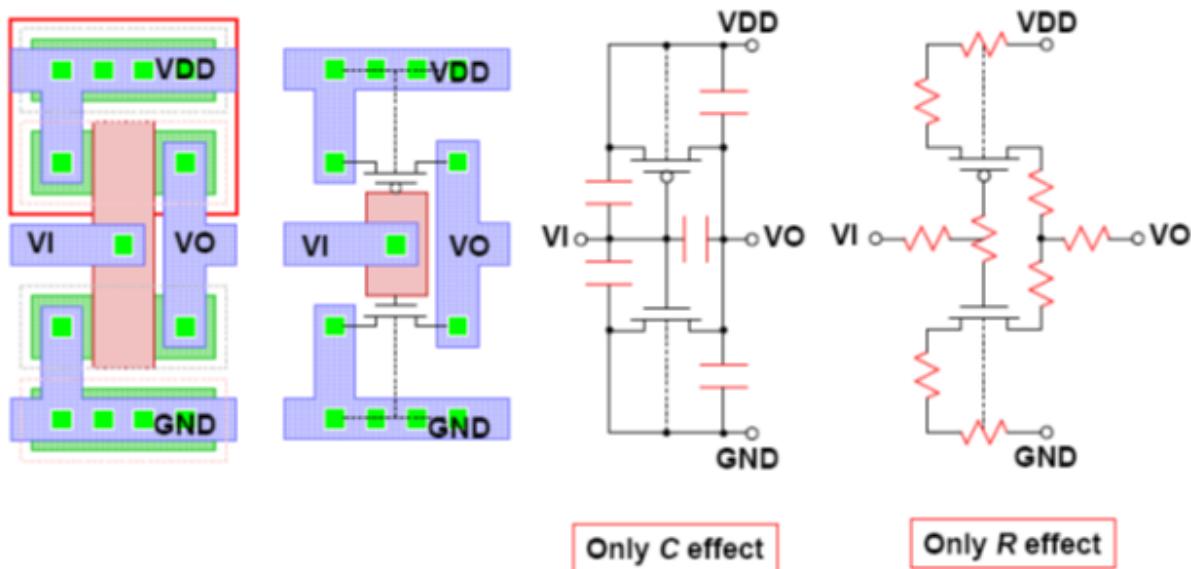
- Use higher metals for the net in which parasitic capacitance are important.
- Increase the spacing of all the nets from the net which is critical (for which parasitic capacitance is important).
- Put some other reference signal (with which parasitic capacitance is not so important) in between the nets for which lower parasitic capacitance required. This is shielding.
- Avoid too much parallel routing of metals.



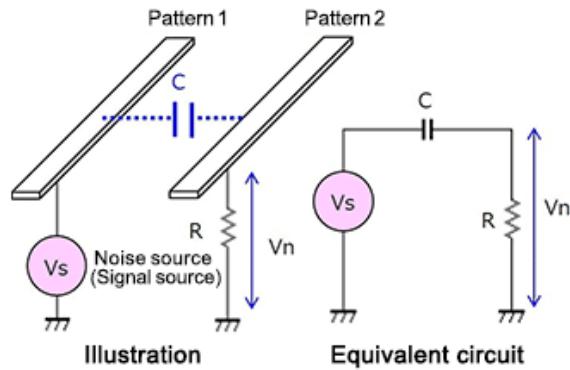


Parasitic Extraction of a MOSFET:

- Evaluate interconnection RC effects



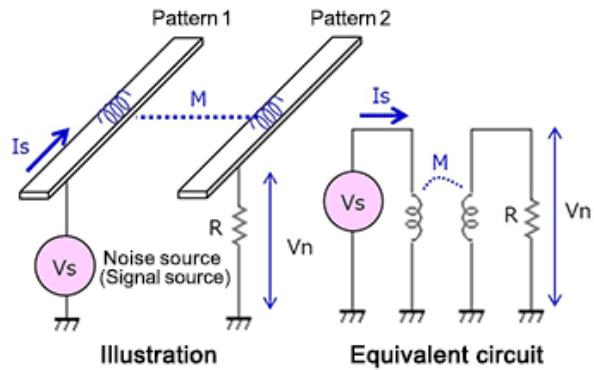
Capacitive Coupling (Electrostatic Coupling)



- Illustration of capacitive coupling between patterns due to stray capacitance
- Noise occurring on the side of pattern 1 causes a voltage V_n relative to GND due to capacitive coupling:

$$V_n = j\omega R C V_s$$

Inductive Coupling (Electromagnetic Coupling)



- Illustration of inductive coupling between patterns due to mutual inductance
- Noise occurring on the side of pattern 1 causes a voltage V_n relative to GND due to inductive coupling:

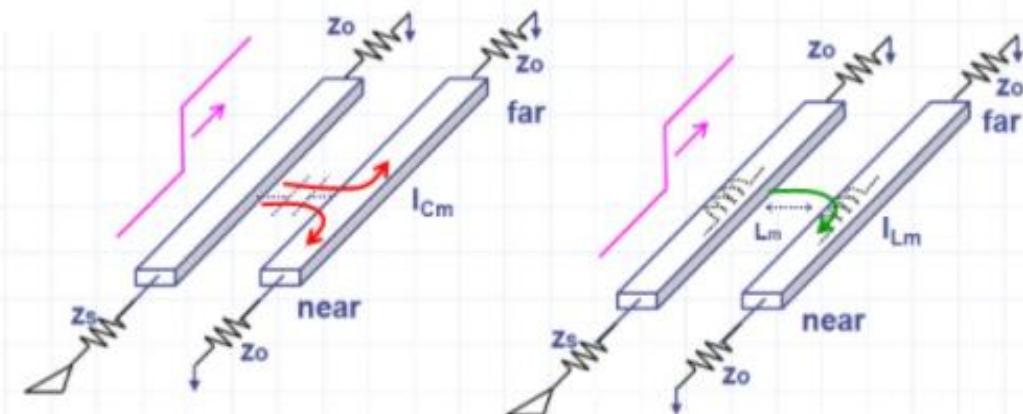
$$V_n = j\omega M I_s$$

Crosstalk induced noise:

The near and far end victim line currents sum to produce the near and far end crosstalk noise.

Coupled currents:

- Current induced by capacitive coupling goes to both directions
- Current induced by inductive coupling goes opposite to the drive current



$$I_{near} = I_{Cm} + I_{Lm}$$

$$I_{far} = I_{Cm} - I_{Lm}$$

12.DRC AND LVS

Design Rule Check (DRC):

- Determines whether the layout satisfies the MOSIS design rules
- Checks for widths, separations and overlaps
- Differ with the process technology used
- DRC is computationally intensive – used on parts of ASIC separately
- Some commercial tools available for DRC are
 1. Caliber by Mentor Graphics
 2. PVS and Assura by Cadence
 3. Hercules by Synopsys

Layout Vs Schematic (LVS):

- LVS tests for the functionality of the layout – created layout is functionally same as schematic
- Performs 3 important steps
 1. Extraction – Determines the semiconductor components represents in the drawing
 2. Reduction – Connects them in parallel or series to complete the circuit
 3. Comparison – Compares layout netlist with schematic netlist to test the functionality
- Reported Errors are
 1. Shorts and Opens
 2. Components Mismatch and Missing
 3. Parameters Mismatch

Video link:

- 1. <https://youtu.be/4nhEYA1rf4w>**
- 2. <https://youtu.be/HnsuU4SIDAQ>**

13.SXCUT AND PSUB LAYER

- SXCUT is the name used by GLOBALFOUNDRIES and Samsung for a CAD layer used to define two or more region of substrate (ground/vss) for LVS purposes that are electrically connected but has different name in schematic, and PSUB used in the TSMC.
- The SXCUT/PSUB layer is also a useful layer for separating out analog and digital grounds (e.g : VSSA and VSSD) **for LVS purposes only**
- SXCUT/PSUB must be used very carefully and only when it is absolutely necessary

SXCUT/ PSUB usage guidelines

- If the SXCUT layer must be used, please make sure the layout does not have any SXCUT layer overlapping the DNW layer (DNW,S3,T3), as this may hide some real shorts in the layout during LVS
- Overlap of SXCUT and DNW layer are usually flagged in ERC checkes, but in some cases, it may not be flagged, If SXCUT layer is used in the design, it is strongly recommended that the LVS methodology described in the next slide is done before tapeout
- Please note that ERC is not on by default, but it is STRONGLY recommended that ERC is run for all tapeouts

SXCUT/PSUB usage – Recommended additional LVS check

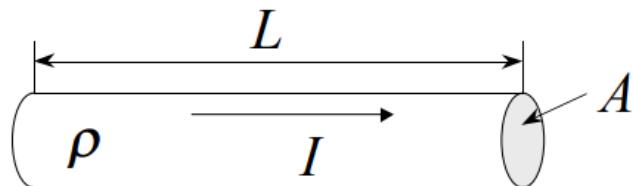
If the SXCUT layer is used in a design, it is recommended that the following additional top level LVS check is done

- On the schematic side, instantiate the top level symbol in a schematic and short together all the pins connected to the P-substrate (e.g:VSSA and VSSD etc)

Two Ground isolation video link:<https://youtu.be/IGSxLyx7NSc>

14.SHEET RESISTANCE

Resistance of a metal line

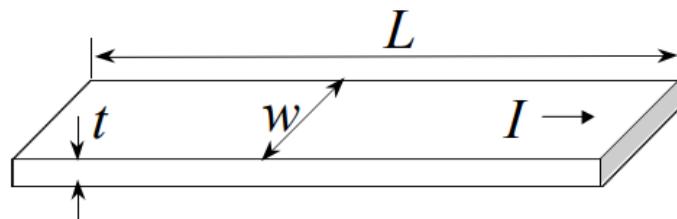


$$R = \rho \frac{L}{A}$$

R = Resistance, ρ = Resistivity

L = Length, A = Area of line cross-section

Sheet Resistance Concepts



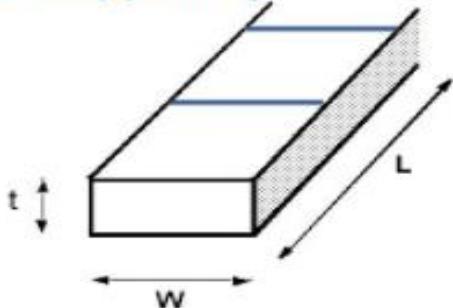
Apply current I and measure voltage V ,

Resistance: $R = V/I = \rho L/(wt)$

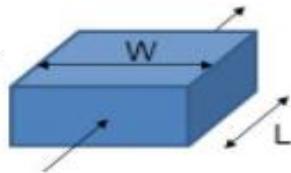
For a square sheet, $L = w$, so $R = \rho/t = R_s$

Unit of R_s : ohms per square (Ω/\square)

Just count number of squares along wire and multiply with R_s



All wires in a layer have the same thickness t



$$R = \frac{\rho L}{t W}$$

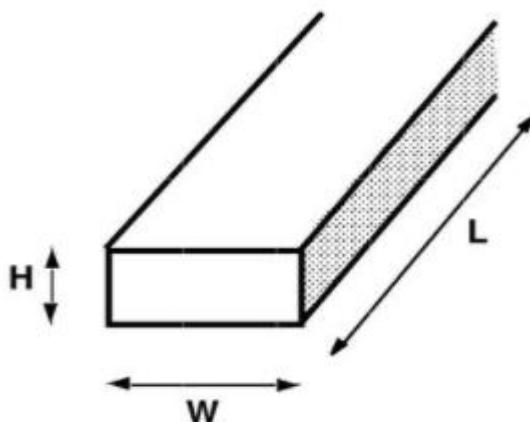
Sheet Resistance

$$R_0 = \rho / t$$

$$R_1 \equiv R_2$$

When $W=L$ the resistance is equal to R_0 , the sheet resistance (i.e. the resistance of a square wire)

Other books use R_s



$$R = \frac{\rho L}{H W}$$

Sheet Resistance
 R_0

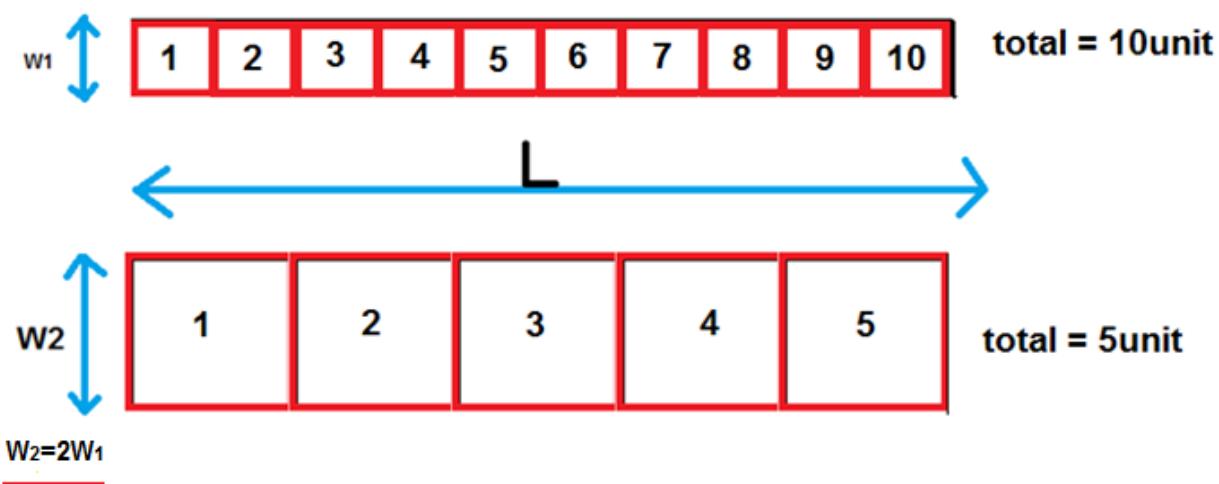
$$R_1 \equiv R_2$$



How to calculate Resistance:

Total Resistance = (resitivitiy) x (total unit)

- Resistance constance value will be given in the Technology Document,
- $R = \rho L/tW$, ρ/t value will be given
- Try to make $L/W = 1$ or make a squer unit
- Total number of squer unit = total length / width
- Total Resistance of wire = (Constance value) X (total number of squer unit)



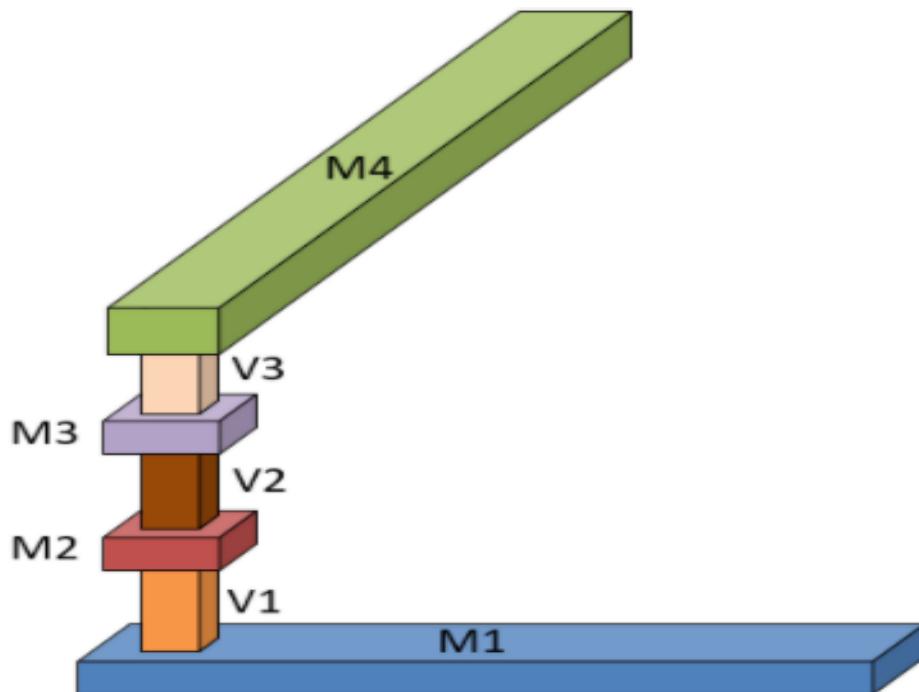
Video link : <https://youtu.be/OoOayYM-WRE>

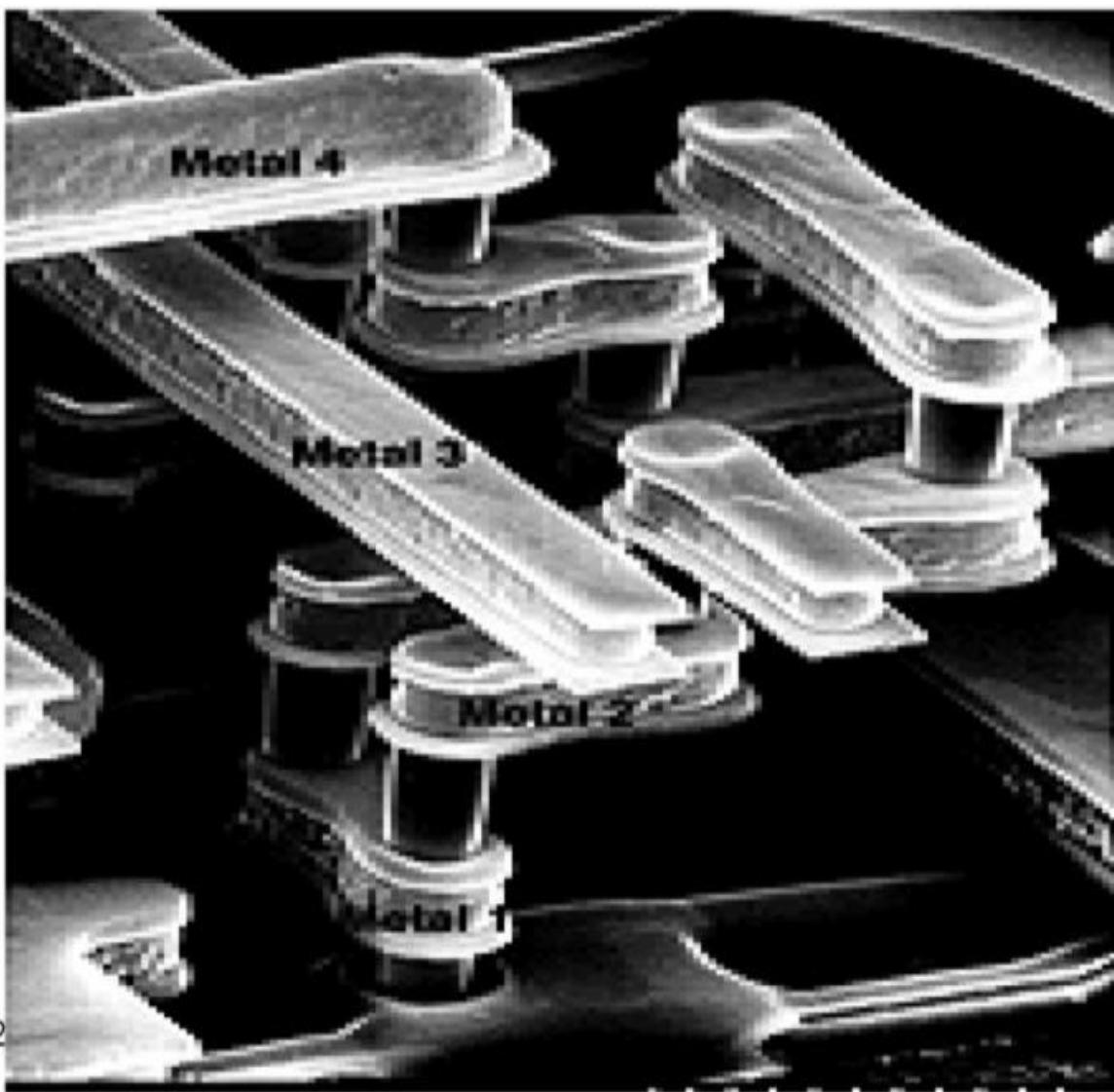
- Via Resistance :

- Vias made from Tungsten in Aluminum processes
- Vias made from Copper in Copper processes
- Via resistance significance in TSMC 180nm/0.18um technology

Diffusion –M1	11.0 Ω
Poly – M1	10.4 Ω
M2 – M1	04.5 Ω
M3 – M1	09.5 Ω
M4 – M1	15.0 Ω
M5 – M1	19.6 Ω
M6 – M1	21.8 Ω

- Use Multiple vias in parallel to reduce effective contact resistance
- Copper processes have lower via resistance





14/02

15.ELECTROMIGRATION (EM)

- Electromigration (EM) is a significant problem in integrated circuits and can seriously damage interconnect wires and vias, reducing the circuit's lifetime
- Due to high current density, the electrons in the metal moved with high acceleration. And these electrons transfer their momentum to other atoms and the atoms get displaced from their original position and might create voids and hillocks.
- Voids will create opens and Hillocks will create shorts between metal layers.

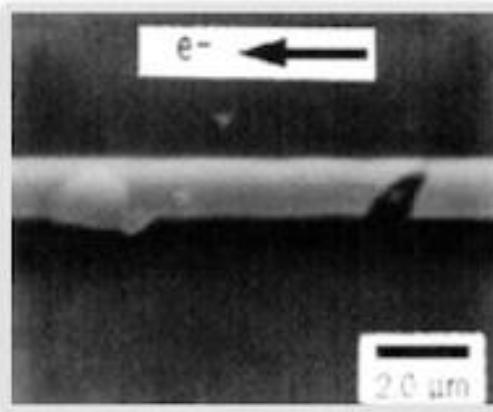


Fig (a): Void (Open)

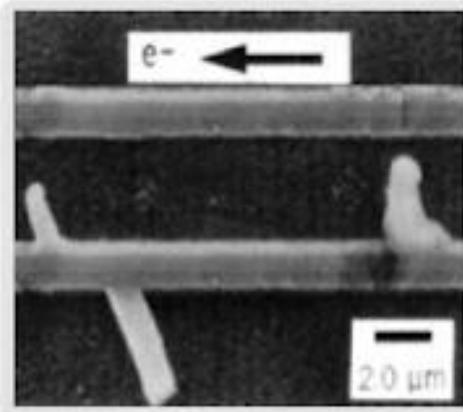


Fig (b): Hillock (Short)

Solution for EM:

- Increase the width of the wire
- Insertion of Buffers
- Upsizing the driver cell
- Switch the net into higher metal layer
- Use metal stack
- Use parallel routing

16.IRDROP

- The power supply in the chip is distributed uniformly through metal layers (Vdd and Vss) across the design. These metal layers have finite amount of resistance.
- When voltage is applied to this metal wires current start flowing through the metal layers and some voltage is dropped due to that resistance of metal wires and current. This Drop is called as IR Drop.
- For example: a design needs to operate at 2 volts and has a tolerance of 0.4 volts on either side, we need to ensure that the voltage across its power pin (Vdd) and ground pin (Vss) in that design does not fall short of 1.6 Volts. The acceptable IR drop in this context is 0.4 volts. That means the design in this context can allow up-to 0.4 volts drop which does not effects the timing and functionality of design

❖ How it effects the timing :

IR Drop is Signal Integrity effect caused by wire resistance and current drawn off from Power (Vdd) and Ground (Vss) grid. According to Ohms law, $V=IR$. If wire resistance is too high or the current passing through the metal layer is larger than the predicted, an unacceptable voltage drop may occur. Due to this un acceptable voltage drop, The power supply voltage decreases. That means the required power across the design is not reaching to the cells. This results in increased noise susceptibility and poor performance.

❖ How to solve IR problems?

- Try to reduce the metal Resistance
- Reduce the metal length by changing placement
- Increase the width of the wire
- Switch the net into higher metal layer
- Use Metal stacking

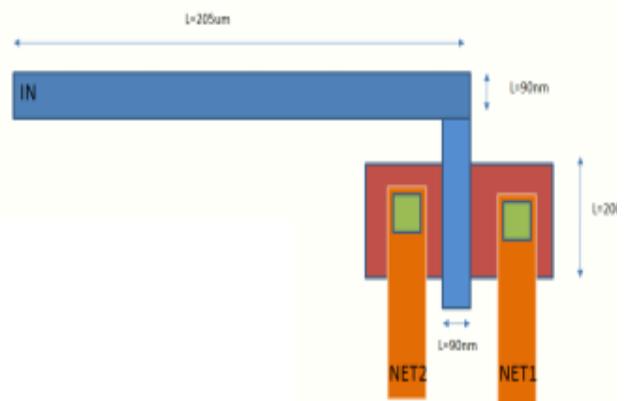
Video link for EM and IR drop: <https://youtu.be/8sZuwzf5EgE>

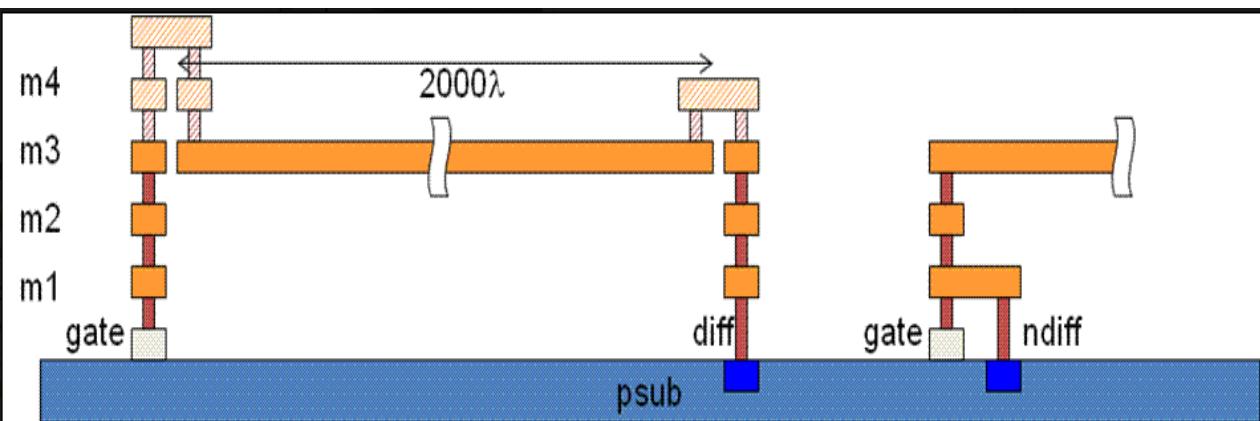
17. ANTENNA ERROR

- During chip processing, interconnect components are subject to plasma etching. If the amount of charge collected on these conductor lines is sufficient, it may damage or destroy the gate of transistor
- Antenna effect occurs only gate terminal because of present of oxide layer which form capacitance.

Antenna basics :

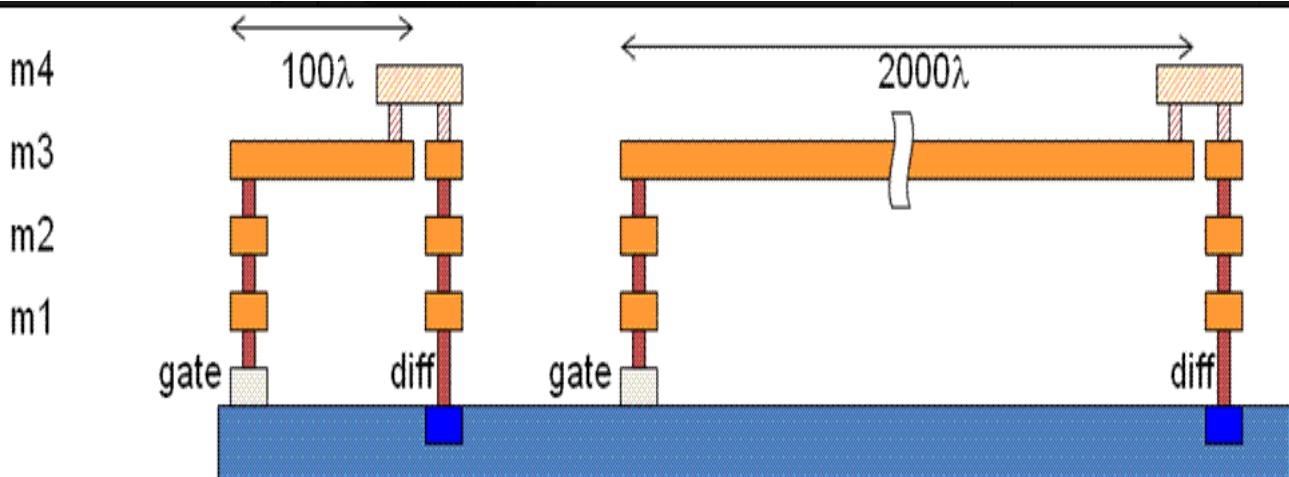
- Mechanism -
 - ✓ Damage is induced in the gate oxide due to plasma processing steps including poly, inter-layer dielectric, contact, via and metal processing
 - ✓ Due to lack of S/D diode to substrate (SOI), Source/Drains are also susceptible
 - ✓ Antenna protection is needed on all device terminals
- Design rule (refer to design manual antenna design section)
 - ✓ A maximum ratio is defined for metal area (or via) to the poly gate area
- Design impact
 - ✓ Design time
 - ✓ Adding diodes impacts parasitic and linearity.





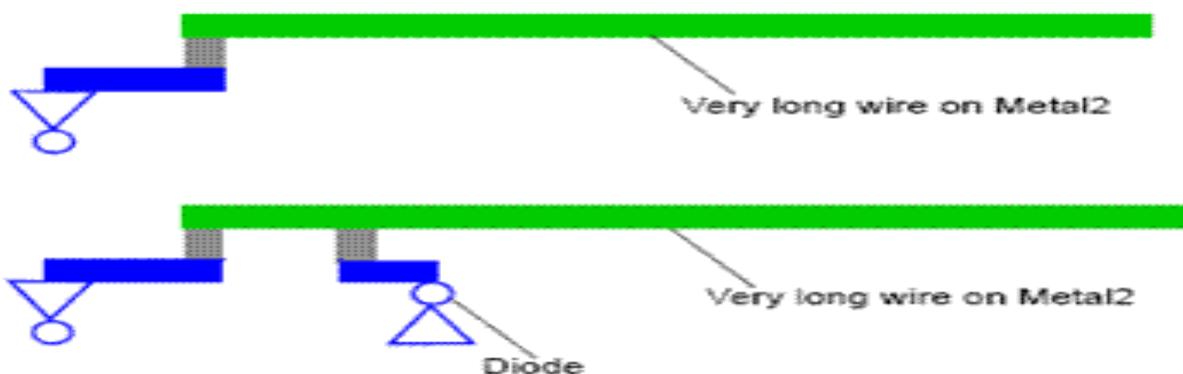
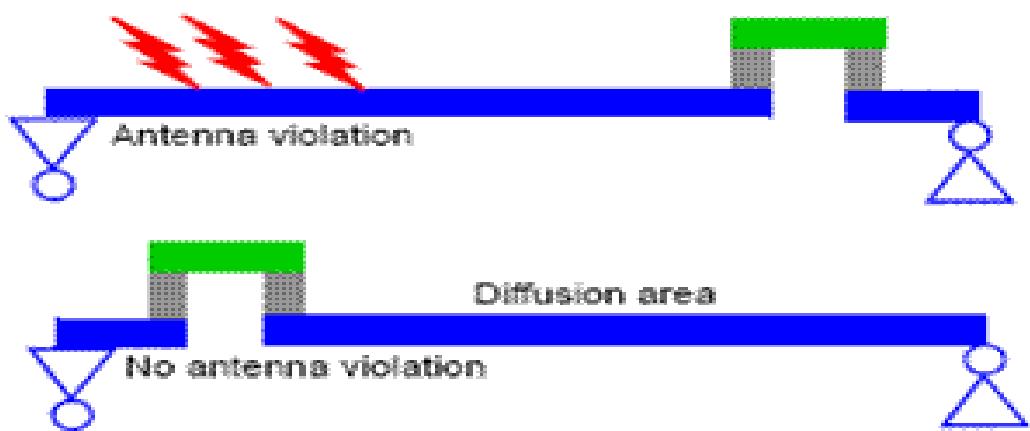
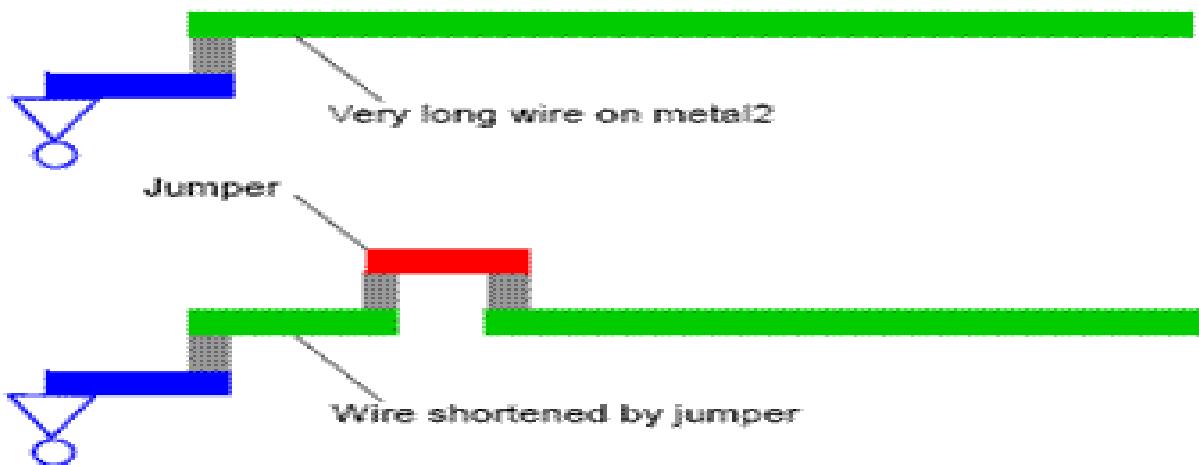
Bridging keeps gate away from long metals until they drain through the diffusion

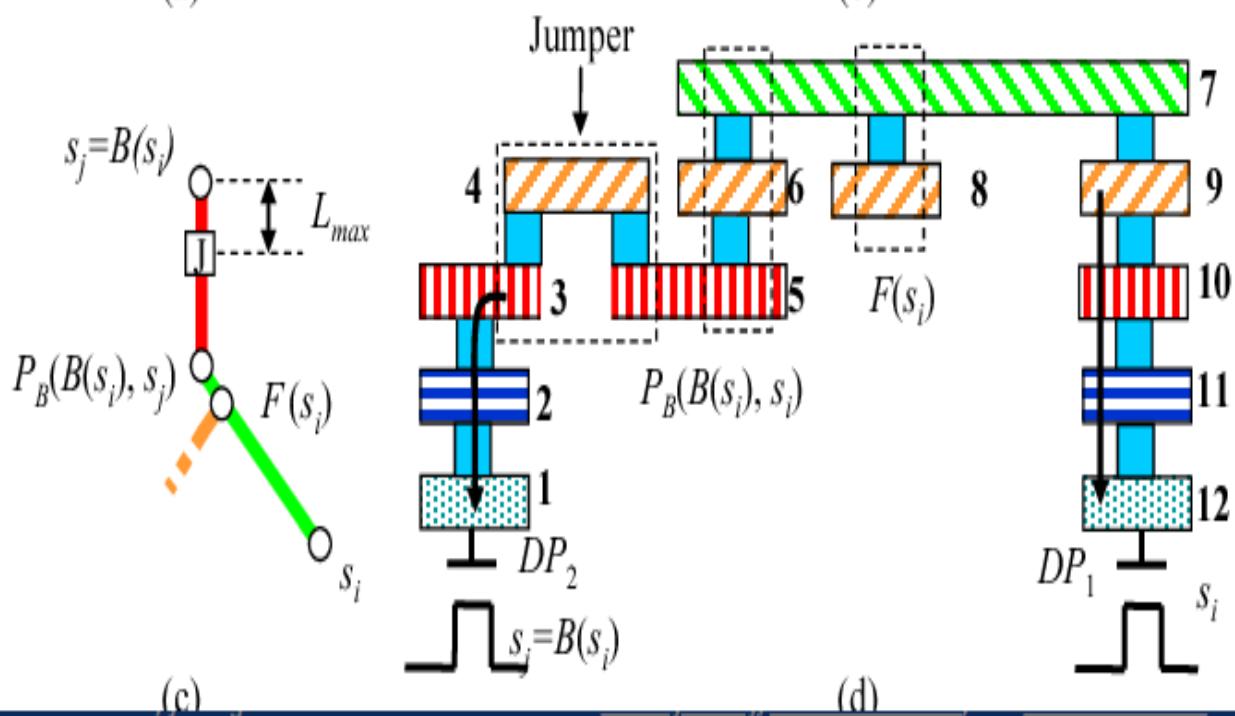
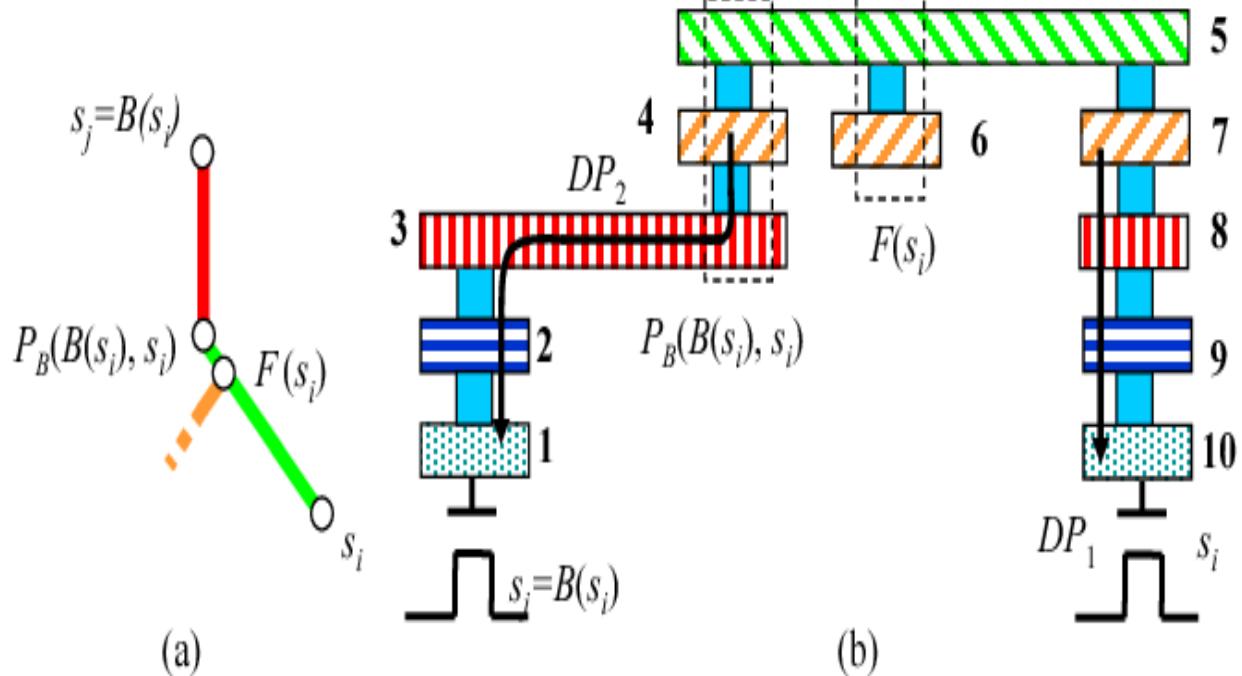
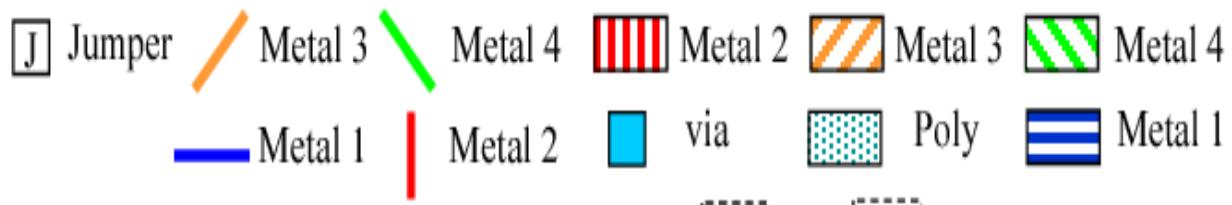
Node diodes are inactive during chip operation (reverse-biased p/n); let charge leak away harmlessly



Safe: m3 is too short to accumulate very much charge; won't kill gate

Dangerous: lots of m3; will probably accumulate lots of charge and then blow oxide





❖ How to solve antenna effect ?

- a. Reduce metal (or via) area until the maximum ration is met
- b. Higher metal jumper to cut the large area lower layer metal (via) to the gate
- c. Use antenna diode (reverse biased) to add more discharge path (area)
- d. Well tie or substrate tie are also counted as “antenna diode” area
- e. Most of the time, ESD diodes help fix the antenna violations at I/O
- f. Antenna violations on nets with negative voltage swing need to be fixed using back-to-back diode

• Metal Jumpers :

Charges accumulated with each etching process gets removed during Planarization process so jumpers help transfer problem to next level

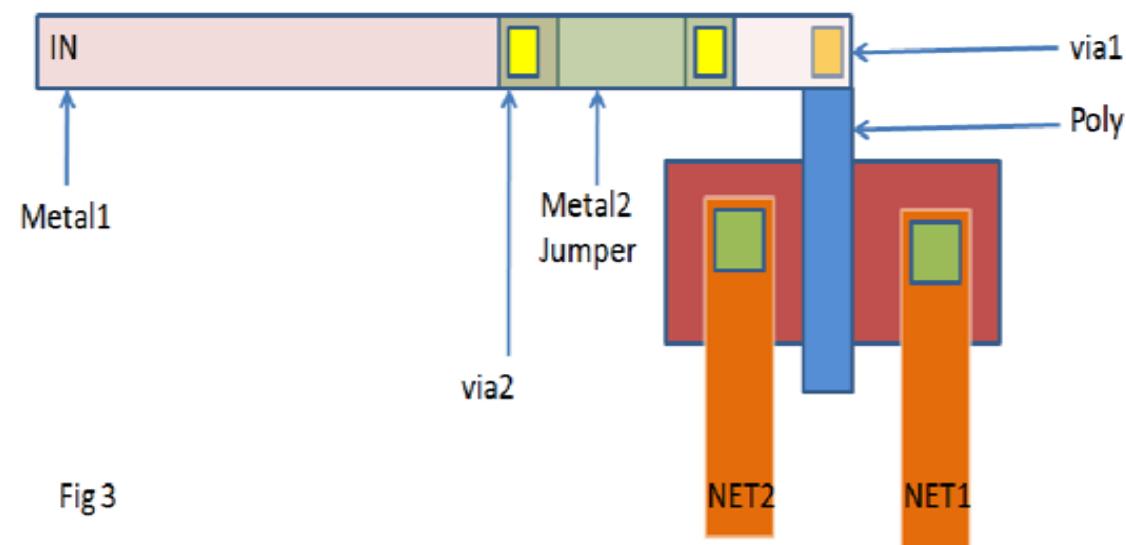
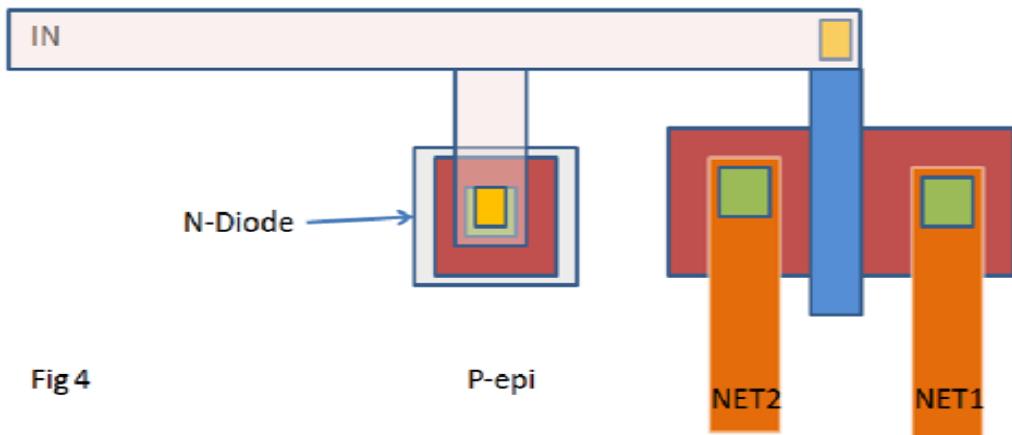


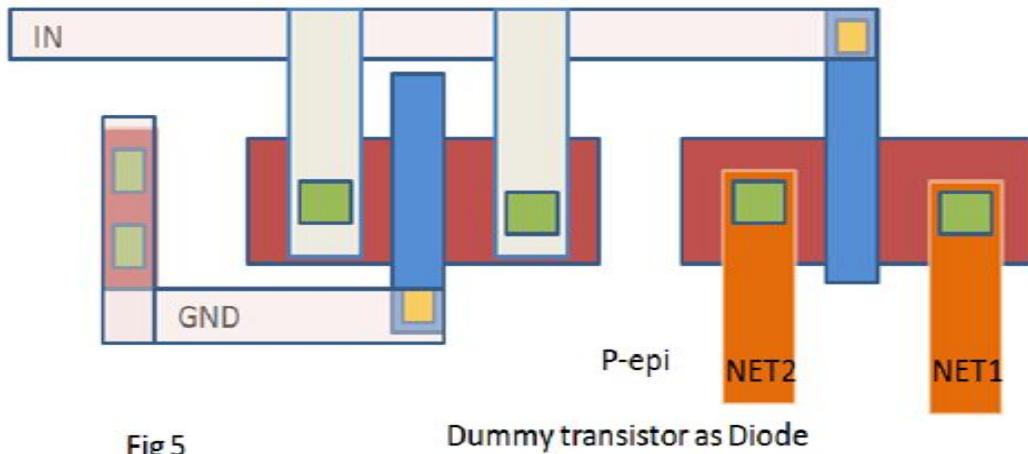
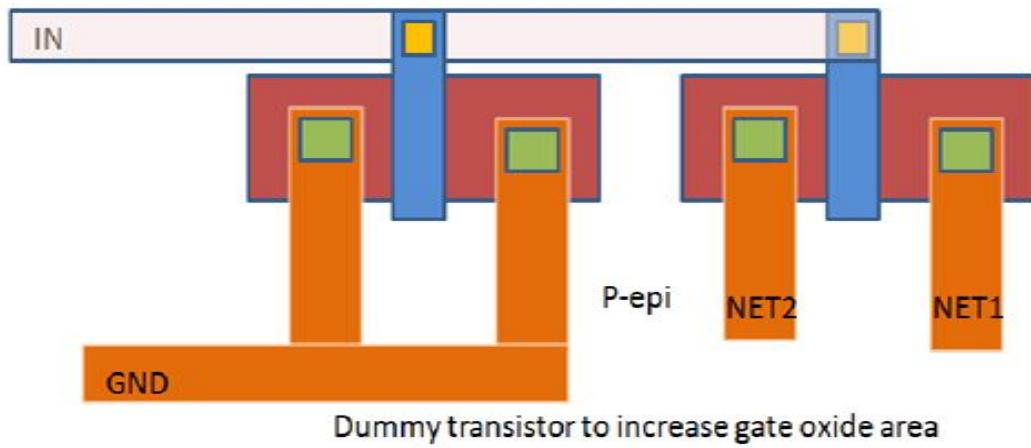
Fig 3

A diode helps dissipate charges accumulated on metals. An antenna diode should be placed as near as possible to gate of device on lower level of metal



- **Adding Dummies :**

- A dummy device can be used to increase the gate area. It helps when ratio is little more than optimum.
- For large ratio, it would lead to the adding lot of dummies hence not preferred.
- A dummy device can be used to form a diode and used. This process takes lesser area than adding conventional antenna diode.
- This process has less complex DRC as compare to DRC in real antenna diode



Video link:

1. <https://youtu.be/WyfiQwtRCC0>

2. <https://youtu.be/T8PqxIDQ5JE>

18.LATCH-UP

SCR:

- The Forward Conduction Mode is the only mode at which the SCR will be in the ON state and will be conducting. We can make the SCR conduct in two different ways, one we can increase the applied **forward bias voltage beyond the breakdown voltage** or else we can apply a positive voltage to the gate terminal.
- The IC has very low voltage, so the forward bias voltage it's not happen. But If the positive voltage applied at the gate of SCR will help the SCR to move to the conduction state. During this mode of operation, the SCR will be operating in forward bias and current will be flowing through it.

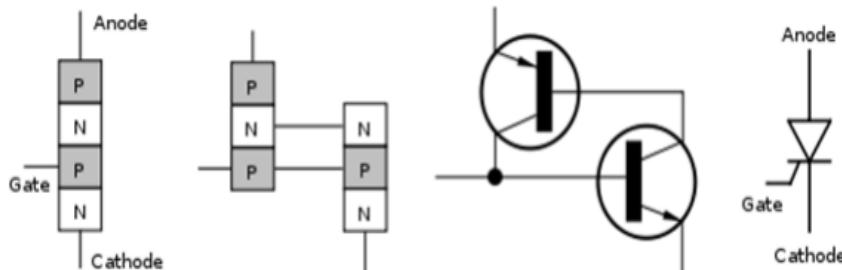
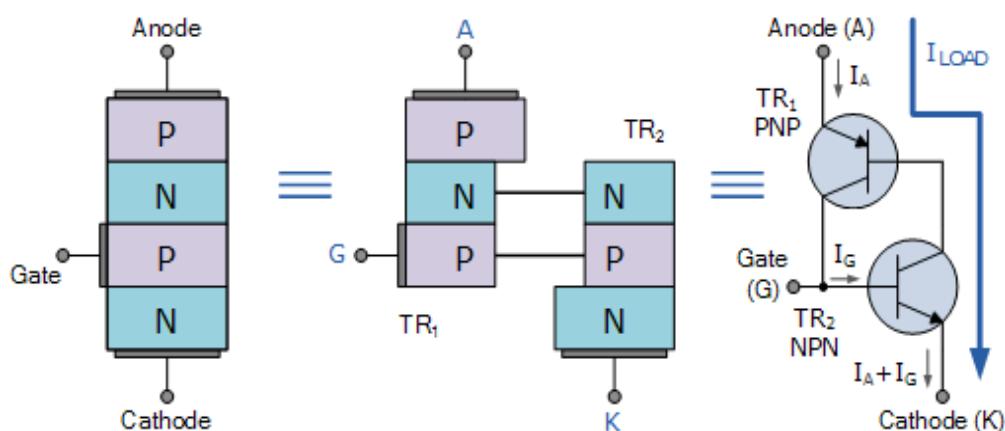
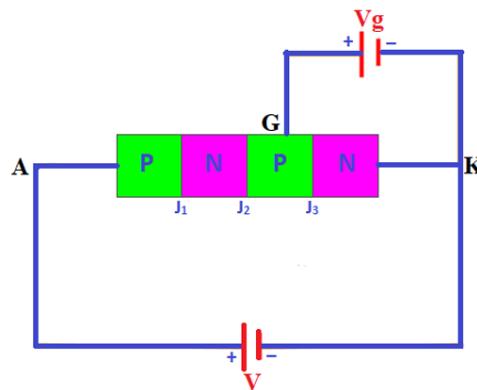
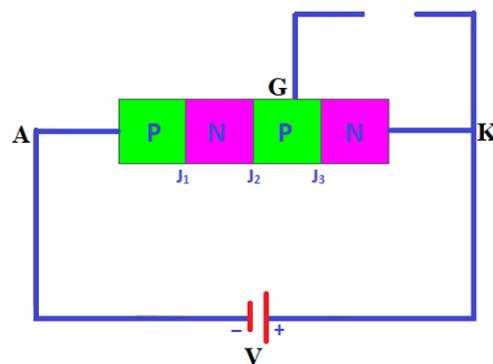


Figure 8.2 Semiconductor controlled rectifier





Forward Conducting Mode of SCR



Reverse Blocking Mode of SCR

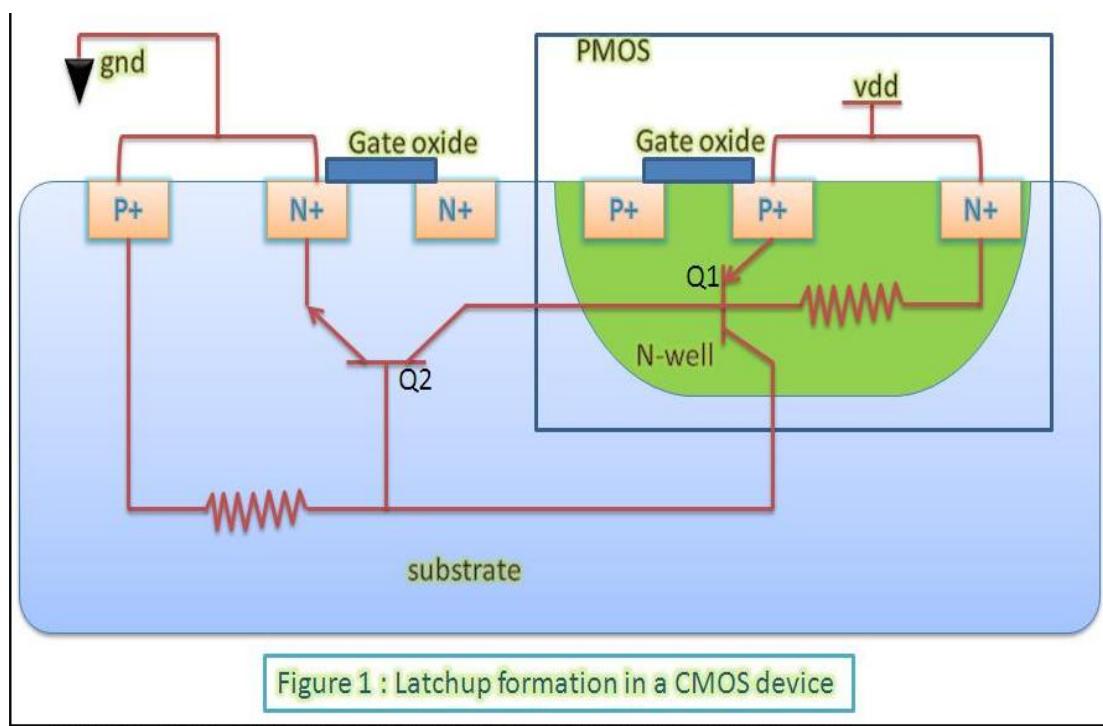
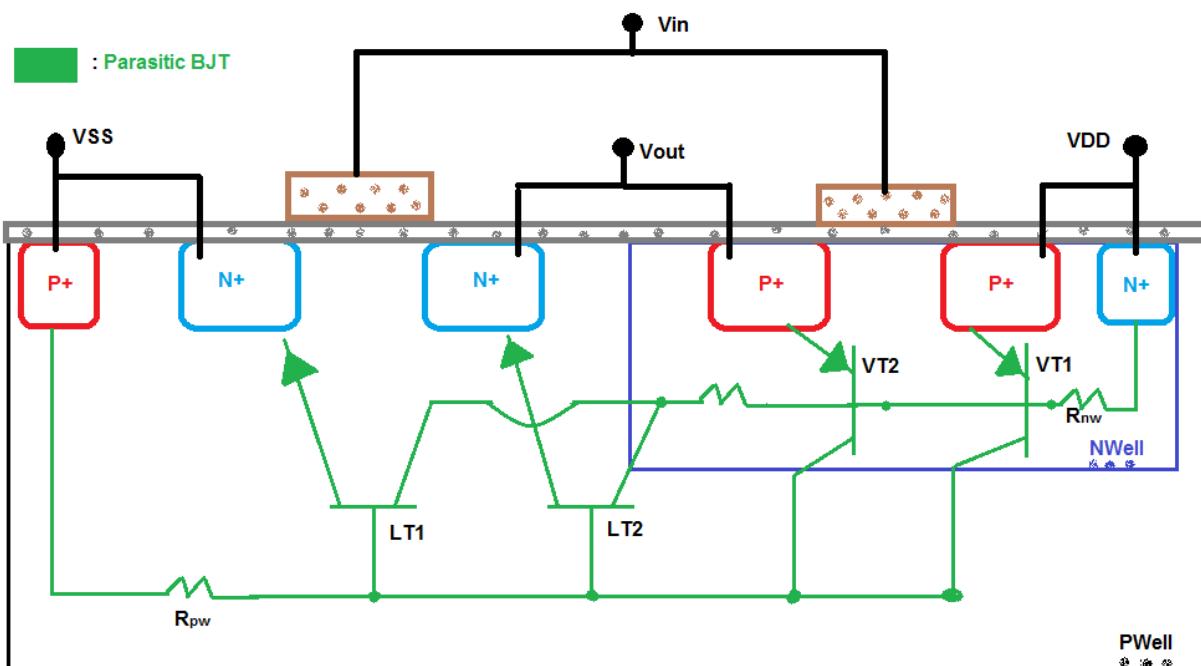
Latch-up:

- Latch-up is type of short circuit which occurs in CMOS structure.
- Latch-Up is a condition where a low impedance path is created between a supply pin and ground pin. This condition is caused by a trigger (current injection or overvoltage)
- Latch-Up is not a risk if the voltage and current levels applied to the device adhere to the absolute maximum ratings.
- Common reason due to positive & negative input & output voltage spike signal.
- Exceeding the rail voltage than a diode voltage.
- There is a parasitic BJT form in CMOS, this spike signal triggered the parasitic BJT.
- And it creates the low impedance path between VDD & VSS. So short circuit happens between VDD & VSS.
- This will damage the IC permanently.

- Another reason ESD & microwave interface & supply voltage is exceeding the absolute maximum rating
- A common cause of latch-up is a positive or negative voltage spike on an input or output pin of a digital chip that exceeds the rail voltage by more than a diode drop
- Another cause is the supply voltage exceeding the absolute maximum rating, often from a transient spike in the power supply.

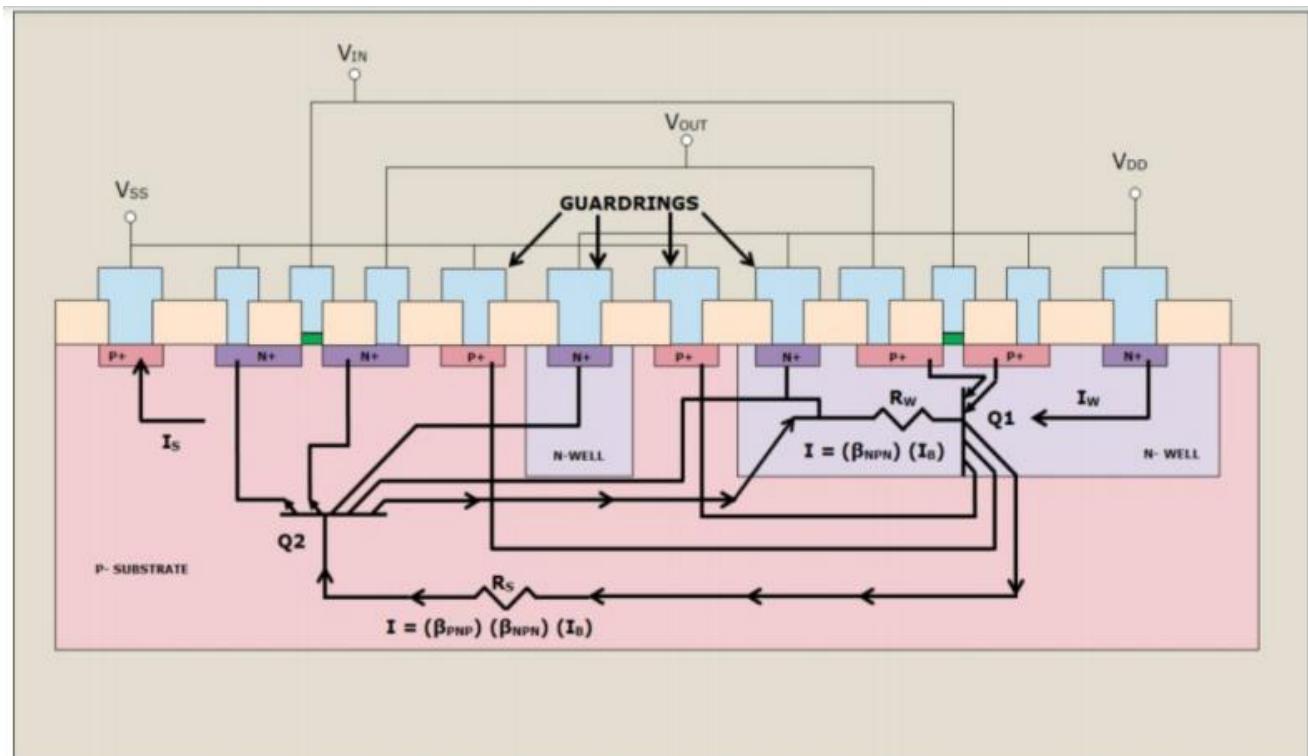
Operation of Latch-Up:

- The equivalent circuit of the parasitic components of CMOS inverter as shown in the below figure.
- When the signal at the output node is 0.7V higher than the VDD (overshooting), the bipolar (VT2) may be turned on first and similarly the bipolar (LT2) will be turned on while the output signal is lower than -0.7V (undershooting).
- For the fact that the collector of each BJT (i.e. VT2) plays the role as base of the other transistor (i.e. LT2) and the collected carriers will reduce the potential difference between emitter and base of the transistor (LT2).
- Under this situation, the positive feedback loop will make the concentration of minority carrier increased to higher than the doping concentrations of both the NW and PW
- Subsequently, the potential barrier NW and PW will be disappeared and then obtains a highly conductivity path between VDD and VSS. This may result in the circuit malfunction, and destroy the device in the worst case



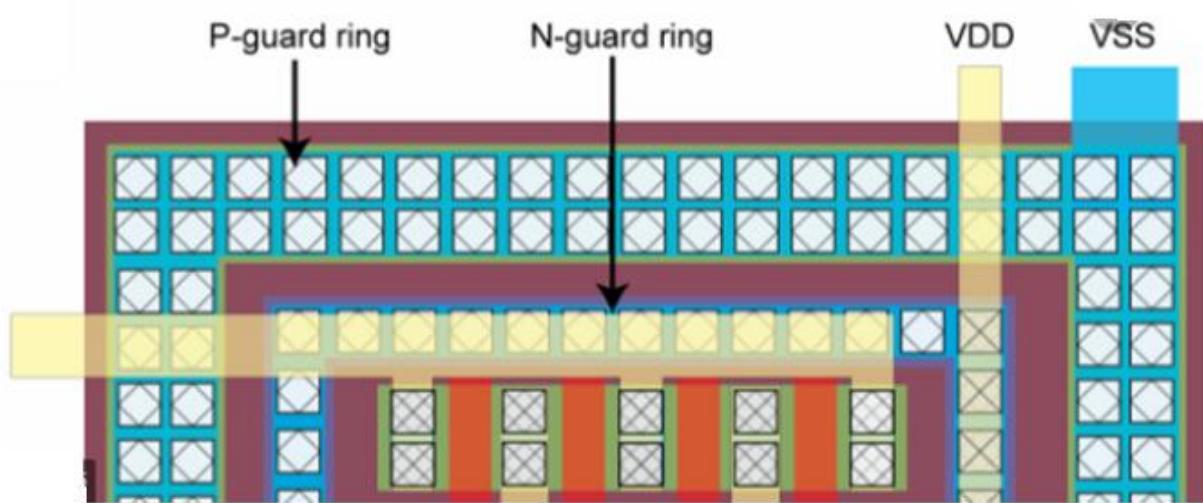
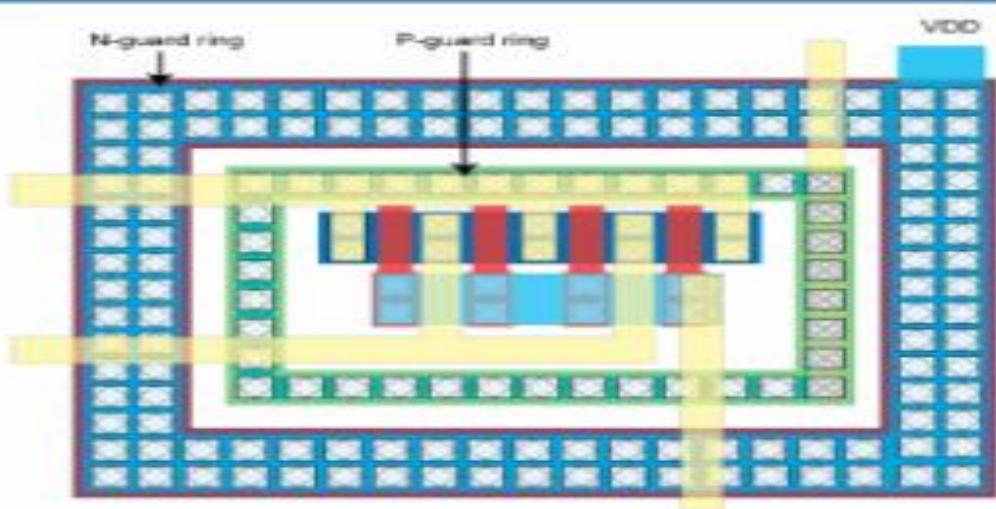
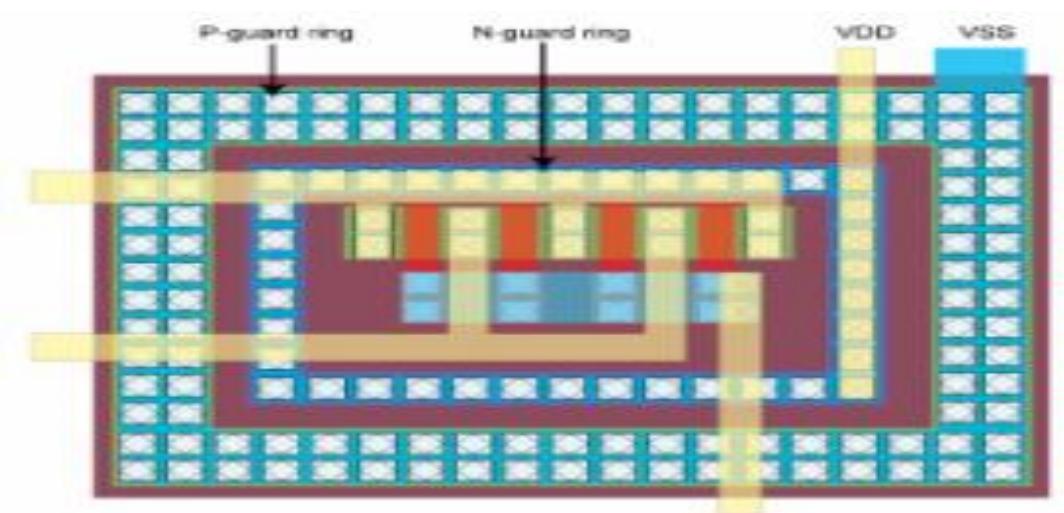
❖ How to solve latch-up problem?

1. NMOS and PMOS should be away from 15um to each other or from ESD diode.
2. NMOS and PMOS should be close double guard ring
3. ESD circuit
4. STI – Silicon trench isolation (between PMOS & NMOS)
5. Deep N-well method
6. Adding the insulation layer between PMOS & NMOS
7. Device is fabricated with lightly doped epitaxy layer on heavily doped substrate

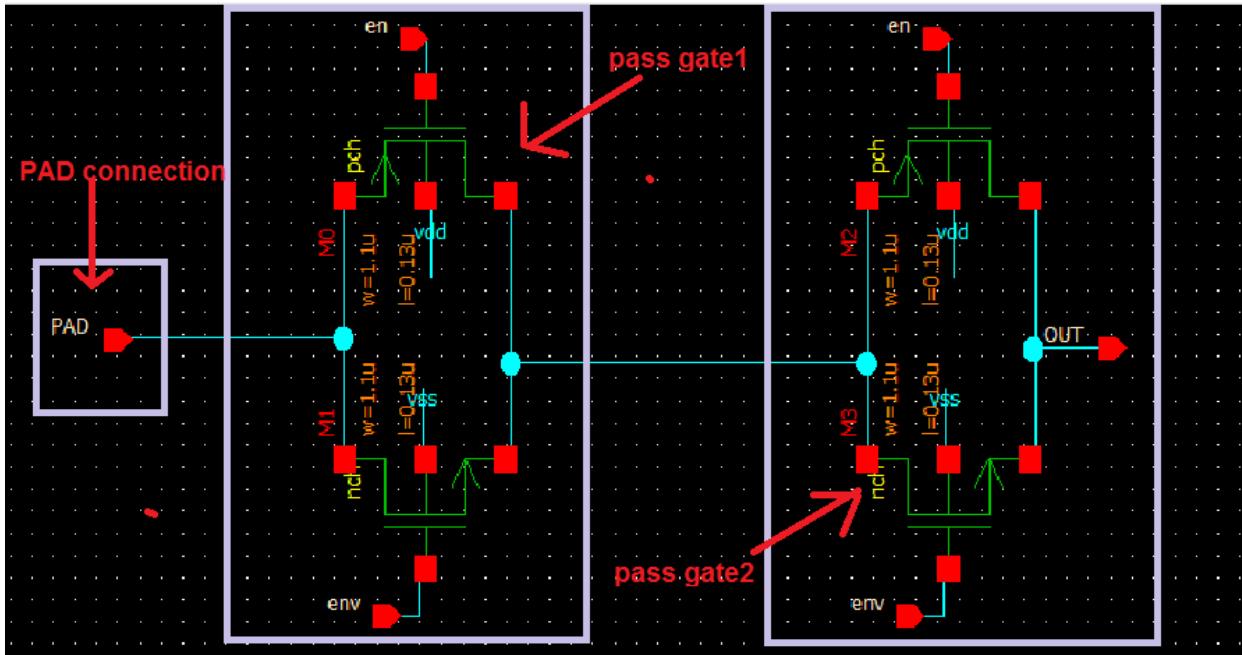


Video link for latch-up

1. <https://youtu.be/EYHFLBI5fig>
2. <https://youtu.be/c-JbhXM0oIo>



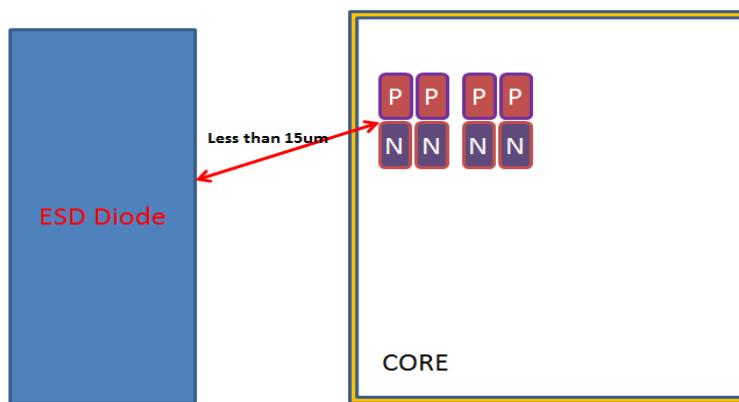
Condition1:



Solution:

1. NMOS and PMOS of pass gate1 should be in **double** close guardring or away from 15um to each other.
2. NMOS and PMOS of pass gate2 its need one guardring but it should be close.
3. Diffusion/OD width of guardringshould be follow to the technology minimum required width.

Condition2:



Solution: NMOS and PMOS should be in close guardring (single)**OR** away from 15um-40um and more to ESD diode.(spacing depend on technology)

19.ELECTROSTATIC PROTECTION AND CLAMP

Electrostatic discharge (ESD)

- Electro Static Discharge (ESD) is sudden flow of static electricity between two electrically charged objects for a very short duration of time.
- One of the most important reliability problems in the IC industry
- ESD protection circuits divert high currents away from the internal circuitry and clamp high voltages during an ESD stress.
- The ESD protection is required to prevent the damage of the GOX of a MOSFET from the static charge buildup.
- The model was intended to represent the interaction of the electrical discharge of a human being, who is charged, with a component or an object. The charged source then touches a component or an object using a finger. The physical contact between the charged human being and the component or object allows for current transfer between the human being and the object.
- Electrostatic Discharge (ESD) is one of the major reliability issues in advanced CMOS technologies. Research has shown that only I/O based ESD protection circuits are inadequate in providing necessary ESD protection. Therefore, it is important to have an effective ESD power supply clamp across the power supply rails so that the ESD event will be discharged through it and protects the circuit core.
- The typical chip-level ESD protection scheme in which an ESD power supply clamp is connected between the two power supply rails. The main goal of ESD protection circuits is to provide a low-resistive discharge path from any two pins on the chip. The circuit core is susceptible to ESD damage if there are only ESD protection circuits at the I/O pads. As shown in Fig. 1, the ESD clamp provides the discharge path for an ESD event that happens between the two power rails (PSD-mode, NDS-mode). The clamp is also part of the discharge path for both PS-mode and ND-mode. Thus, it is important to have an effective ESD power supply clamp across the power supply rails.

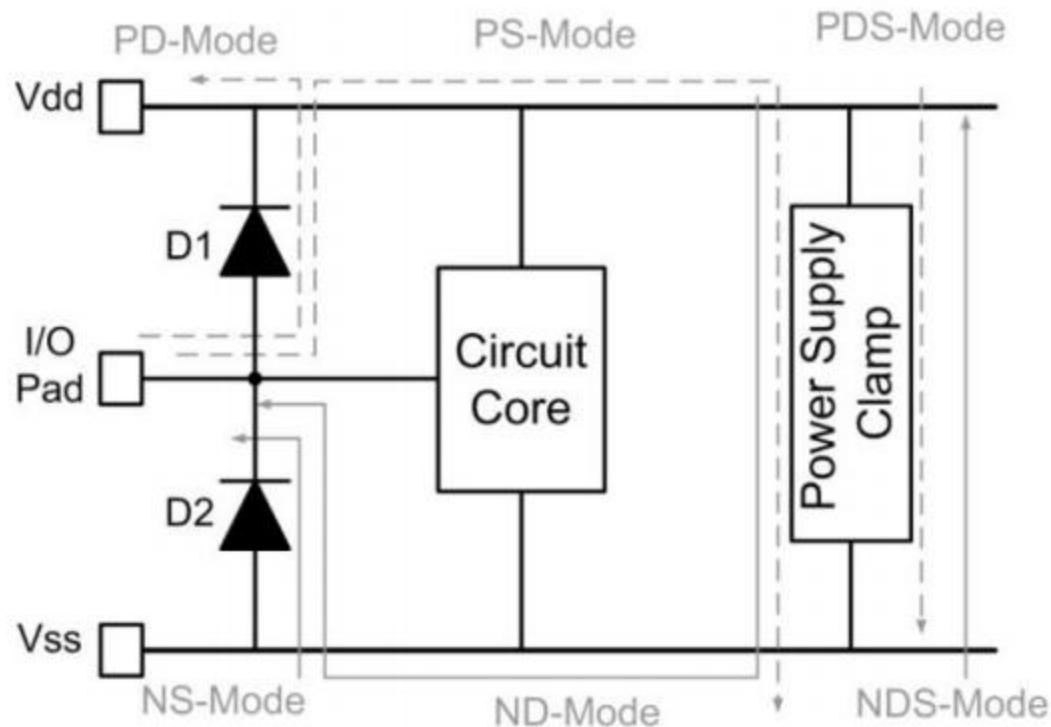


Fig. 1. Typical chip-level ESD protection scheme under different ESD stresses.

Diode clamps

- Diodes turn on if pad voltage:
 - Exceeds $VDD +0.7V$
 - Drops below $VDD -0.7V$

Resistor

- Limits the current
- Protects secondary protection

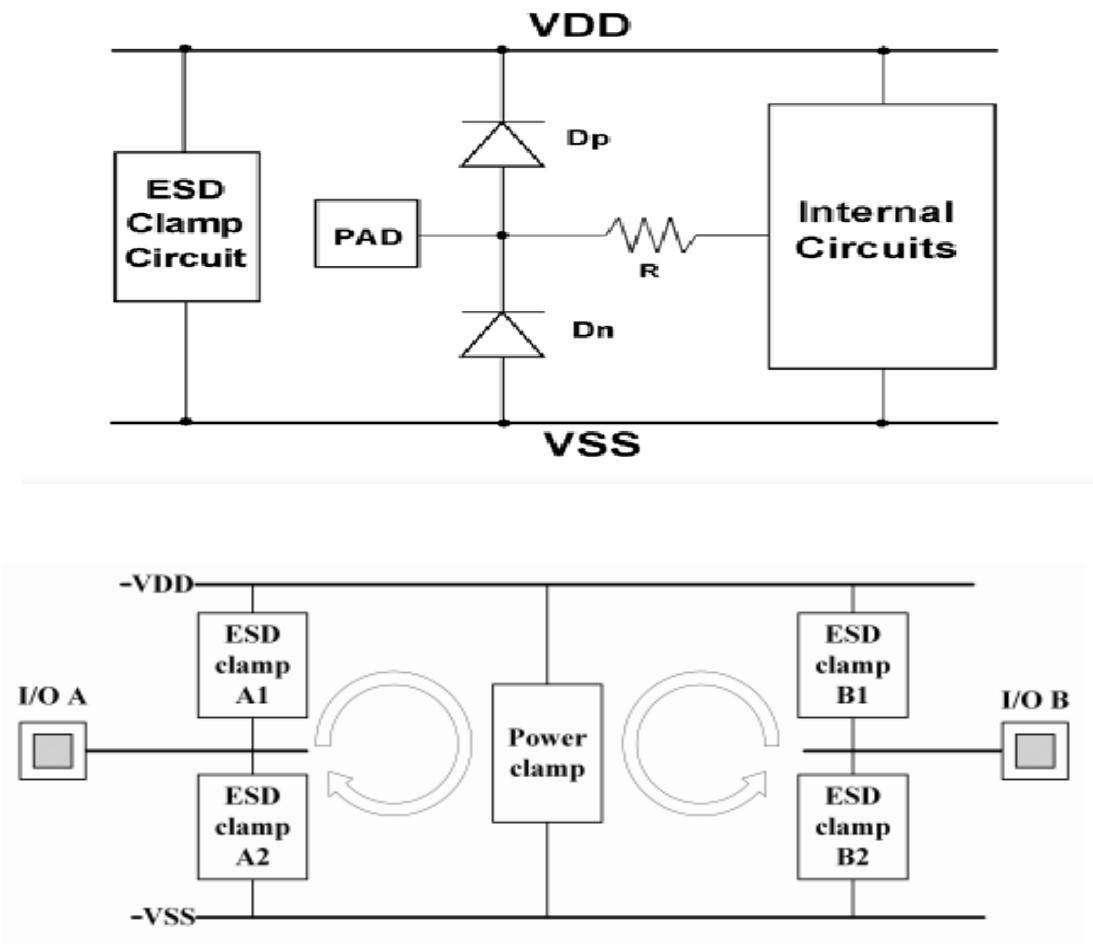
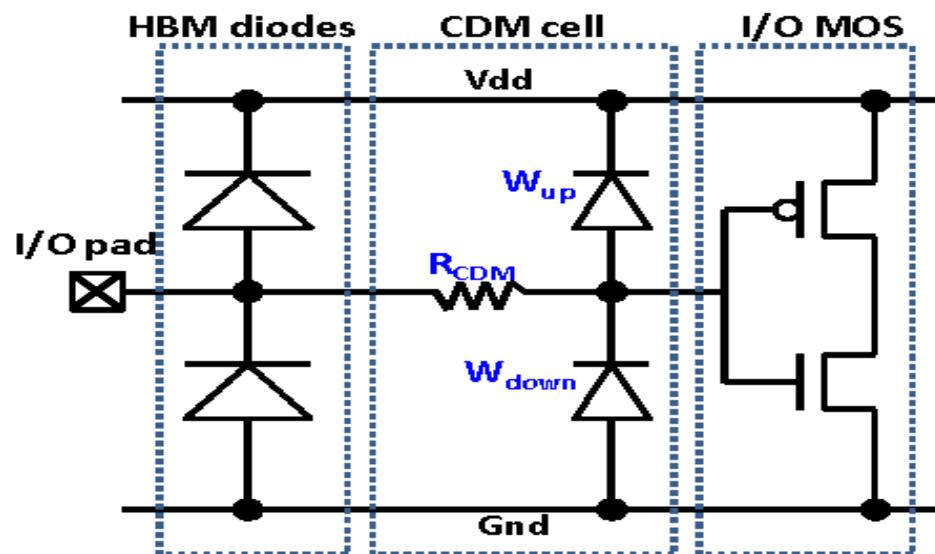
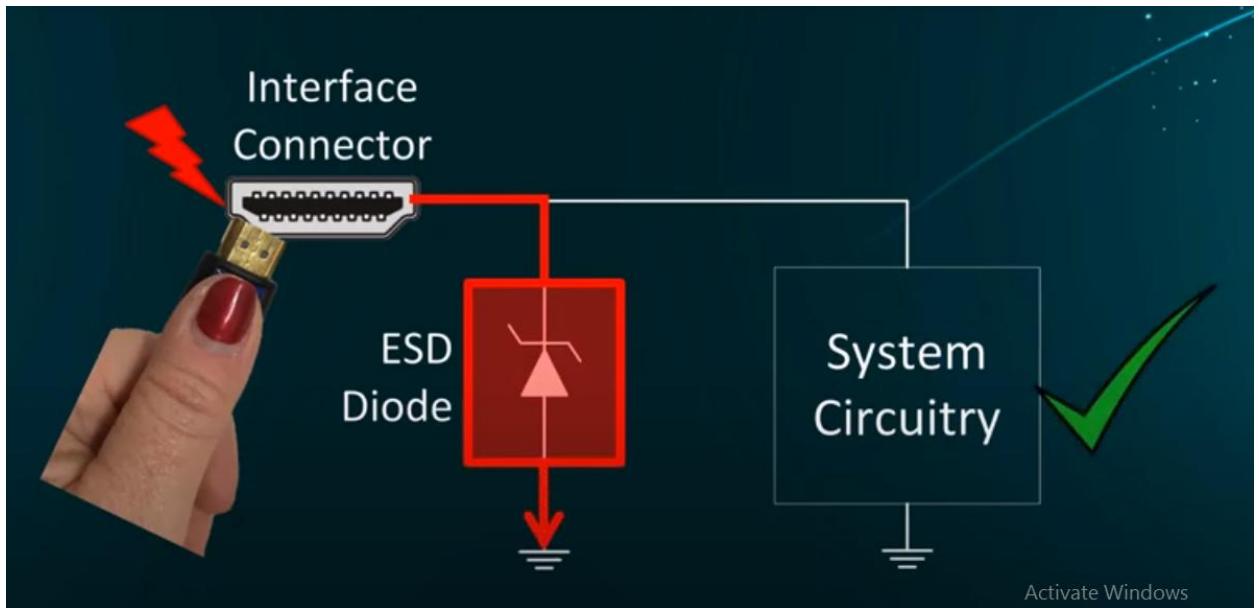
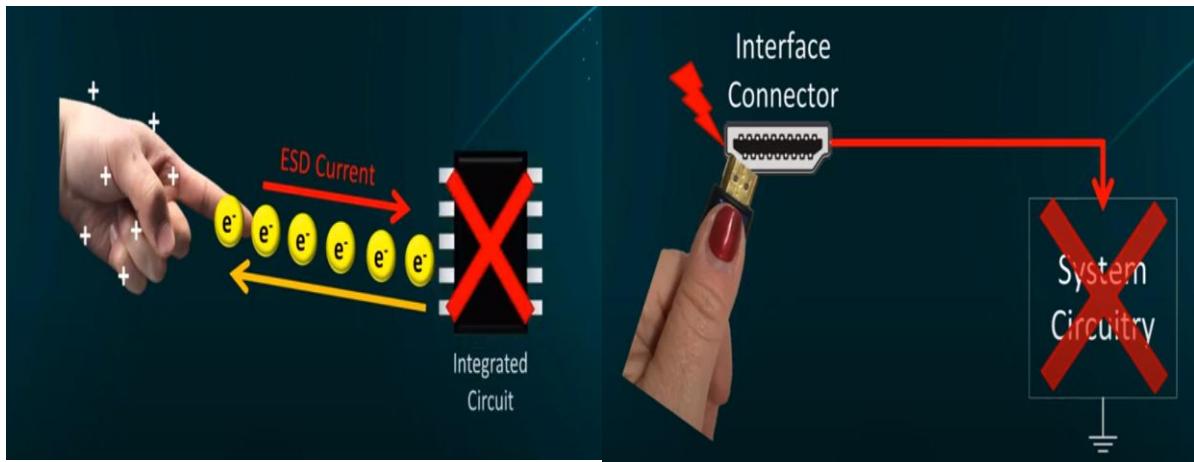


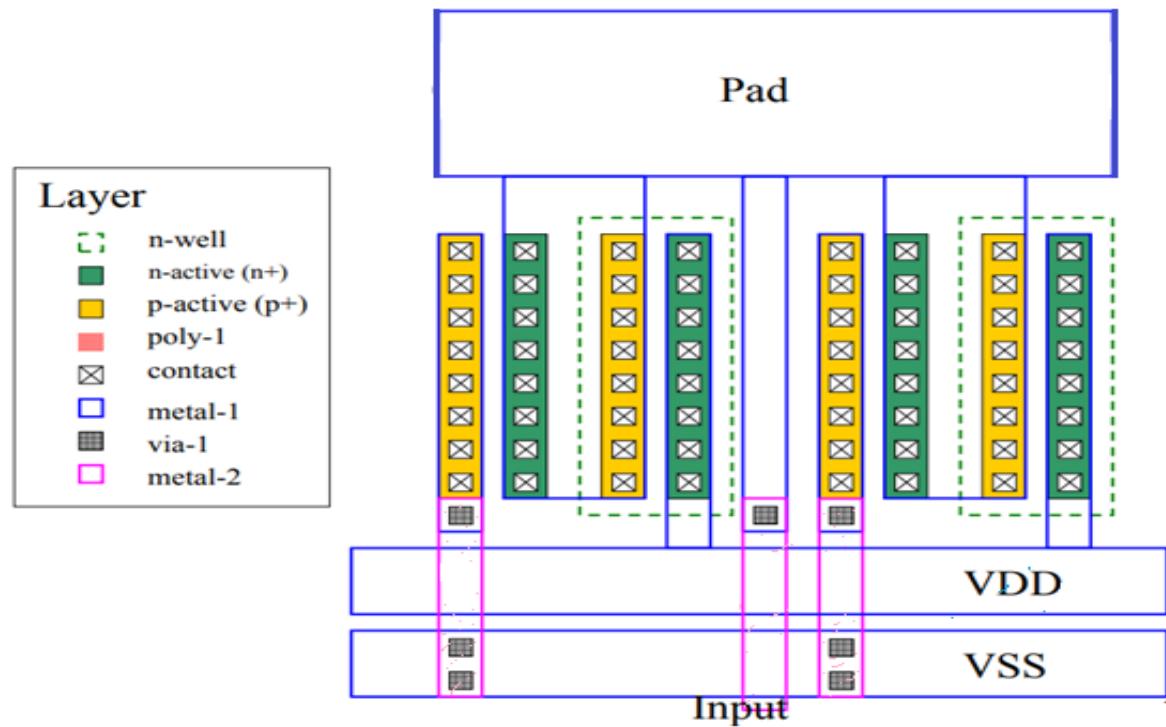
Figure 1.8: Power-grid-based ESD Protection Network



primary ESD protection and Secondary CDM protection at

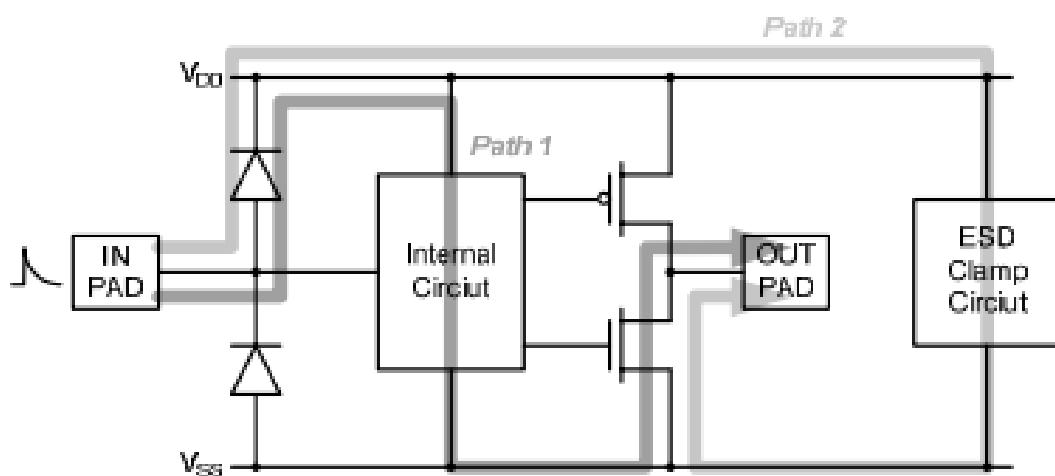


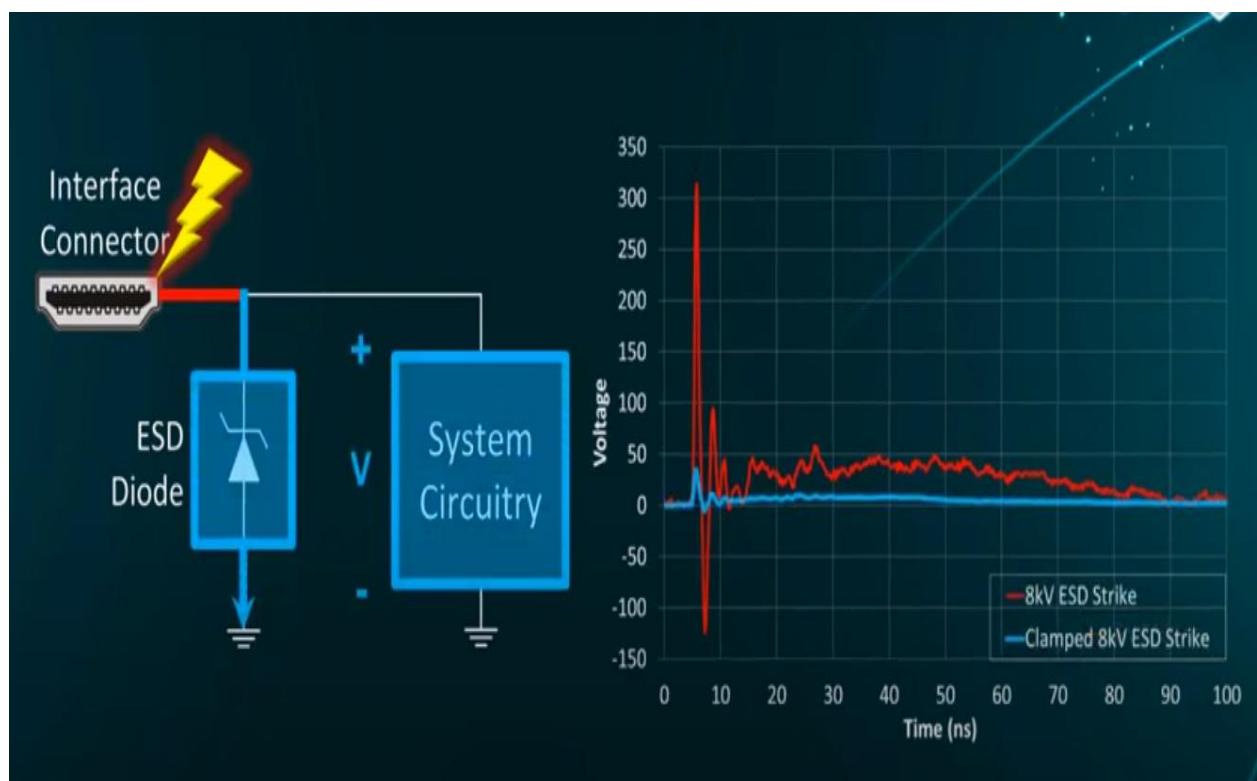
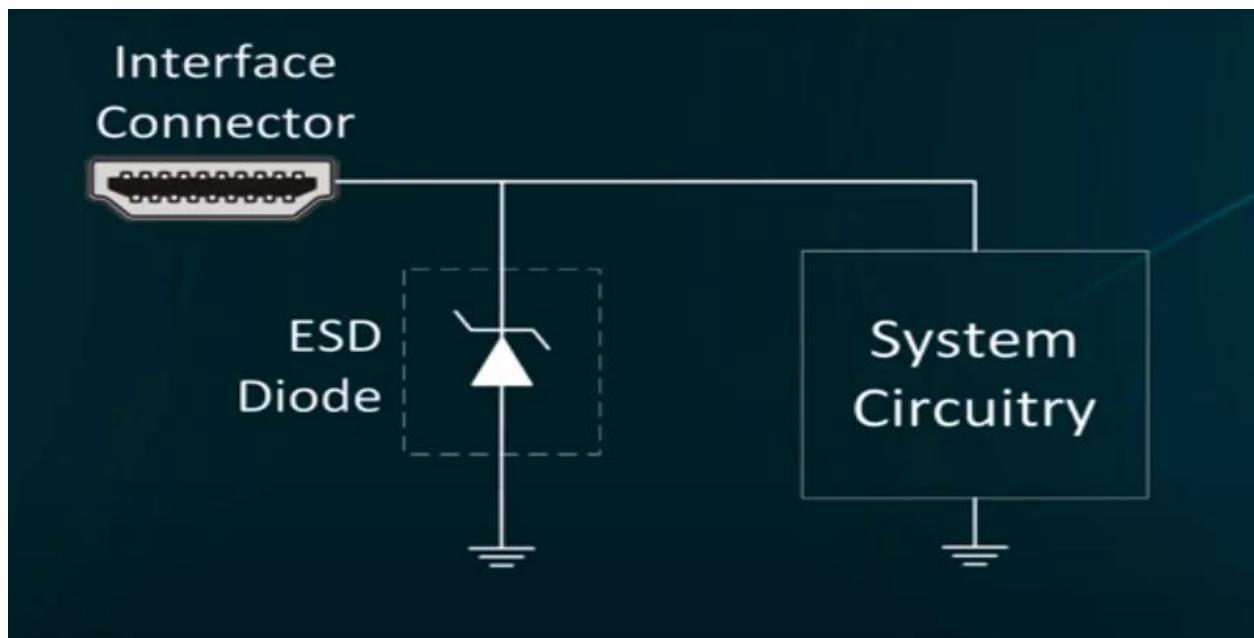
Layout sample of pad ESD protection:



CLAMP:

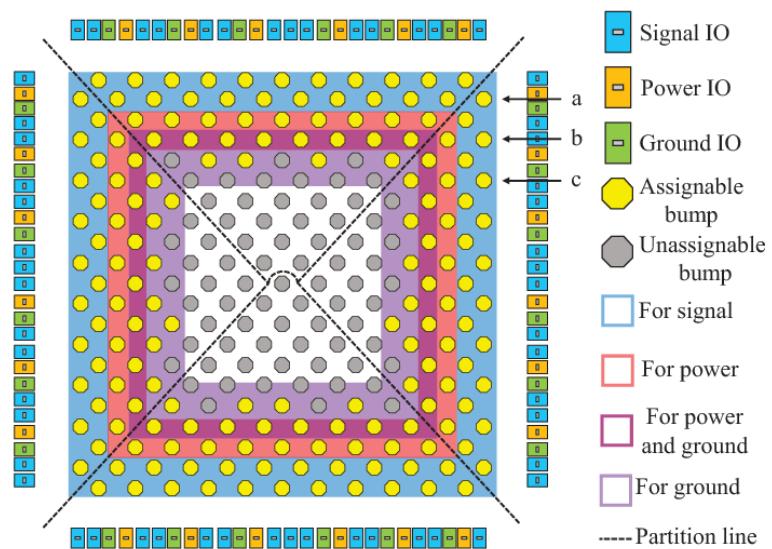
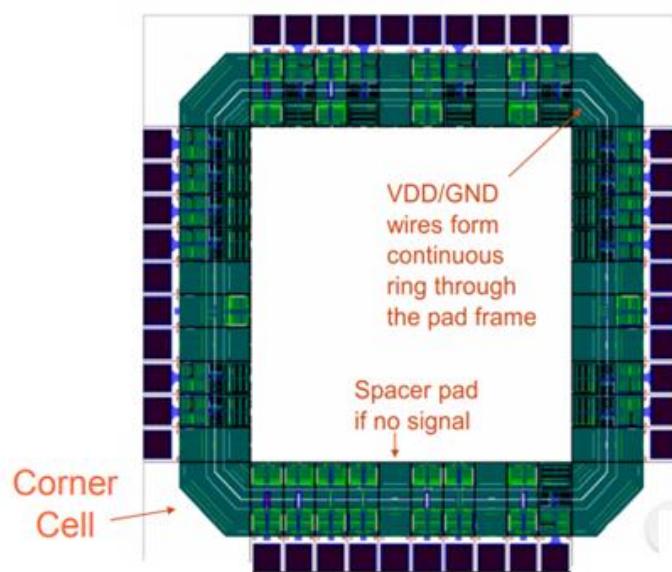
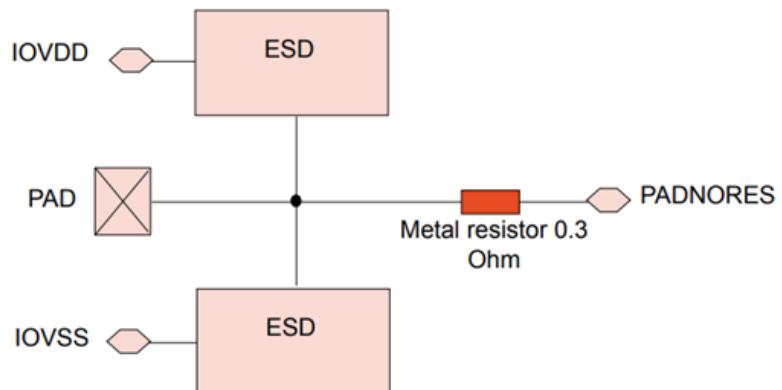
The first step in designing an ESD power supply clamp is technology characterization in which several aspects about the technology used and the application must be known. Firstly, the breakdown voltage of the circuit core due to an ESD stress should be known. Secondly, the target ESD protection level should also be known in order to design an ESD protection circuit that limits the voltage across the circuit core below the breakdown voltage. For example, the oxide breakdown voltage for 65 nm CMOS technology is 5 V and the target HBM is 1.5 kV [4]. Therefore, under 1.5 kV HBM stress the voltage between the stressed pins should be less than 5 V to ensure the correct functionality. Thirdly, based on the available components those can be utilized as ESD clamping elements in the technology used and the target application, the design decisions can be taken to tradeoff the ESD performance, the design area, and the leakage current. For instance, silicon controlled rectifier (SCR) has a high current carrying capability; however, it cannot be used in an application where ionizing radiation or hot switching is expected.





20.I/O PADS OR BOND PADS OR WIRE BONDING

- Input/Output circuits (I/O pads) are intermediate structures connecting internal signals from the core of the integrated circuit to the external pins of the chip package.
- The circuitry on a chip has to connect with other circuits. These may be chips or display devices, transducers or electro-mechanical devices and the capacitance connected to the chip could be very large. In some cases the devices being driven will require or supply TTL signal levels, in others they may be liable to be short circuits, have high noise levels or be liable to discharge spikes of several kV. Each of these situations will require the imposition of circuitry to interface the chip to the external environment. Most IC designers avoid the problem of pad design and take pad drivers from standard libraries.
- Physically, pads are the squares of metal, generally 100-150 um square, that are connected to the pins of the package with bonding wires. The word *pad* is often used to also include the circuitry that is used to interface the CMOS logic within the IC (typically composed of near minimum-geometry transistors) to the outside world.
- At least two pads in each circuit will be used to connect the chip to the V_{DD} and V_{SS} power supply lines, while other pads will be used for input connections and output connections. Some pads may also be required to be bi-directional, (for use both with input signals and output signals). In such cases there is usually a control connection to determine the direction of signal transfer.
- An important function for all pad driver circuitry is the protection of the chip circuitry against destruction due to overvoltage pulses or sustained overvoltages. These may be due to electrostatic discharges or due to faults on other circuitry that cause unexpectedly high voltages to be applied to the chip pins.



21. ELECTRICAL RULE CHECK / SOFTCHECK

ERC: Checking for all electrical connections that are considered dangerous.

Generally checks the wells and substrate contacts

- Any floating N or P -Wells
- Any floating substrate
- N- Well tap connected to VSS/GND and P-Well tap connected to VDD/PWR
- Any unconnected inputs or shorted outputs
- Input gate of logic cell is directly connected to VDD or GND.
- Any floating poly
- Any floating DNW

22. DENSITY

Why density fix is important

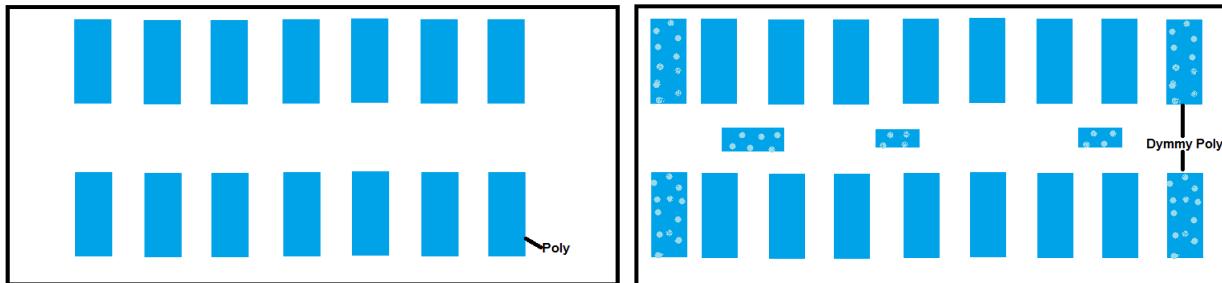
- When fabricating chips, the silicon is polished; it is preferable to have uniform concentration of all layers all throughout the chip, so it would be polished uniformly.
- Maximum density causes the chip to be polished unevenly, causing less dense area to be polished out.

Instruction of density error

- There are some DRC rules obligating that the density of certain layers in the layout does not get below or above a certain value.
Ex: Limits are typically $30\% < \text{density} < 80\%$ for metal
- The Density error not much important at the cell/block level, but usually they are taken care of at the chip assembly step.
- However we don't need to meet the density rules manually, usually there are fill tools that come with the tool that will place dummy layer in the area where the density is too low.
- Maximum density is exactly the opposite of the minimum density error.
- After the dummy fill is done, we don't connect them to any potential. They are left floating in the layout. Leaving the dummy fill floating will not affect the circuit working because we take care of the DRC rules which will specify spacing between the metal drawn layer and the metal dummy fill layer.
- If any net in the layout is parasitic critical, we usually don't put dummy fills over it to avoid any kind of capacitive effects.
- Due to the density issue, Due to a large variations in density can cause thermal expansion stress
- Inserting dummy fill can help make uniform layers, this practice may degrade the performance of sensitive analog circuitry of high-speed digital blocks due to added parasitic capacitance

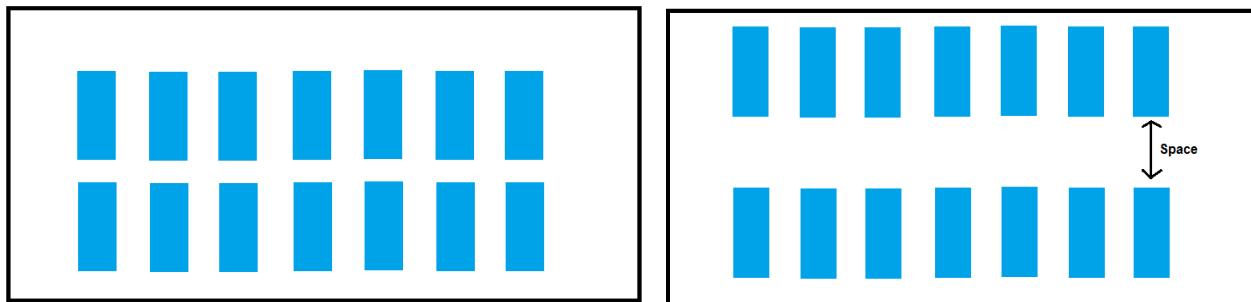
How to fix min density of Poly

- Add dummy Poly or dummy device on the vacant space.



How to fix Max density of Poly

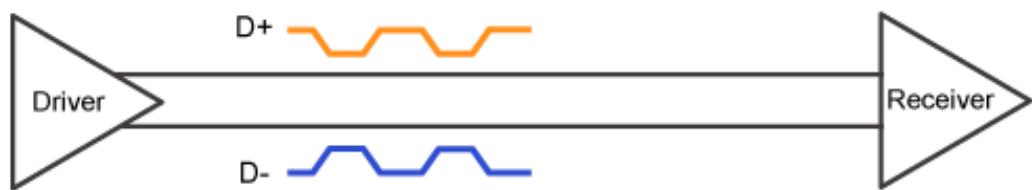
- Make space between the Poly/Device.



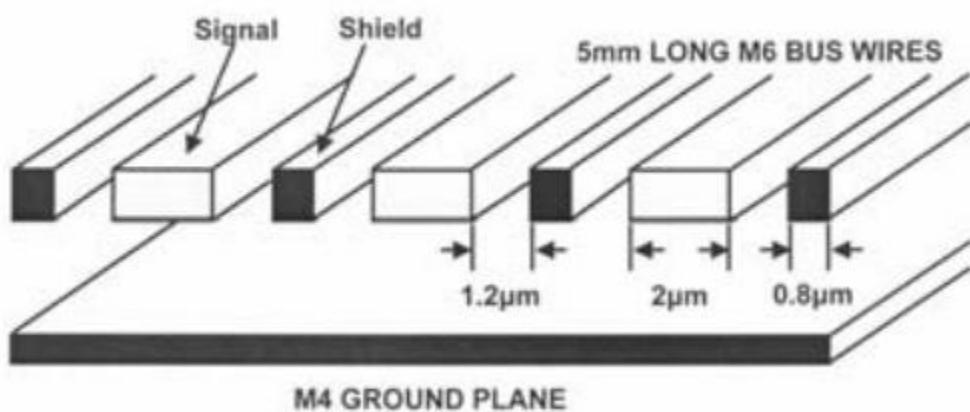
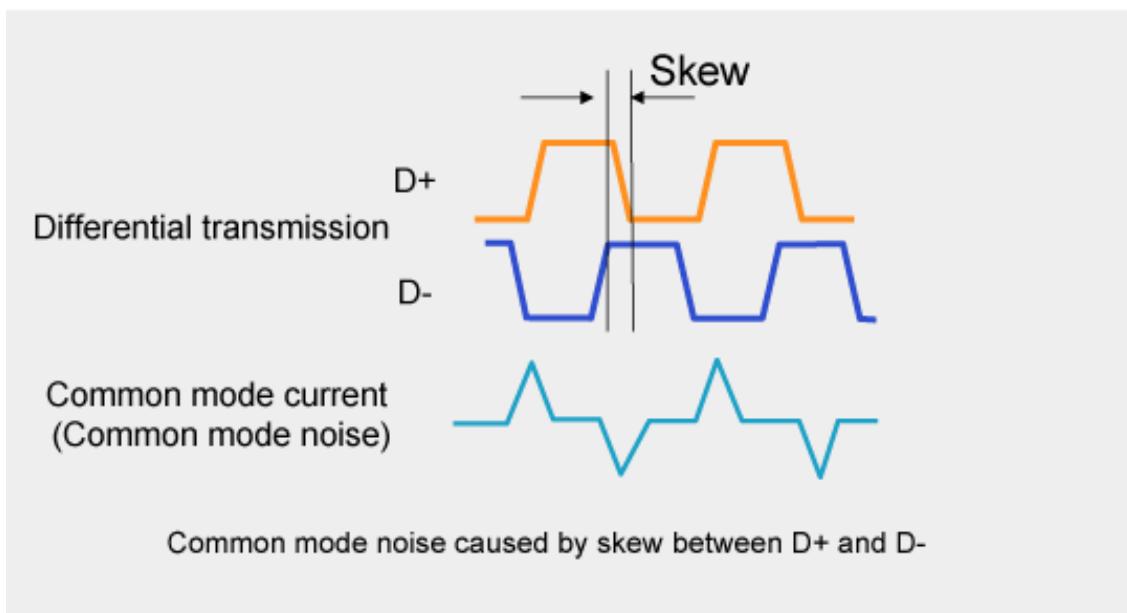
23. HIGH SPEED LAYOUT DESIGN

High Frequency clock Transmission:

- The effect of random signal lines on the on-chip inductance is quantitatively investigated, using an S-parameter-based methodology and a full wave solver, leading to an empirical model for high-frequency inductance. The results clearly indicate that the random signal lines as well as designated ground lines provide return paths for gigahertz-frequency signals. In particular, quasi TEM-wave-like propagation mode is observed above 10 GHz, revealing a unique relationship between capacitance and inductance of the signal line. Incorporating the random capacitive coupling effect, our frequency-dependent RLC model is confirmed to be valid up to 100 GHz.
- Short, medium, and long on-chip interconnections having line widths of 0.45-52 μm are analyzed in a five-metal-layer structure. We study capacitive coupling for short lines, inductive coupling for medium-length lines, inductance and resistance of the current return path in the power buses, and line resistive losses for the global wiring. Design guidelines and technology changes are proposed to achieve minimum delay and contain crosstalk for local and global wiring. Conditional expressions are given to determine when transmission-line effects are important for accurate delay and crosstalk prediction



High-speed data communications by differential transmission



Crosstalk:

Crosstalk is the unwanted coupling of signals between parallel traces. Proper routing and layer stack-up through microstrip and stripline layouts can minimize crosstalk.

To reduce crosstalk in dual-stripline layouts, which have two signal layers next to each other, route all traces perpendicular, increase the distance between the two signal layers, and minimize the distance between the signal layer and adjacent plane.

Use the following steps to reduces crosstalk in either microstrip or stripline layouts:

- Widen spacing between signal lines as much as routing restrictions will allow. Try not to bring traces closer than three times the dielectric height.
- Design the transmission line so that the conductor is as close to the ground plane as possible. This technique will couple the transmission line tightly to the ground plane and help decouple it from adjacent signals.
- Use differential routing techniques where possible, especially for critical nets (i.e., match the lengths as well as the gyrations that each trace goes through).

- If there is significant coupling, route single-ended signals on different layers orthogonal to each other. Minimize parallel run lengths between single-ended signals. Route with short parallel sections and minimize long, coupled sections between nets. Crosstalk also increases when two or more single-ended traces run parallel and are not spaced far enough apart. The distance between the centers of two adjacent traces should be at least four times the trace width. To improve design performance, lower the distance between the trace and the ground plane to under 10 mils without changing the separation between two traces.

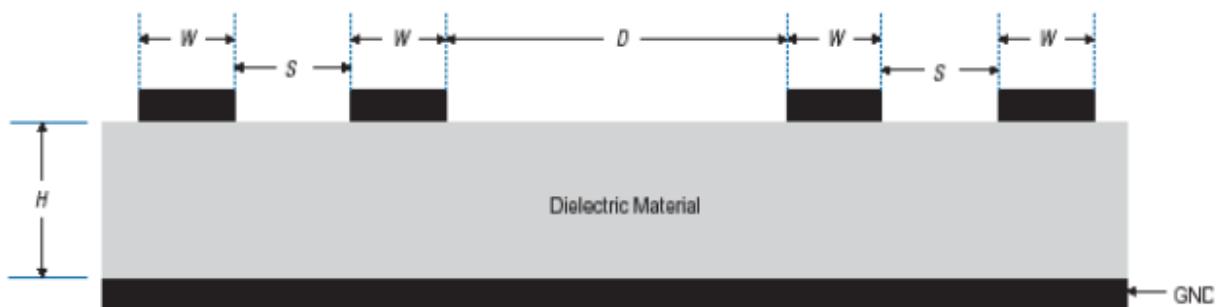
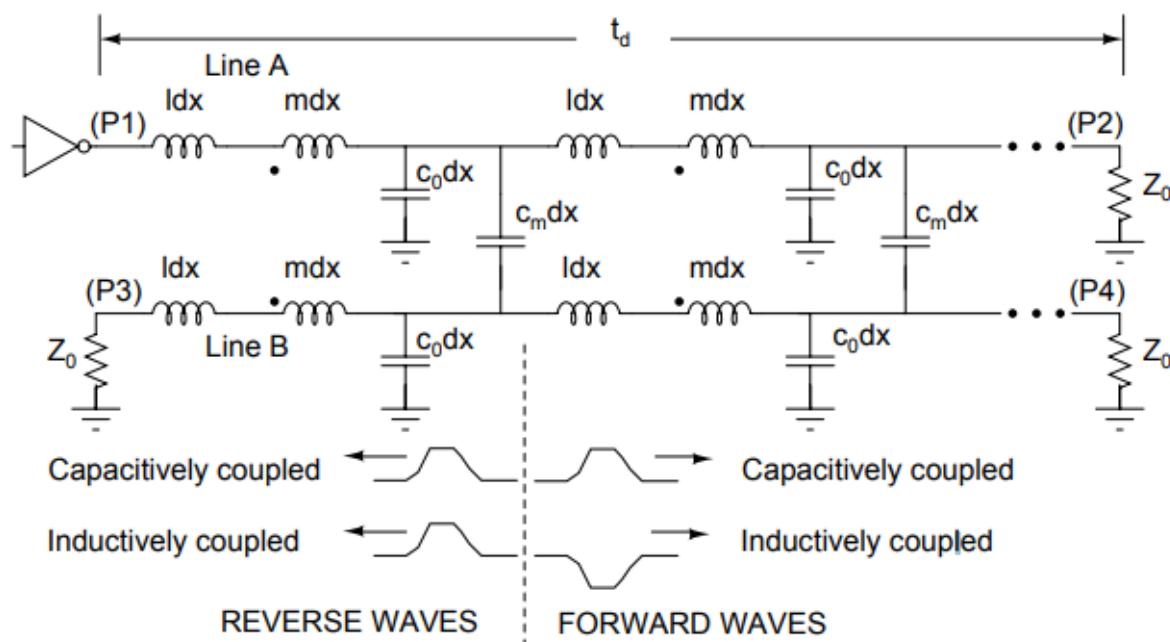


Figure 10. Example of Separating Traces for Crosstalk

Forward and reverse crosstalk waves created by capacitive and inductive coupling:



Video link :<https://youtu.be/0A941cqVP2A>

Signal Integrity

For a single-ended trace, like clock transmission line, it could be improved using the following guidelines:

- Keep clock traces as straight as possible. Use arc-shaped traces instead of right-angle bends.
- Do not use multiple signal layers for clock signals.
- Do not use via in clock transmission lines. Via can cause impedance change and reflection.
- Place a ground plane next to the outer layer to minimize noise. If you use an inner layer to route the clock trace, sandwich the layer between reference planes.
- Terminate clock signals to minimize reflection.
- Use point-to-point clock traces as much as possible.
- Make sure $D > 2S$ to minimize the crosstalk between the two differential pairs.
- To minimize reflection noise, place the differential traces $S = 3H$ as they leave the device.
- Keep the distance between the differential traces (S) constant over the entire trace length.
- Keep the length of the two differential traces the same to minimize the skew and phase difference.
- Avoid using multiple via, because they can cause impedance mismatch and inductance.

ClkroutingBends

A right angle in a trace can cause more radiation. The capacitance increases in the region of the corner, and the characteristic impedance changes. This impedance change causes reflections.

- Avoid right-angle bends in a trace and try to route them at least with two 45° corners. To minimize any impedance change, the best routing would be a round bend (see Figure).
- Separate high-speed signals (for example, clock signals) from low-speed signals and digital from analog signals; again, placement is important.
- To minimize crosstalk not only between two signals on one layer but also between adjacent layers, route them with 90° to each other.

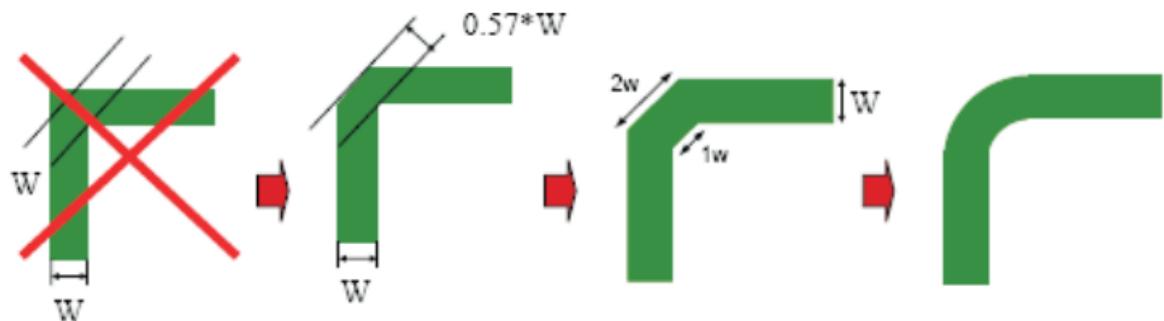


Figure: Poor and Good Right Angle Bends

The use of vias is essential in most routings, but the designer has to be careful when using them. They add additional inductance and capacitance, and reflections occur due to the change in the characteristic impedance. Vias also increase the trace length.

- Avoid vias in differential traces. If it is impossible to avoid them, use vias in both traces or compensate the delay also in the other trace.

Clock Buffering Mechanisms

- Clock signal is global in nature.
 1. Clock lines are typically very long.
 2. Long wires have large capacitances, which limit the performance of the system.
 3. RC delay plays a big factor.
- RC delay cannot be reduced by making the wires wider.
 1. Resistance reduces, but capacitance increases.
- To reduce RC delay, buffers are used.
 1. Also helps to preserve the clock waveform.
 2. Significantly reduces the delay.
 3. May occupy as much as 5% of the total chip area.

Power Noise Filtering or Decap

To decrease the low-frequency (below 1 kHz) noise caused by the power supply, filter the noise on power lines at the point where the power connects to the chip and to each device. Place a capacitor where the power supply lines enter the chip. If you use a voltage regulator, place the capacitor immediately after the pin that provides the VDD signal to the device(s). (Capacitors not only filter low-frequency noise from the power supply, but also supply extra current when many outputs switch simultaneously in a circuit.) To filter power supply noise, place decoupling capacitors as close as possible to each VDD and GND, make it less resistance from Decap to pwr/Gnd pump or regulator.

Power Distribution

A system can distribute power throughout the chip with either power planes or a power bus network/grid. You can use power planes on multi-layer metals that consist of two or more metal layers that carry VDD and GND to the devices. Because the power plane covers the full area of the chip, its DC resistance is very low. The power plane maintains VDD and distributes it equally to all devices while providing very high current-sink capability, noise protection, and shielding for the logic signals on the chip. It is recommended to use lower planes to distribute power. The power bus network which consists of two or more wide metal traces that carry VDD and GND to devices. When designing with power bus networks, be sure to keep the trace widths as wide as possible. The main drawback to using power bus networks is significant DC resistance. It is recommended to separate analog and digital power planes. For fully digital systems that do not already have a separate analog power plane, it can be expensive to add new power planes. However, you can create partitioned islands (split planes). An example board layout with phase-locked loop (PLL) ground islands is shown on Figure.

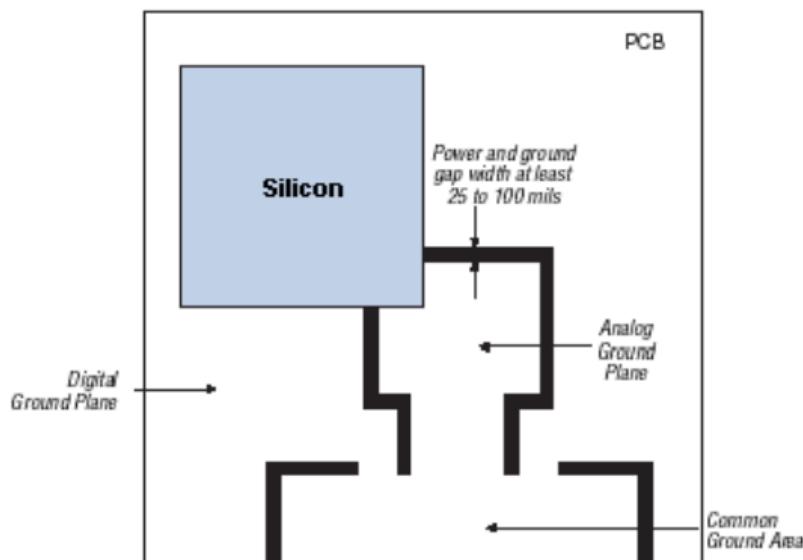


Figure 12. Example Board Layout with Phase Locked Loop (PLL) Ground Islands

If your system shares the same plane between analog and digital power supplies, there may be unwanted interaction between the two circuit types. The following suggestions will help to reduce noise:

- For equal power distribution, use separate power planes for the analog (PLL) power supply. Avoid using trace or multiple signal layers to route the PLL power supply.
- Use a ground plane next to the PLL power supply plane to reduce power-generated noise.
- Place analog and digital components only over their respective ground planes.

Ground Bounce

As digital devices become faster, their output switching times decrease. Faster switching times cause higher transient currents in outputs as they discharge load capacitances. These higher currents, which are generated when multiple outputs of a device switch simultaneously from a logic high to a logic low, can cause a board-level phenomenon known as Ground Bounce. Many factors contribute to ground bounce. Therefore, no standard test method predicts ground bounce magnitude for all possible chip environments. Determine each condition and each device's relative contributions to ground bounce by testing the device under these conditions. Load capacitance, socket inductance, and the number of switching outputs are the predominant conditions that influence the magnitude of ground bounce in programmable logic devices.

It is recommended to follow some design methods to reduce ground bounce:

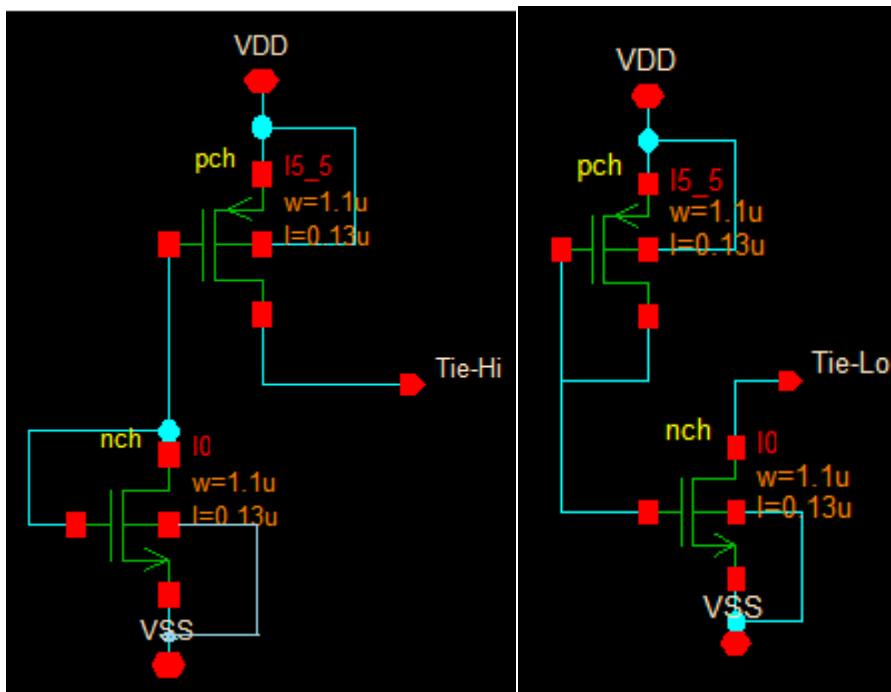
- Use wide, short traces between the via and capacitor pads, or place the via adjacent to the capacitor pad.
- Traces stretching from power pins to a power plane (or island, or a decoupling capacitor) must be as wide and as short as possible. This reduces series inductance, and therefore, reduces transient voltage drops from the power plane to the power pin. Thus, reducing the possibility of ground bounce.

- Connect each ground pin or via to the ground plane individually. A daisy chain connection to the ground pins shares the ground path, which increases the return current loop and thus inductance.
- Add the recommended decoupling capacitors for as many VDD/GND pairs as possible.
- Place the decoupling capacitors as close as possible to the power and ground pins of the device.
- Add external buffers at the output of a counter to minimize the loading on silicon device pins.
- Configure the unused I/O pin as an output pin and then drive the output low. This configuration acts as a virtual ground. Connect this low driving output pin to GNDINT and/or the boards ground plane.
- Turn on the slow slew rate logic option when speed is not critical.
- Place the power and ground pins next to each other. The total inductance will be reduced by mutual inductance, since current flows in opposite directions in power and ground pins.
- Use a bigger via size to connect the capacitor pad to the power and ground plane to minimize the inductance in decoupling capacitors.
- Use surface mount capacitors to minimize the lead inductance.
- Each GND pin/via should be connected to the ground plane individually.

24.TIE-HI & TIE-LO

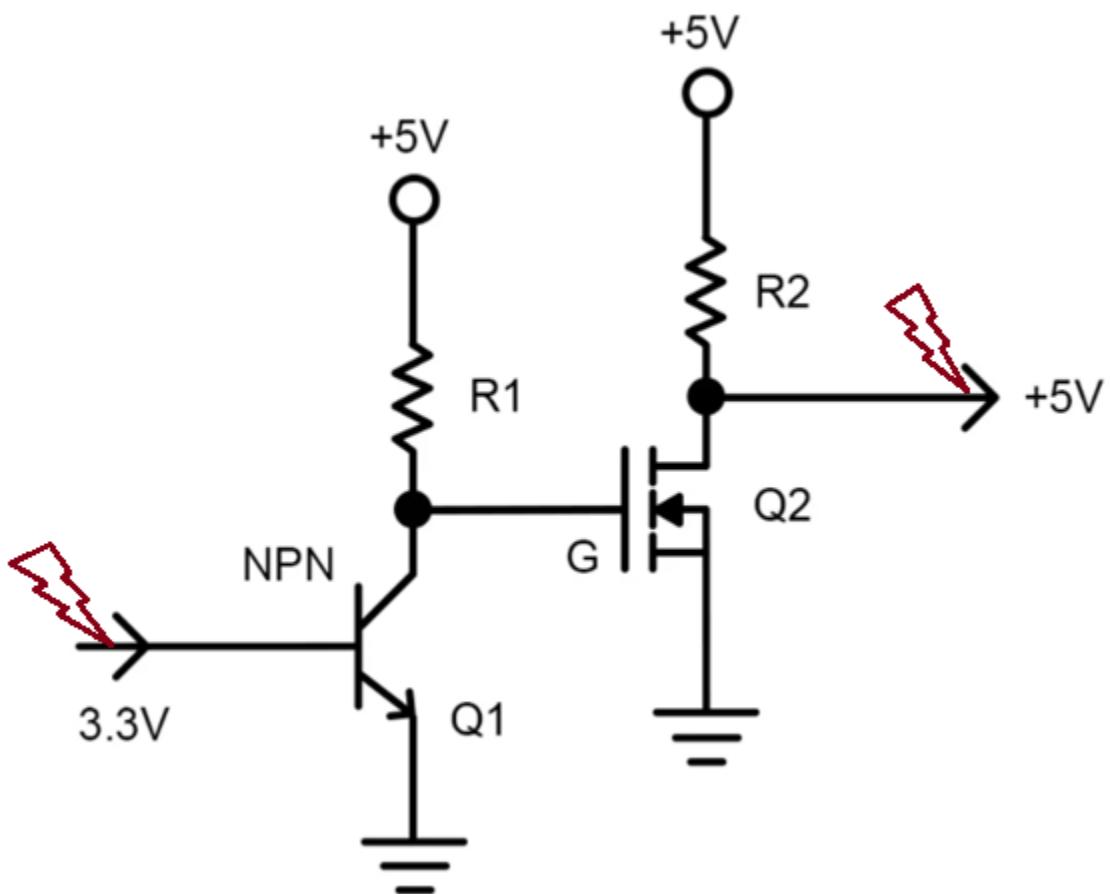
The Ti cell is used for protect the transistor. With its very few connections and a relatively simple circuit, tie-hi cell is found to be one of the most important place in IC design, thus protecting all your logic gates from the spurious disturbances happening at VDD, normally called as 'voltage droop'

The general idea is if you know or intend to provide constant logic '0' or logic '1' to any inputs of your gates (for tie-hi, its logic '1'), then refrain from directly connecting it to VDD, instead add a tie-hi cell in between VDD and gate static input. Let tie-high receive the VDD and provide a clean output of logic '1' to static input of gate. And there's proof, that any disturbance in power or VDD line will just be filtered out at output of tie-hi cell, thus protecting your gate oxide from damage



25.LEVEL SHIFTER

A logic level shifter, or a voltage level translator, is used to translate signals from one logic level (voltage) to another. Nowadays, most of the system runs on 3.3V or 5V. Logic level is simply a HIGH and LOW level of voltage for a certain board or IC. Knowing this, a logic level shifter is necessary to create a path between processors, sensors, or boards of different voltage levels.



26. SERDES

What is a SERDES ?

- SERDES = serializer – deserializer
- Used to transmit high speed IO-data over a serial link in I/O interfaces to speeds upward of 2.5Gbps.
- SerDer TX: transmit parallel data to receiver overhigh speed serial-link.
- SerDes RX: receive data from seriallink and deliver parallel data to next-stage.

SERDES LINKS: In modern times to take advantage of both topologies, often applications involve both parallel and serial communications. A SerDes transceiver is one such application. The main function of the SerDes system is to transmit data at high speeds over a channel and receive the correct data at the receiver end. It facilitates the transmission of data between parallel I/O ports over a serial link by sequentializing data. Thus, the number of channels and pins are reduced. The SerDes becomes especially beneficial for high data rates when the issues with parallel buses are more heightened. For the implementation, CMOS technology proves to be the ideal choice due to its low cost and low power consumption. The SerDes is primarily a mixed-signal circuit with several analog as well as digital blocks to perform different functions.(See Figure 2.1).

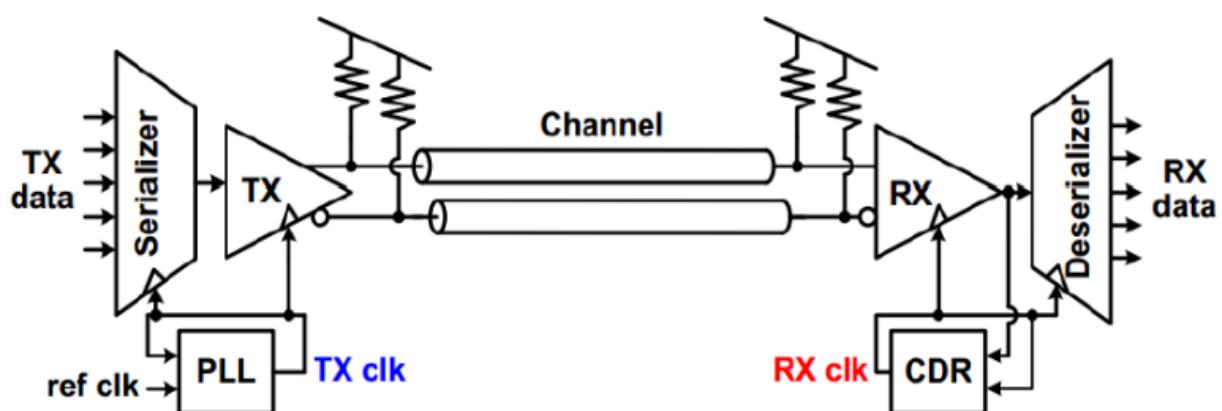


Figure 2.1: The SerDes Link

Advantage: Fast signaling, robust, high signal integrity.

- Low cost: Due to lesser number of pins and lower board space usage. The extra space can be reused for other components.
- Less Crosstalk: Fewer conductors in proximity, thus serial links minimize cross-talk
- Clock skew issues: Parallel links face clock skew issues between parallel channels. This becomes redundant for serial links as they usually use asynchronous and unclocked data transmission.

Clock and Data Recovery (CDR)

Sometimes the data is sent over the channel without an accompanying clock signal. The CDR block performs the role of generating a clock from the received data. It receives the data at an approximately known frequency and then uses phase alignment to align the generated clock to the transitions in the data. This clock and retimed data is then used by the deserializer to recover the data. The aligned clock now allows the data to be sampled at the middle of the eye and hence supports the accurate recovery of data. The middle of the eye is also the most optimum point to take a data sample as it has the least jitter usually.

Phase Locked Loop (PLL)

The Phase-locked-loop forms an essential part of the SerDes circuitry as it generates the high frequency reference clock used to drive the serial transmitter. The reference clock inputs are often required to meet tight electrical and jitter requirements as the quality of the clock can have a significant impact on the links performance. The PLL is primarily used as a frequency synthesizer and is responsible for stepping up the clock frequency by using a divider in the feedback loop.

A PLL in the most basic of terms is a feedback system. It constitutes of a Phase Detector, charge pump loop filter and a voltage controlled oscillator. Figure 3.2 shows a basic phase locked loop.

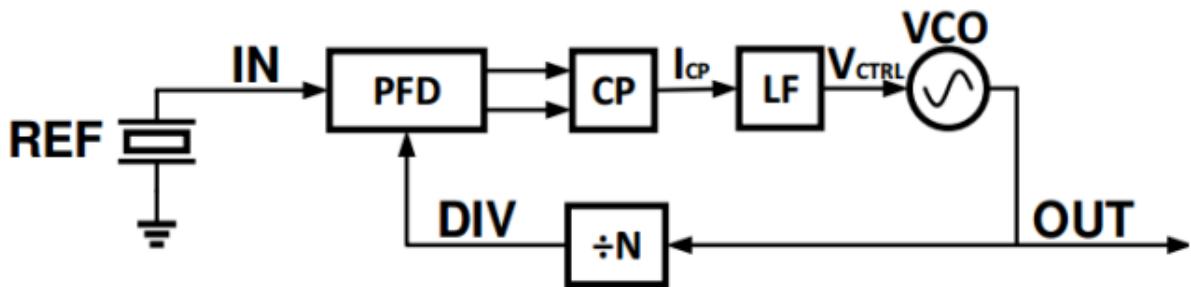


Figure 3.2: A Basic Phase Locked Loop

The PLL is said to achieve lock when the phase difference between the two input signals becomes constant and the corresponding frequencies become equal. The phase frequency detector (PFD) generates an error pulse for a phase difference between the two signals. This error signal is often amplified and converted into an analog signal by a charge pump. The analog output of the pump is passed onto the loop filter which suppresses the high frequencies, enabling the DC component called control voltage. This control voltage is the controlling input of the VCO which determines the oscillation frequency of the VCO. The VCO changes its frequency to accumulate enough phase for the PLL to achieve lock. The VCO output is fed back into the PD for comparison.

27. BUFFER CHIPS / DATA BUFFER

1.

In computer science, a **data buffer** (or just **buffer**) is a region of a physical memory storage used to temporarily store data while it is being moved from one place to another. Typically, the data is stored in a buffer as it is retrieved from an input device (such as a microphone) or just before it is sent to an output device (such as speakers). However, a buffer may be used when moving data between processes within a computer. This is comparable to buffers in telecommunication.

Buffers are typically used when there is a difference between the rate at which data is received and the rate at which it can be processed

2.

In simple terms, a data buffer is an area of physical memory storage that temporarily stores data while it is being moved from one place to another. This becomes increasingly necessary in data centers, autonomous vehicles, and for machine learning applications. The challenges with these applications are advanced signal equalization, and increased capacity and bandwidth. Data buffering techniques — either as a discrete chip within the memory module, or integrated into an SoC.

The data buffer chip was conceived during the development of DDR3. It captures data to and from the CPU and the DRAM. Buffer chips manage read and write commands from the CPU to the DRAM. When buffer chips are used, it is known as a load reduced dual inline memory module (LRDIMM). Those chips boost total memory capacity at the same performance and speed.

Buffers are used for many purposes, including:

- Interconnecting two digital circuits operating at different rates,
- Holding data for later use,
- Allowing timing corrections to be made on a data stream,
- Collecting binary data bits into groups that can then be operated on as a unit,
- Delaying the transit time of a signal in order to allow other operations to occur.

Video link:

1.Serdes: <https://youtu.be/608TPRam3G8>

28.COMPLETE ESD EXPLANATION

ESD Protection Circuit

What is an ESD

- Static electricity is the charge generated on the surface of dielectric materials. Static electricity is discharged when positively and negatively charged objects are brought into contact with or close to each other. This phenomenon is called an electrostatic discharge (ESD). When a charged human body touches an electronic device, the resulting ESD can be several thousand volts.
- Everyone who's ever walked across a carpet then gotten a shock when they touched a doorknob knows exactly what an ESD event is. In an integrated circuit (IC), an ESD event typically induces electrical currents on the order of 0.1–10 amps, which lasts between 10^{-6} and 10^{-3} seconds, and dissipates energy on the order of 10–100 watts. ESD protection methods shunt these ESD currents through unpowered devices (ESD protection devices) along intended ESD discharge paths, while clamping the voltage at a safe level, preventing any functional degradation to the protected devices.

Introduction to Electro Static Discharge (ESD) in ICs

ESD (Electrostatic discharge) failure is one of the critical reliability problems which have been present in the IC industry for a long time. ESD failure accounts for almost 35% of the total IC field failures and cost several billion dollars to the industry annually. There has been an external protection method for ESD failures, but the electronics industry started adopting dedicated on-chip ESD protection methods some time ago. Active research on ESD fundamentals could answer many issues, but still there are grey areas to be addressed. Predictive ESD CAD design is one of the areas of recent interest. Since the development in the IC industry is fast paced, new ESD problems get introduced continuously. The aim of the current work is to predict the optimum distance ("optimum distance" in this document represents the "maximum distance" at which the ESD clamp can be placed in an IO ring with reduced chances of ESD failure) for the placement of ESD

clamps so that chances of an ESD failure is reduced. This helps save time, effort and money. An Electronic Design Automation (EDA) tool “IO Planner” is developed here so that the placement of clamps is done to a level where a chance of IC failure due to ESD events is reduced. Any IC design engineer can use the “IO planner” tool as it does not demand deep ESD knowledge from the engineer.

An ESD event is basically a charge balancing act between two objects at different potentials. It can happen through direct contact or through induced electric field. An ESD strike from different sources is illustrated in Figure 1-1.

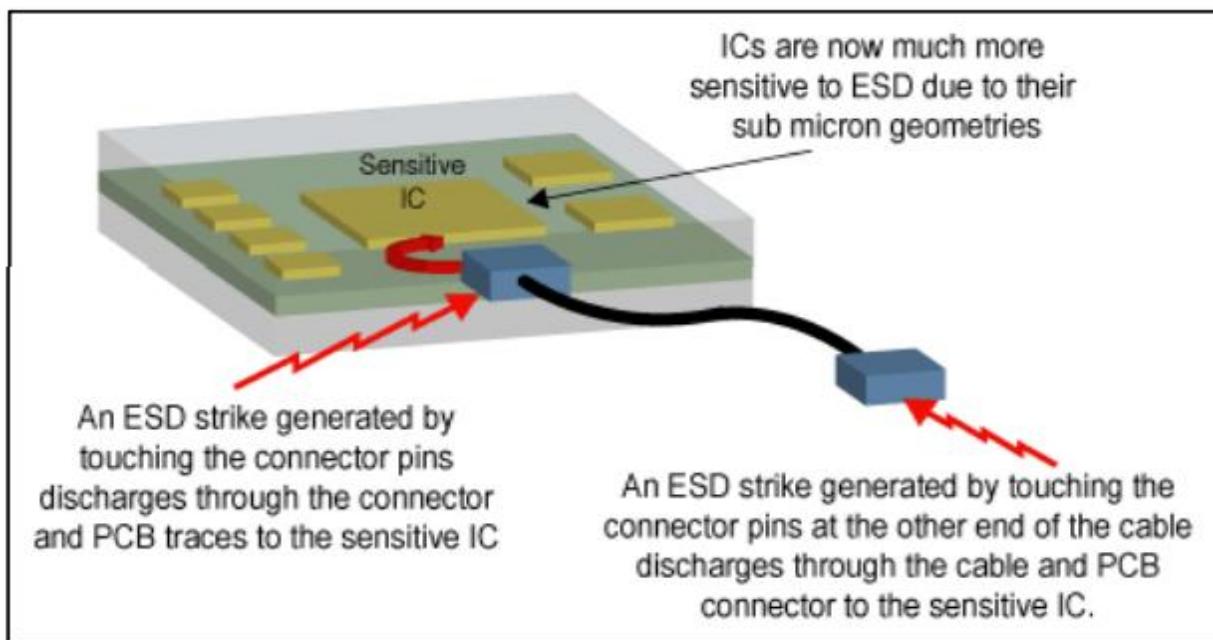


Figure 1-1 ESD strikes from different sources

High voltage or current stress occurring for a very short interval of time is the characteristic of an ESD event. Thus the impact of the event is so high that it can damage the IC. The different damages caused by ESD failures are shown in Figure 1-2. Study of ESD is a field where electrical, thermal and mechanical engineering join hands. The solution to protect ICs from ESD events is to discharge the high current via a low impedance shunting path or clamp the PAD voltage to a sufficiently low level or ground.

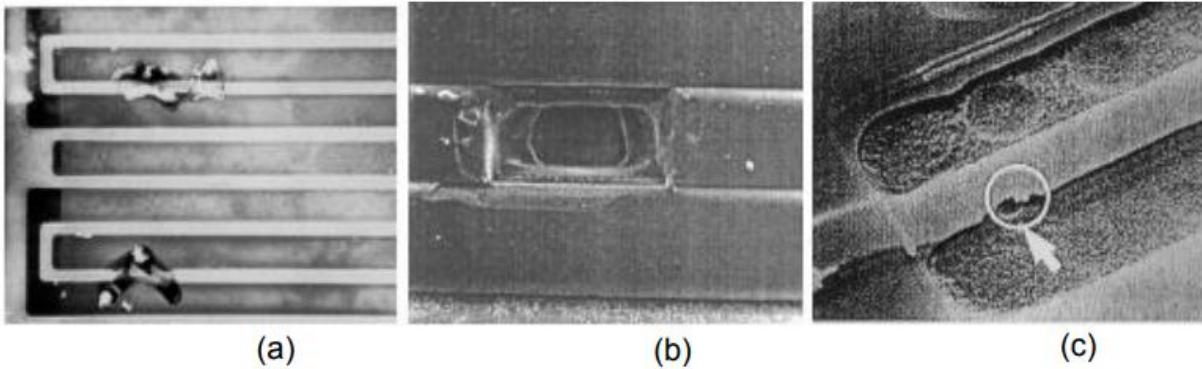


Figure 1-2 Damages caused by ESD failures: (a) junction breakdown, (b) metal/via damage, (c) gate oxide damage

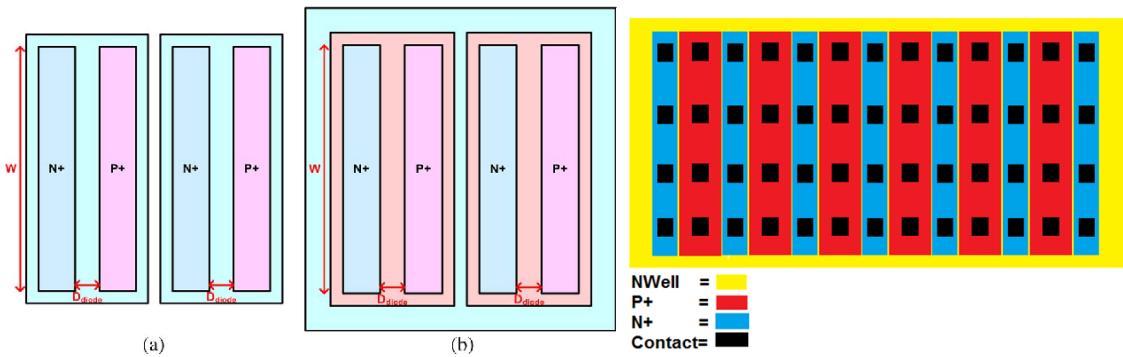
Common ESD Failures:-

ESD failures are disastrous and can lead to instant malfunction of IC chips.

Junctions and oxides are prone to damages. The basic mechanisms for ESD induced failure are:

- 1. Junction burnout in silicon:** This is the most common HBM failure mechanism – the injection of an ESD transient of energy can drive the junction into a breakdown.
- 2. Oxide breakdown:** Another major category of ESD damage arises when a high voltage applied across the oxide layer – high-voltage overstress – causes the dielectric to breakdown. As the dielectric breaks down, it starts conducting current. The heat from the current flow can produce hot spots and melt down the dielectric, silicon, and other materials.
- 3. Metallization burnout:** This produces openings in the interconnection paths. It happens when the temperature – the I^2R heat – reaches the material's melting point. It is commonly a secondary effect, occurring after junction or oxide failure.

ESD Diode:

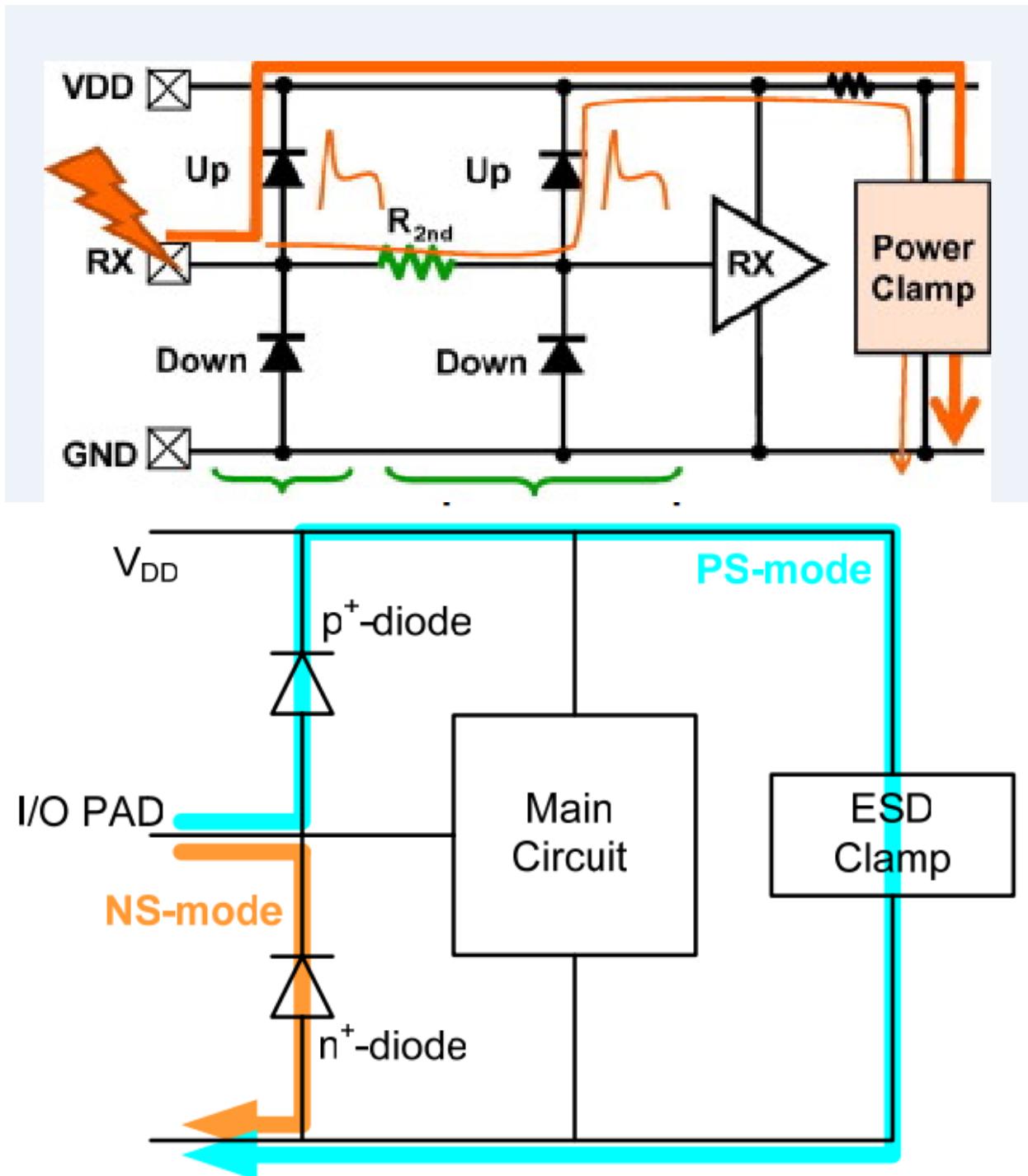


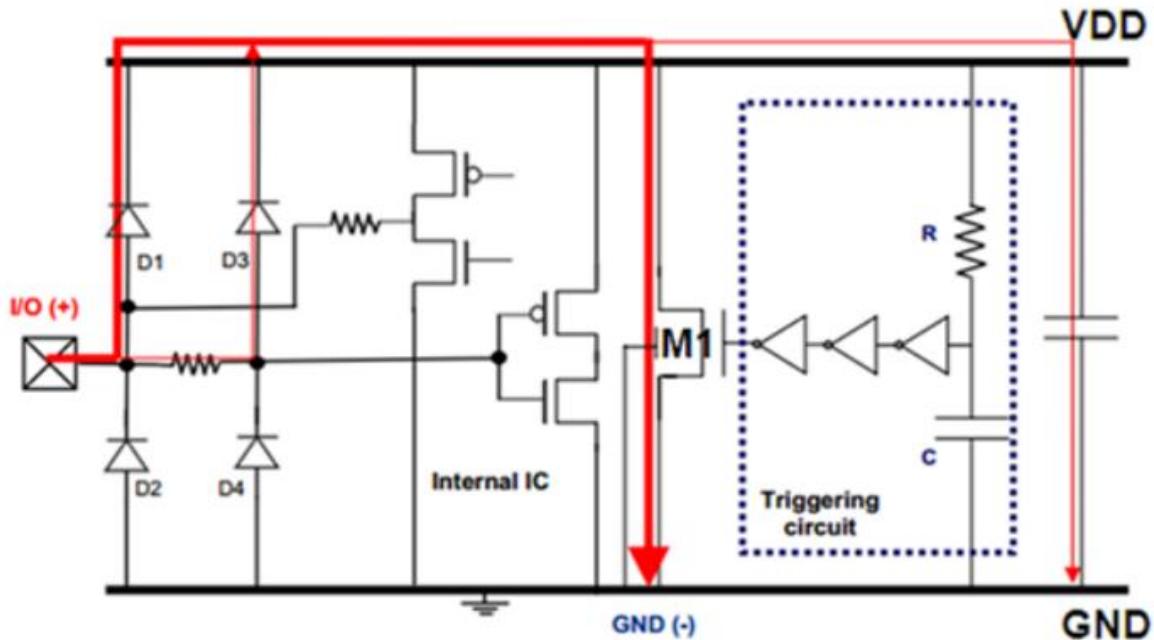
Layout top view of conventional (a) P diode stackup and (b) N diode stackup.

The diode stackup has been used as on-chip electrostatic discharge (ESD) protection for some applications in which the input/output signal swing is higher than VDD or lower than VSS. A novel ESD protection structure of diode stackup is proposed for effective on-chip ESD protection

ESD Protection ckt

There are different on-chip protection methods used for ESD protection. The rail-based ESD protection circuit is one among them. In general, this scheme has a primary network, secondary network and an RC-triggered power clamp



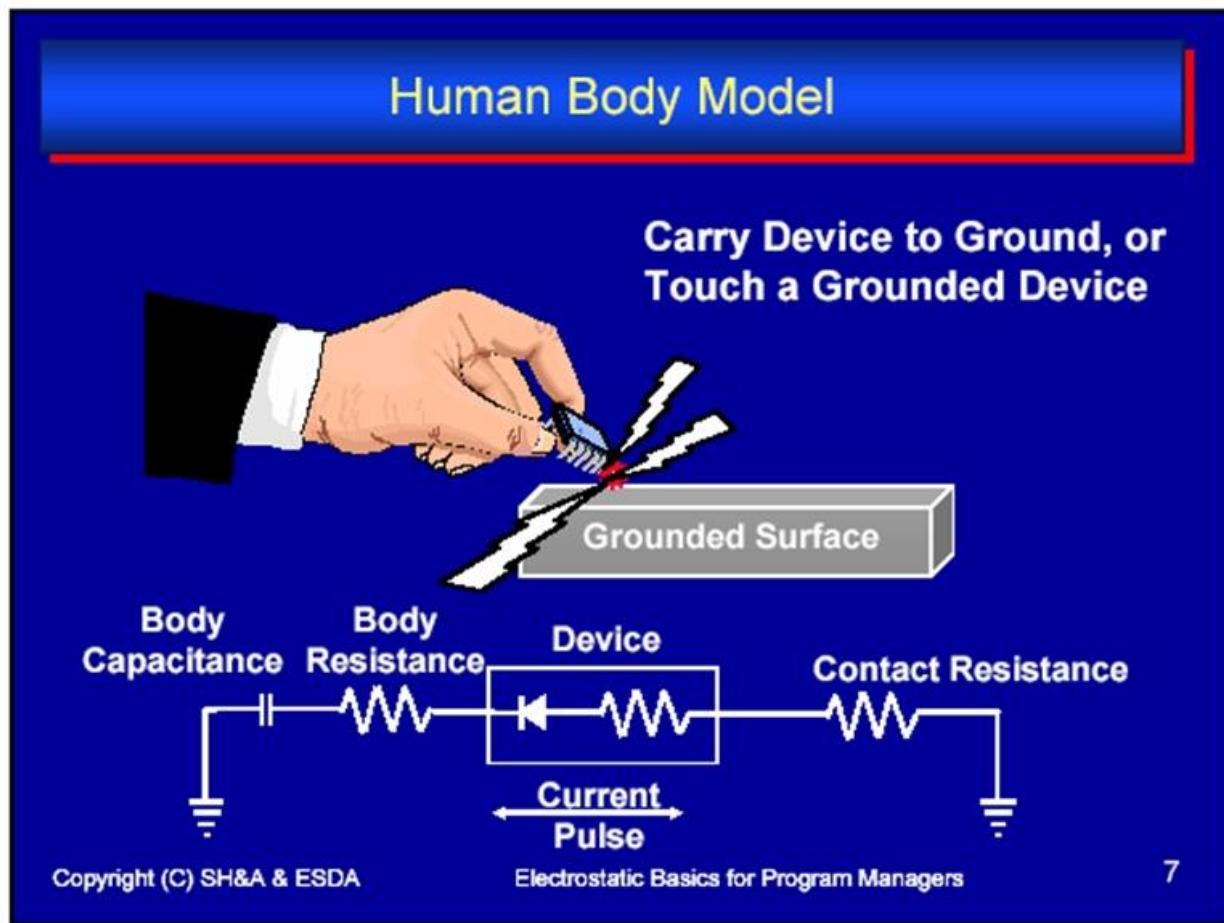


Here, diodes D1, D2 form the primary network and D3, D4 form the secondary network. M1 represents the clamping device which is triggered by the RC arrangement and three series inverters. When the positive ESD pulse is induced on the PAD, there is no direct path from the PAD to GND; diodes D1 and D3 are triggered on and the ESD current flows through them to the power supply line and to GND through the triggered power clamp. Design of each of the involved devices is done such that it is able to handle the ESD current. For a negative ESD pulse, current enters the supply line, flows through the power clamp and through the diodes D2, D4 to the GND line. Placement of the power supply clamps determines the effectiveness of the IO protection in a rail-based ESD protection arrangement. Therefore, it is advantageous if the 51 optimum distances to place power clamps can be estimated so that the probability of failure in silicon is reduced. The work presented here analyzes the ESD network and suggests the optimum distance to place power clamps.

ESD Test Models

Different test models, categorized by their origin, are used to simulate different ESD events upon which ESD protection circuits are tested and qualified. The models are the Human Body Model (**HBM**)/primary network, and the Charged Device Model (**CDM**)/secondary network.

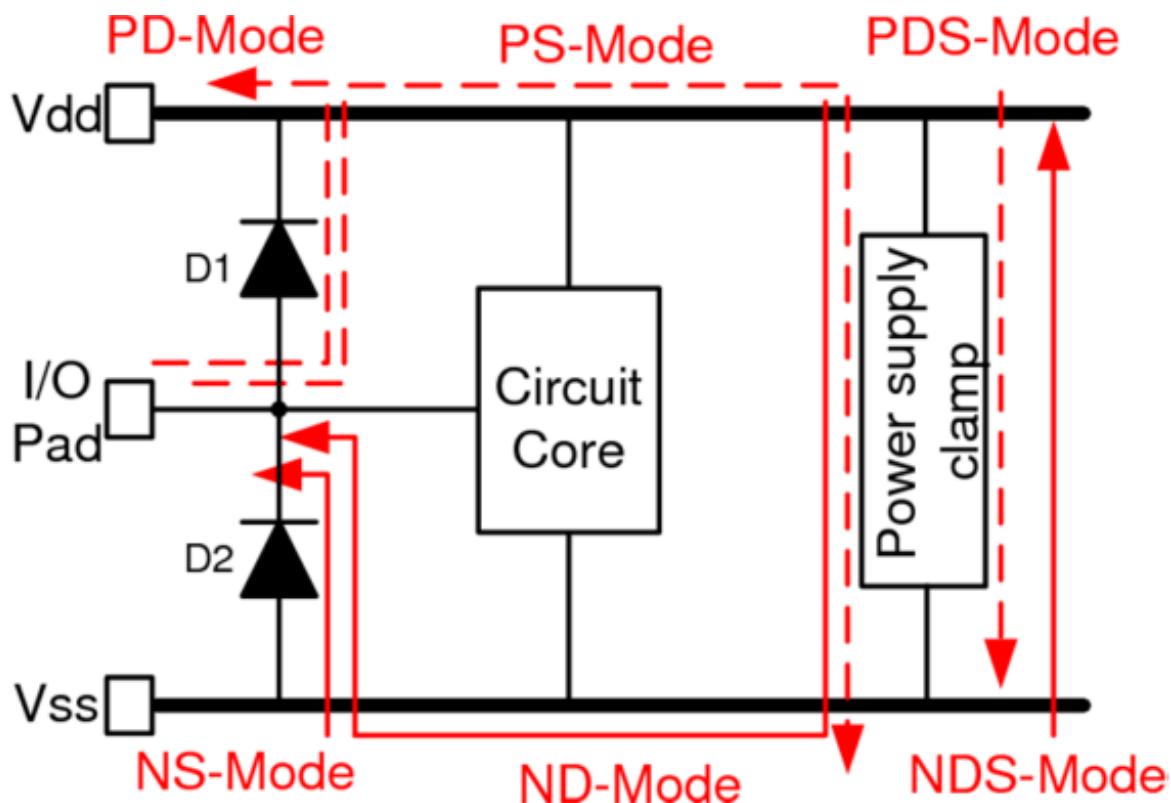
Primary ESD / HBM



The HBM represents the ESD event which occurs when a charged human body or a charged material touches the electronic circuit. Charge transfer happens through physical contact.

HMB Characteristics

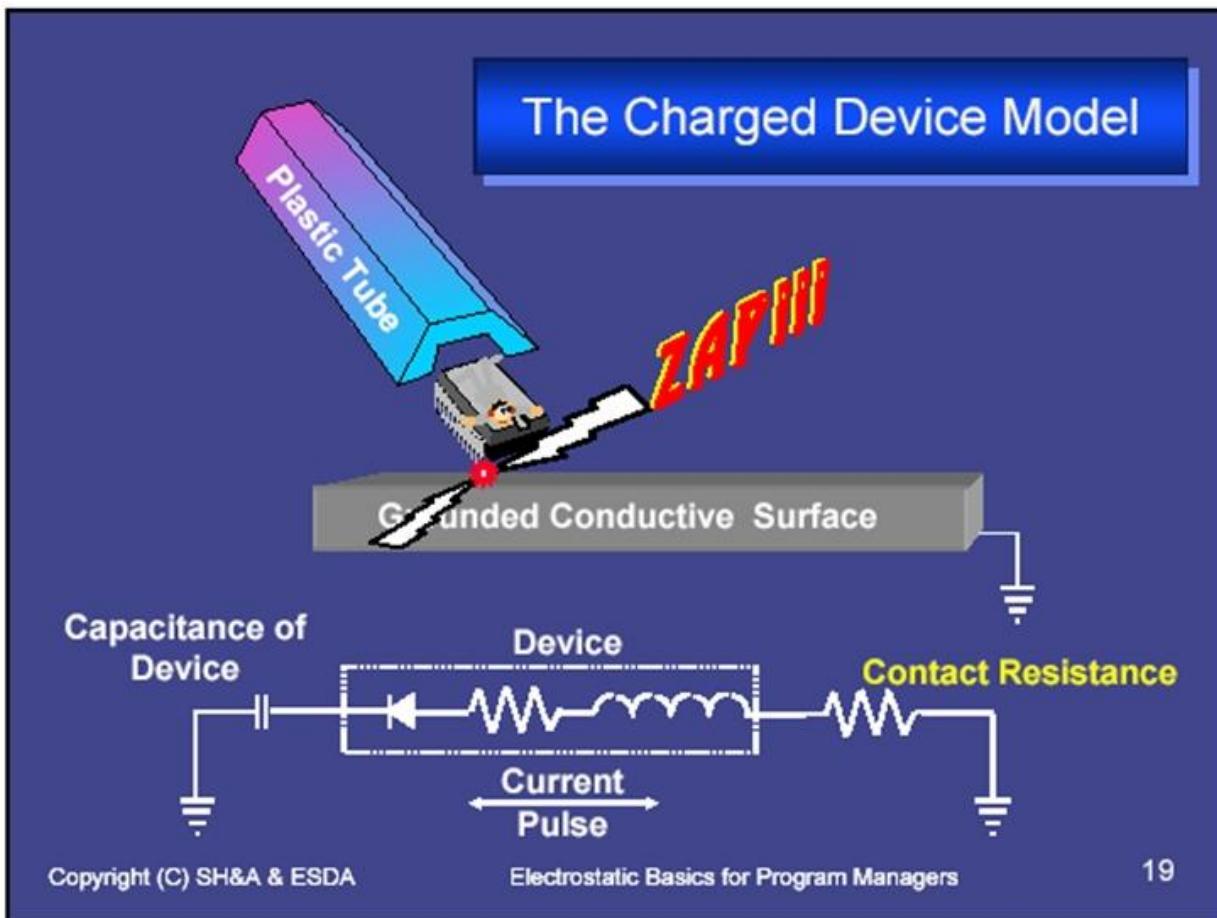
- HBM models the ESD of a human body.
- Qualification level $\approx 2\text{kV}$
- Pulse width $\approx 150\text{ns}$
- Peak current $\approx 1.3\text{A}$
- rise time $\approx 5\text{-}30\text{ns}$



The main goal of ESD protection circuits is to provide a low-resistive discharge path between any two pins in the chip (Vdd/Vss). The ESD clamp provides the discharge path for an ESD event that happens between the two power rails (Positive stress at VDD node while VSS node is grounded (PDS-mode), Positive stress at VSS node while VDD node is grounded (NDS-mode)).

- Primary ESD protection: as close as possible to pad.

Secondary ESD / CDM

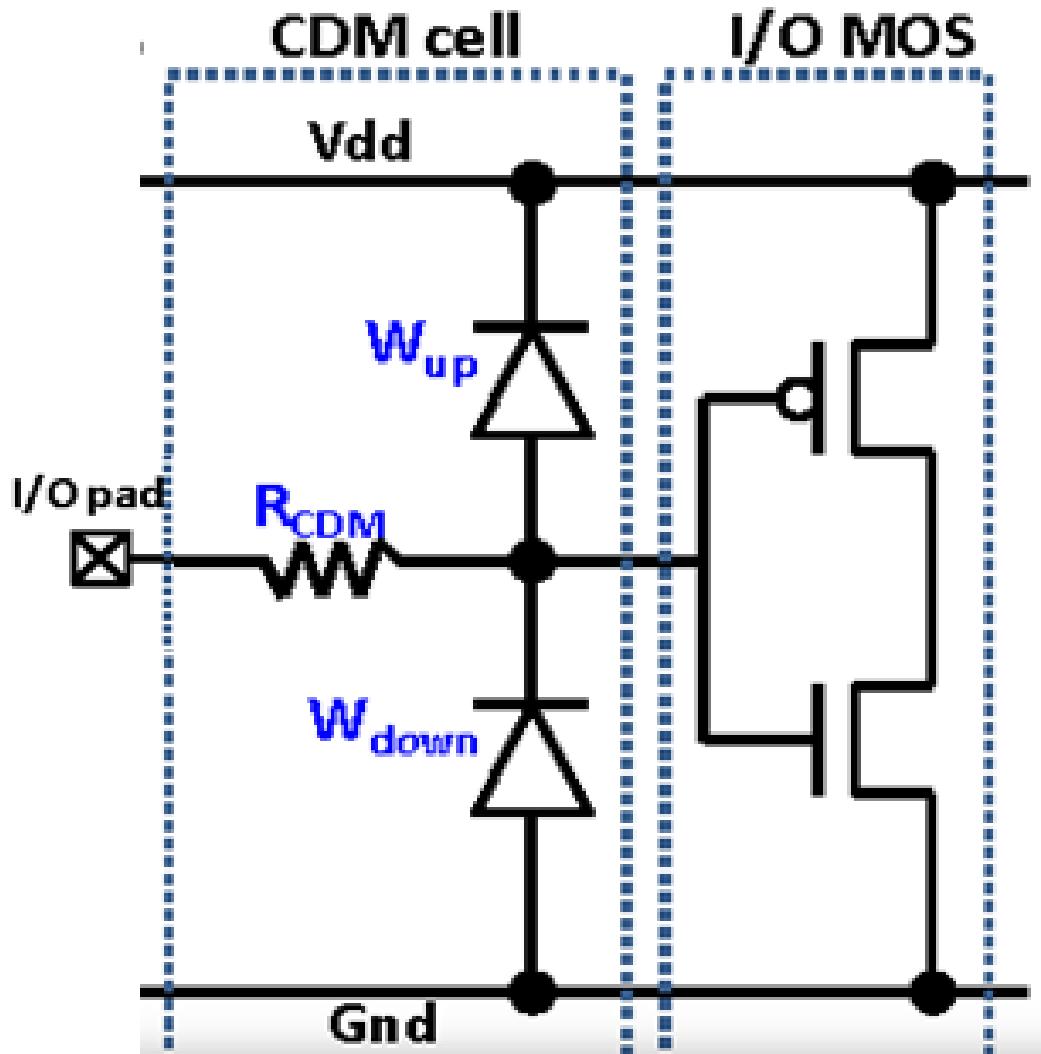


The charged device model simulates the event where the un-grounded electronic parts accumulate charge during manufacturing or assembly and then discharge to ground. Unlike the HBM and MM event, CDM event involves a single pin on the module.

- CDM models the ESD of charged integrated circuits.
- Inductance in the model is mainly due to the inductance of bond wires.
- Peak current $\approx 10A$, rise time $\approx 1\text{ns}$.
- Gate oxide breakdown is the signature failure of CDM stress, in contrast to the thermal failure signature of HBM stress.
- CDM stress has the fastest transient and has the max. peak current.
- CDM stress is the most difficult ESD stress to protect against.

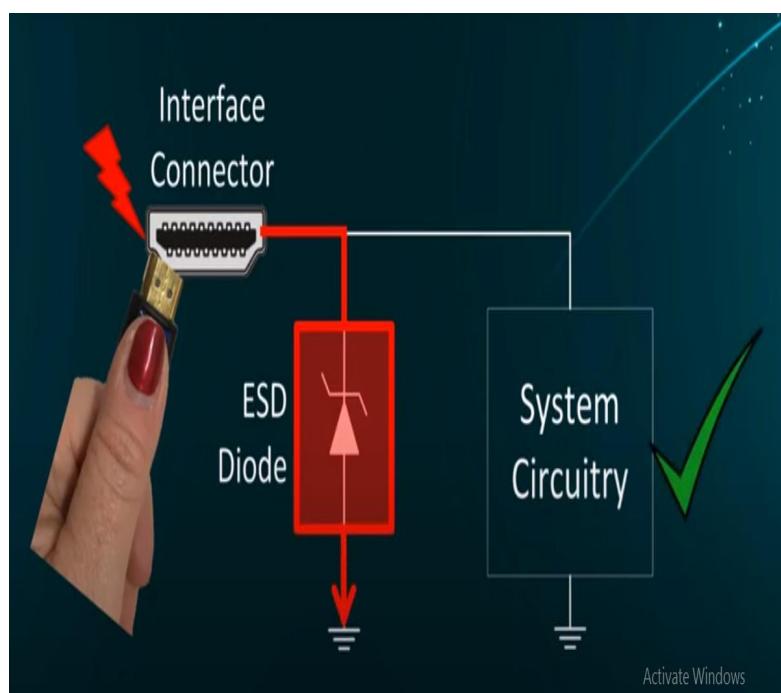
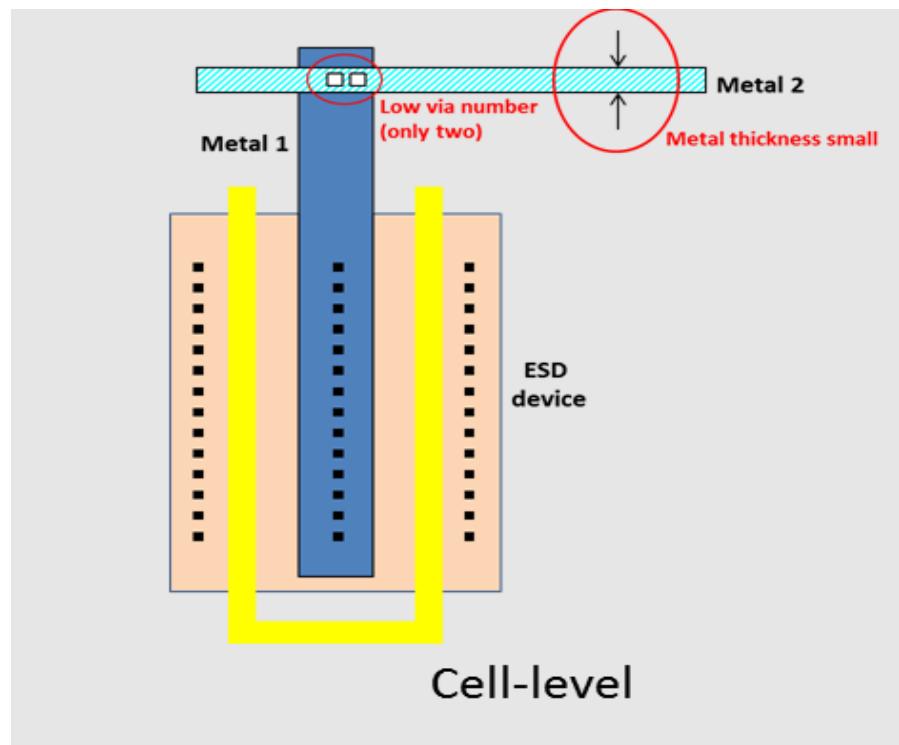
CDM Characteristics:

- Qualification Level $\approx 500V$
- Pulse Width $\sim 1.5\text{ns}$
- Rise Time $\approx 100-300\text{ps}$
- Peak Current $\approx 1-15\text{A}$



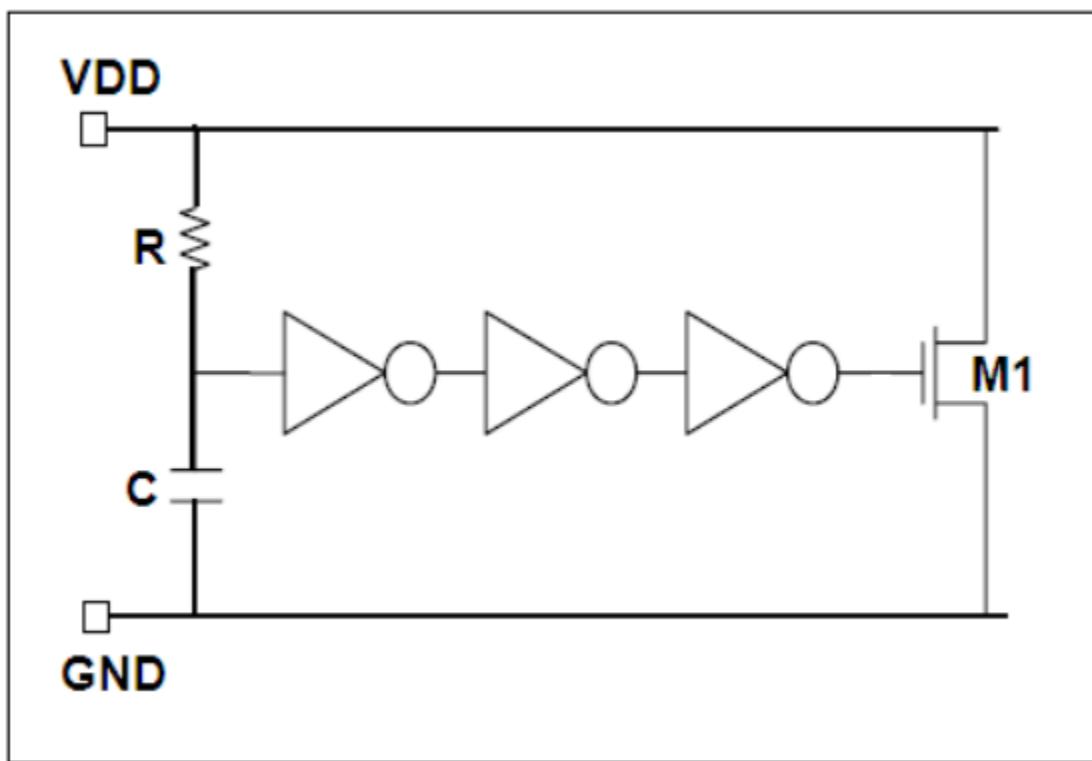
ESD to Device connection

Resistance b/w BUMP to diode connection should be less than resistance b/w BUMP to device connection



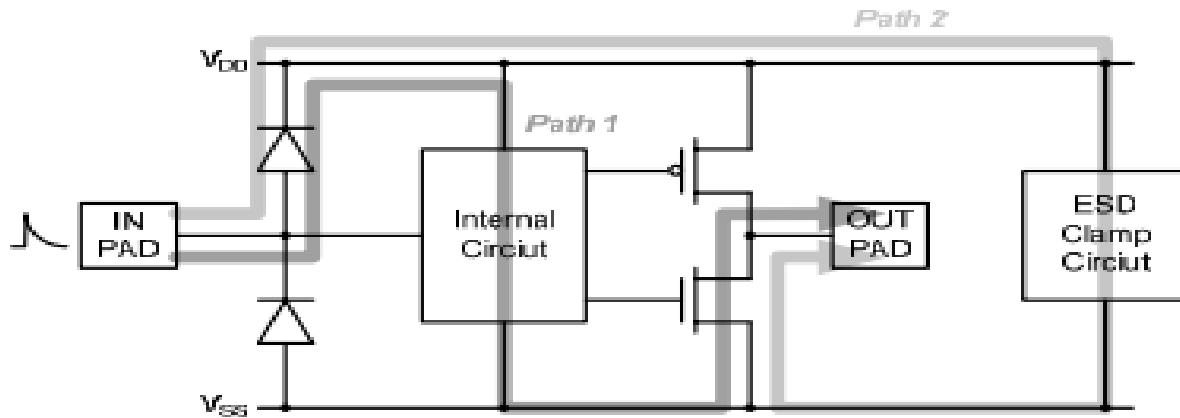
POWER CLAMP

Power clamps are present inside each power supply IO. They are used to shunt current to ground from power supply lines when an ESD event occurs or any event which can potentially damage the electronic components. At least, a chip will have three such pins namely, VDDIO (IO-level power supply), VDDC (core-level power supply) and VGND (ground for both supplies). The RC-triggered power clamp is a simple and efficient implementation to achieve the same. One such typical RC-triggered power clamp is given in Figure



Design of the RC network and the 3-stage inverter circuit is done according to the trigger requirements for the clamp. The width of M1 (size of the transistor) is based on the maximum value of current it will have to shunt.

A power IO functions to provide a power supply for the chip. Power clamps which are part of the IOs, help reduce risks due to ESD events for the chip. A rule of thumb is to have as much power IOs, especially for the core power supply, as possible so that the chip is supplied with stable and robust power.

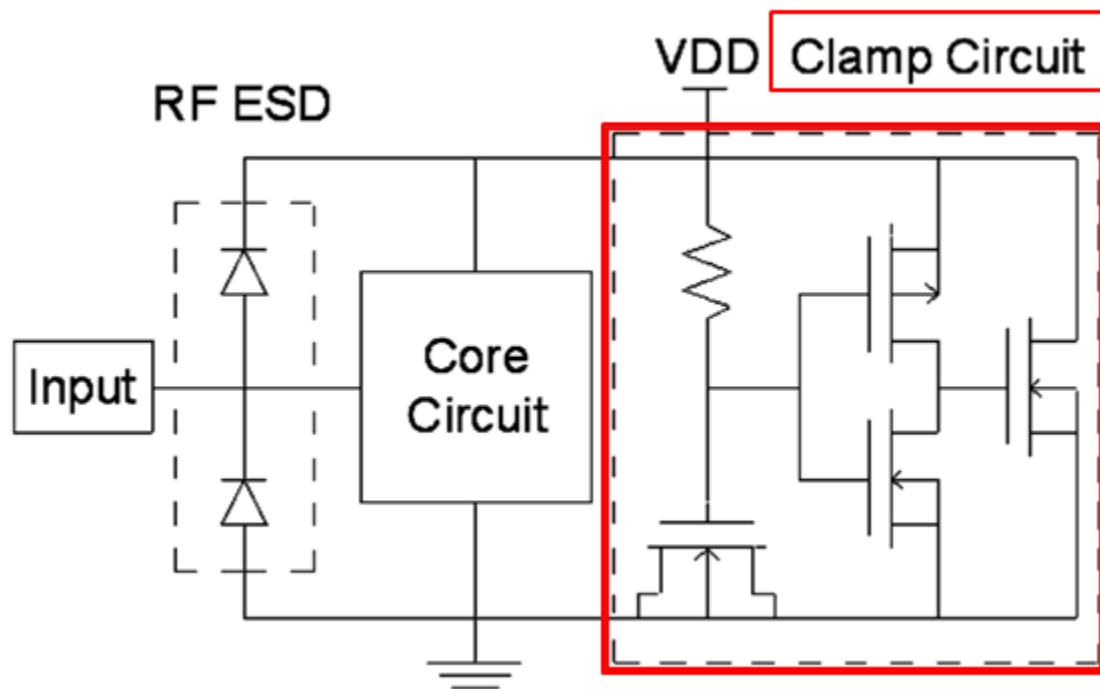


ESD Clamps can be grouped into two categories: static and transient.

The static clamps provide a static or steady-state current and voltage response. A fixed voltage level activates static clamps. As long as the voltage is above this level, the clamp will conduct current. A diode, MOSFET and SCR based clamps are known as static ESD clamps.

Transient clamps take advantage of the rapid changes in voltage and/or current that accompanies an ESD event. During this transient, an element is turned on very quickly and slowly turns off. This type of clamp conducts for a fixed time when it is triggered. An RC network determines the time constant. These clamps are typically triggered by very fast events on the supply lines.

RC-triggered MOSFET



ELECTROSTATIC DISCHARGE (ESD) protection has become an important task in the pursuit of the increased reliability of semiconductor ICs. Furthermore, while the required ESD protection level is HBM 2 kV in most cases, in the Smart Power IC market the customer requirements are more demanding because Smart Power ICs are used in harsher environment than general logic IC. In smart power technology, high-voltage MOSFETs have been used as on-chip ESD protection devices. Among them, the lateral diffused MOS (LDMOS) and the drain extended MOS (DEMOS) have been widely used as common ESD protection device because it can function both as output driver and as ESD protection

LATCHUP OR LATCHUP-LIKE FAILURE IN ESD POWER CLAMP:

When high voltage MOSFETs such as LDMOS (lateral double diffused MOS) are used in ESD protection clamps, they are used in “gate-grounded (gg)” configuration. However, as stated in the Introduction, these ggLDMOSs have very low snapback holding voltages. The TLP-measured I-V characteristic of a ggLDMOS. In the TLP measurements, pulses having 100 ns duration and 10 ns rise/fall time were used. theggLDMOS snaps back to only 9 V after the triggering. This strong snapback characteristic of the LDMOS is related to the turn-on behavior of the parasitic bipolar transistor and the occurrence of the Kirk effect (base push-out effect) .

When a ggLDMOS is used as I/O clamps, a low holding voltage is helpful to sustain a much higher ESD current because it leads to the smaller power dissipation. However, when these are used in power clamps, a low holding voltage poses a threat. The power clamps may be triggered on by noise transients or by a glitch on the power lines during normal circuit operations. If the holding voltage of ESD power clamp devices is smaller than the power supply voltage, the power clamp may not be turned off after being accidentally turned on. This may cause very serious “latchup” or “latchup-like” failures and may lead to IC malfunction or even destruction.

This latchup related danger can be avoided if the power supply clamp is designed in such a way that the devices in the clamp do not enter the snapback region. No-snapback LDMOSs are one of the ways to achieve this and active research is going on in this area. However, the design of a new such device usually involves process modification, which is very expensive. Therefore, in this work, we developed a circuit-level solution in which the device in the power clamp does not enter snap-back region.

Principle of Operation

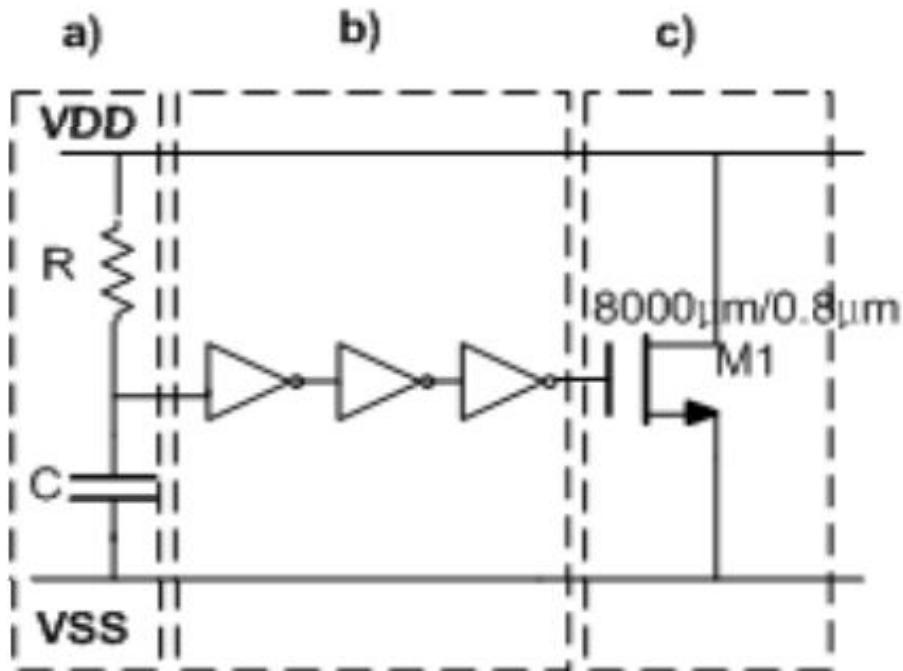


Fig2 :Schematic of the conventional RC-triggered MOSFET ESD power clamps

a: Rise Time Detector / RC- frequency discrimination circuit

b: Inverter drive circuit

c: MOSFET big clamp

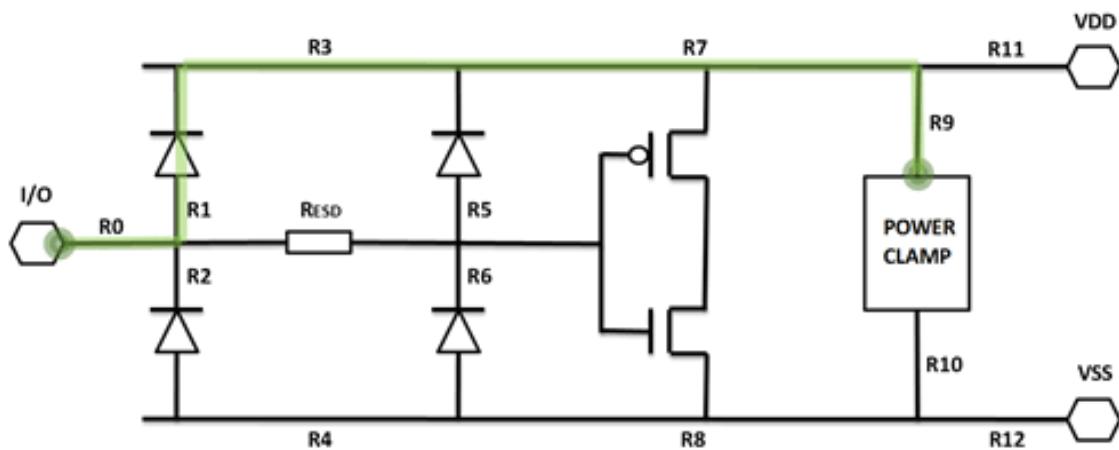
We propose the use of “Darlington scheme” to overcome the latchup or latchup-like issues in the design of power clamps for Smart Power ICs. Darlington scheme is an enhanced RC-triggered MOSFET ESD power clamp. Therefore, we briefly review the basic mechanism of RC-triggered MOSFET ESD power clamps first, and then we move on to the explanation of the proposed power clamp.

Fig. 2 shows a schematic of conventional RC-triggered MOSFET ESD power clamps. It consists of three elements: a) RC-frequency discrimination circuit, b) inverter drives circuit, and c) big MOSFET clamp. Under an ESD transient event, the RC-frequency

discrimination circuit detects a short ESD pulse, and drives the gate of MOSFET into the active turn-on mode, forming a low-impedance discharge path to shunt ESD current. Because of the raised gate voltage, the MOSFET operates below the triggering and does not enter into the snapback region. Therefore, it can reduce the latchup or latchup like dangers. The draw-back of the RC-triggered power clamps is the large area requirement. In the RC-triggered MOSFET clamps, ESD current flows only through the channel region of the MOSFET because of the absence of the parasitic bipolar action associated with the snapback. Therefore, a very large MOSFET is required. Furthermore, the RC-frequency discrimination circuit and the inverter drive circuit occupy additional considerable silicon area.

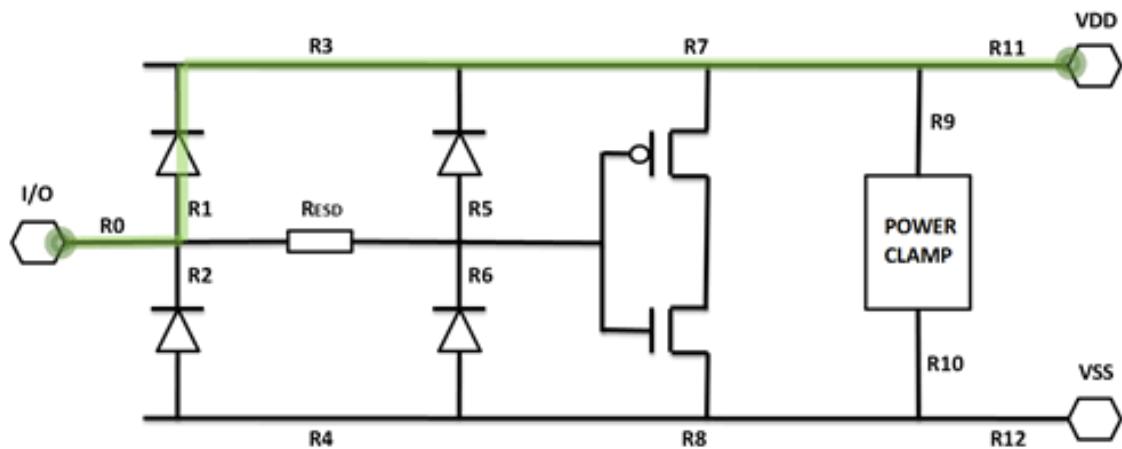
To alleviate this area budgeting problem, we propose the Darlington scheme in Smart Power ICs. By using Darlington configuration as the drive circuit, the silicon area of the discrimination/drive circuit of the power clamp could be reduced by a factor of two when compared to that of the conventional RC-triggered circuits.

Fig. 3 shows the schematic of the proposed ESD power clamp circuit using the Darlington scheme. It consists of three parts, similar to the simple RC-triggered power clamp. However, the inverter drive circuit in the conventional circuit was replaced by a small LDMOS (M2) and a resistor. In the event of an ESD, M2 is turned on from the gate coupling effect, and a current flows through the resistor R1. This current raises the potential at node “n1” turning on the big LDMOS (M1). Because most of the ESD current flows through M1, M2 does not need to be large. Therefore the inverter stages like the one in Fig. 2 are not necessary between the RC network and M2. Therefore, the triggering section of the proposed circuit can be made much smaller than that of the conventional circuit



(a)

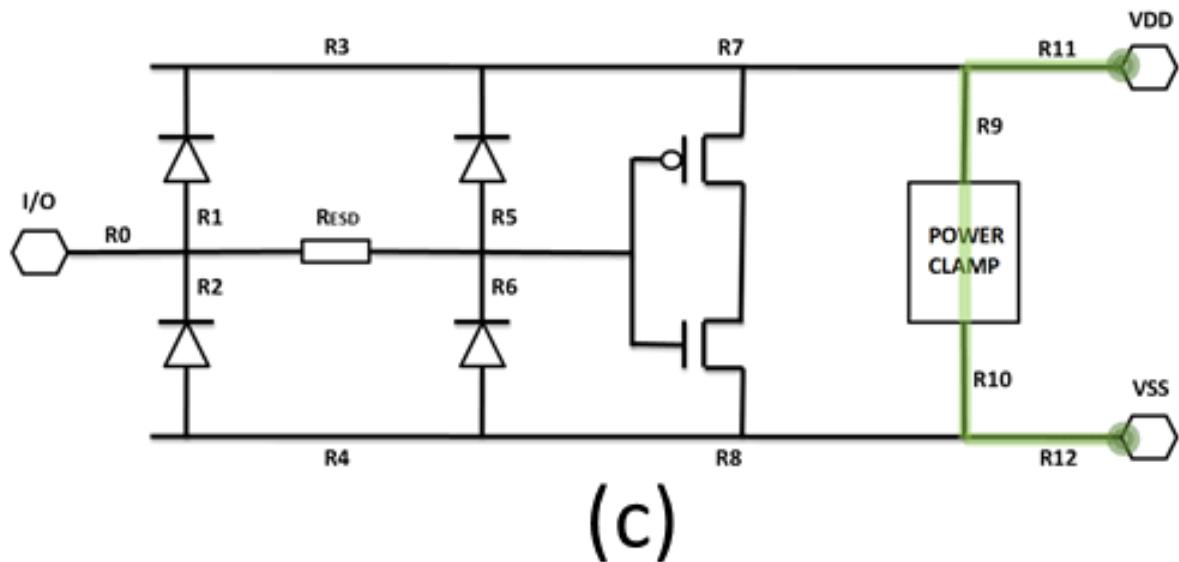
(a): The overflow of electric charge shunts through the pull-up diode, and dissipates through the power clamp. For the pull-up diode to effectively shunt the ESD currents, the total resistance value of $R_0 + R_1 + R_3 + R_7 + R_9$ must be within a maximum allowed limit.



(b)

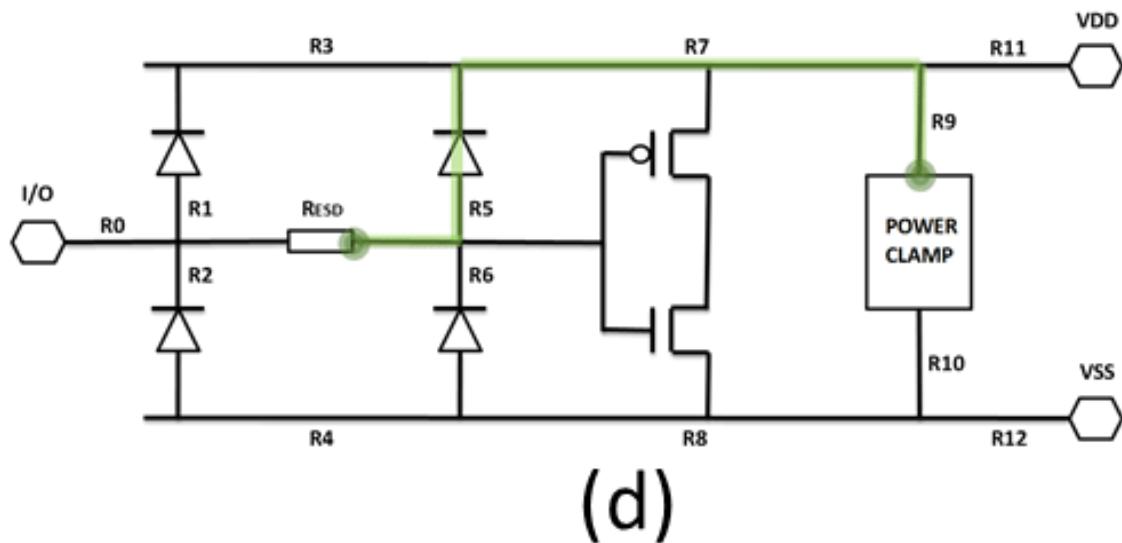
(b): The

overflow of electric charge shunts through the pull-up diode, and dissipates through the VDD power source. The total resistance value of $R_0 + R_1 + R_3 + R_7 + R_{11}$ must be within a maximum allowed limit.



(c)

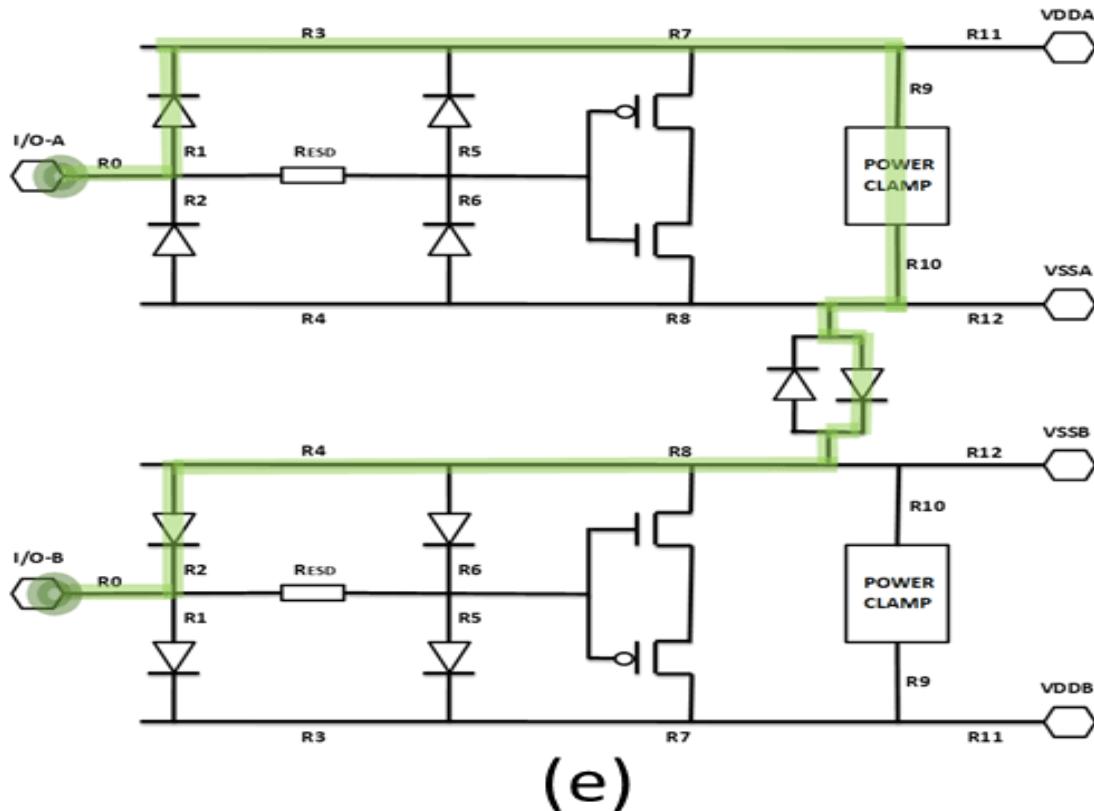
(c): The possible ESD discharge path runs between power pad VDD and ground pad VSS, which goes through the power clamp. The total resistance value of $R_{11}+R_9+R_{10}+R_{12}$ must be within a maximum allowed limit.



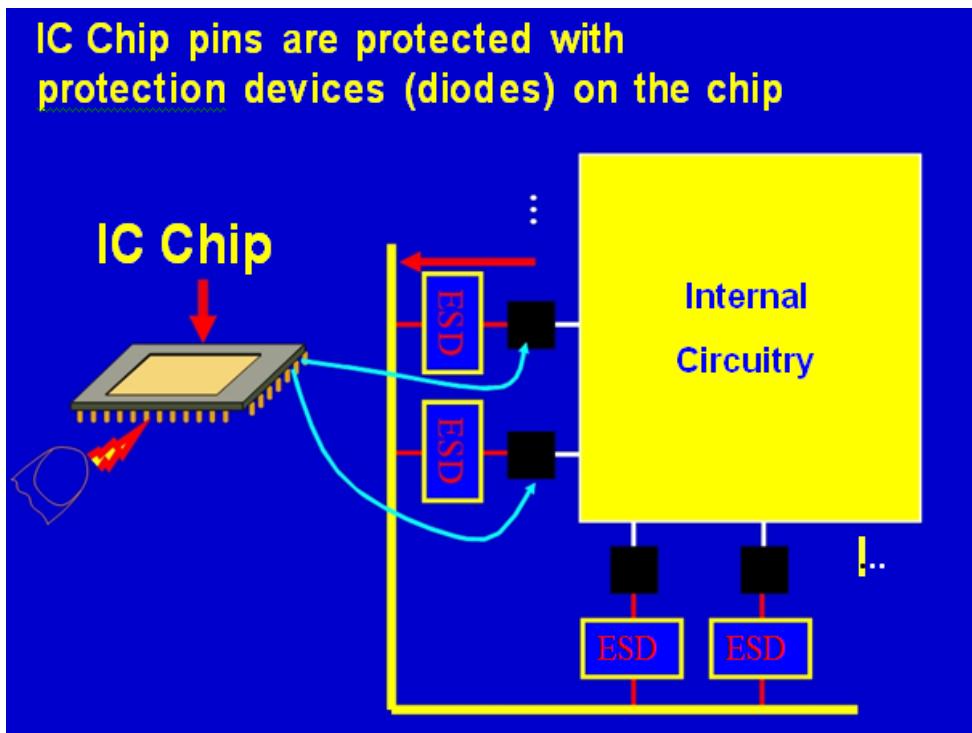
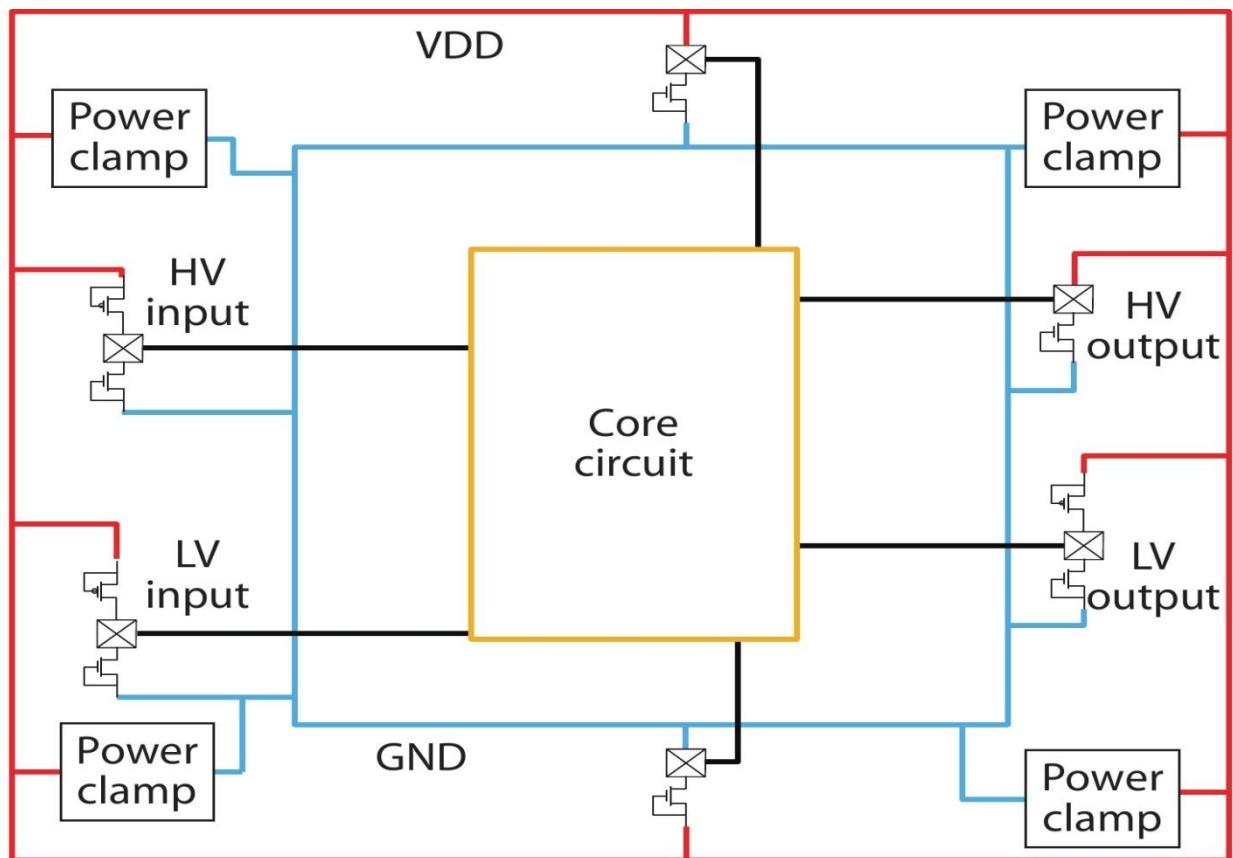
(d)

(d): The possible ESD discharge path starts from the ESD resistor, passes through secondary ESD protection (i.e., a pull-up diode), and ends at the power clamp. The total resistance value of $R_5+R_7+R_9$ must be within a maximum allowed limit.

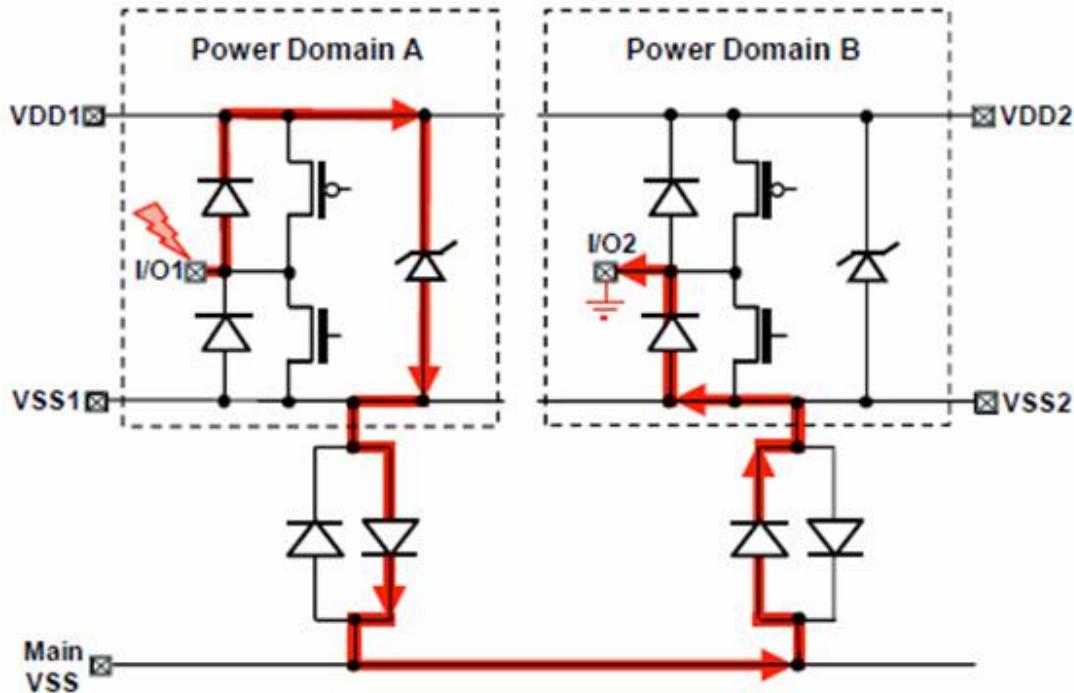
In (a)-(d), all metal interconnects and vias along the ESD discharge path must be robust enough to withstand the ESD currents during the ESD event.



(e): The possible ESD discharge path starts at I/O pad A, goes through a pull-up diode, a power clamp, the B2B diode, the ground bus VSS_B , a pull-down diode, and eventually ends at I/O pad B. Not only must all metal interconnects and vias along this ESD discharge path be robust enough to withstand the ESD currents, but their effective total resistance value must be within a maximum allowed limit.



B2B diode



Across multiple power domains or at full-chip level

. These elements are combined to form the back-to-back diode. This structure is appropriate to external, first stage, protection for high voltage ICs. Combining in the same component, high capacitors for filtering or decoupling, and efficient ESD protection, is a great challenge in system-level consideration.

Layout

In the circuit layout it is extremely important to have good floor planning and power planning in place as well as a clean layout that passes DRC (Design Rule Check) and LVS (Layout Versus Schematic) checks.

- Metallization is done up to Higher for drivers so that it has the ability to handle the current even during a malfunction without breaking down.
- Drain routings for drivers extend as fingers and connect to the PAD which ensures that there is no break down due to inability to handle current.
- Substrate contacts and guard rings are provided so that chances of latch-up are reduced.
- ESD diodes are laid out as it is in the standard library.
- Decaps are added wherever possible, to help stabilize the power supply.
- Lower metal are used for signal routing.
- Higher metal are used for power routing and also metals are staked ; care is taken to make the power plan as robust as possible by having wider metals and maximum number of vias.
- The layout has no LVS or DRC errors.

ESD Network Analysis

An ESD network analysis is done here in order to suggest the optimum distance for power clamp placement. The aim of the approach is to find out how much margin is available for the voltage drop at a power bus. According to the available margin, the optimum distance at which a power clamp can be placed is suggested. Components of an ESD network are shown in Figure 4-1. Only the parasitic resistance is considered for this analysis since parasitic inductance is very small in GPIOs and the parasitic capacitance value which may exist would actually take up some current and it can only reduce the chances of ESD related failure.

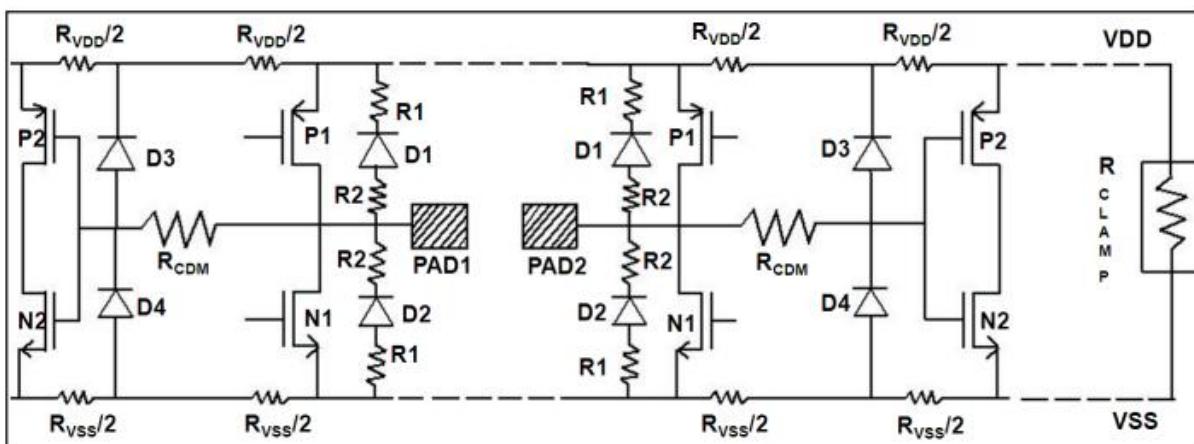


Figure 4-1 ESD network

D1, D2 :- Primary ESD protection network. D1 is referred to the HBM up-diode and D2 refers to the HBM down-diode in the following sessions of the report.

D3, D4 :- Secondary ESD protection network. D3 is referred to the CDM up-diode and D4 refers to the CDM down-diode in the following sessions of the report.

R_{CDM} :- CDM resistance

R_{CLAMP} :- Clamp resistance

R_{VDD} :- VDD bus resistance

R_{VSS} :- VSS bus resistance

R1 :- Diode to supply parasitic resistance

R2 :- Diode to PAD parasitic resistance

P1 :- PMOS driver

N1 :- NMOS driver

P2, N2 :- Receiver transistors

Two failure mechanisms for transistors are considered here. They are junction breakdown (source-drain junction) and gate-oxide breakdown during both an HBM event and a CDM event, for both PMOS and NMOS transistors associated with it. Hence, there are three variables available and there can be up to eight unique cases to be considered to ensure that the IC does not fail due to an ESD event. When there is an HBM event, current flow is from PAD to PAD, whereas for a CDM event, it is from PAD to VSS. Kirchhoff's voltage law is applied for the loop related to each case and the maximum possible value for RVDD and RVSS is calculated. By knowing the resistance per μm for power buses, optimum distance to place power clamps is calculated and suggested.

Figure 4-2 shows the current flow direction for an HBM event for which a junction breakdown parameter for NMOS driver (N1 of IO-1) is evaluated. The HBM pulse originates from PAD1 and flows to PAD2 through R2, D1, R1, RVDD, RCLAMP, RVSS and D2.

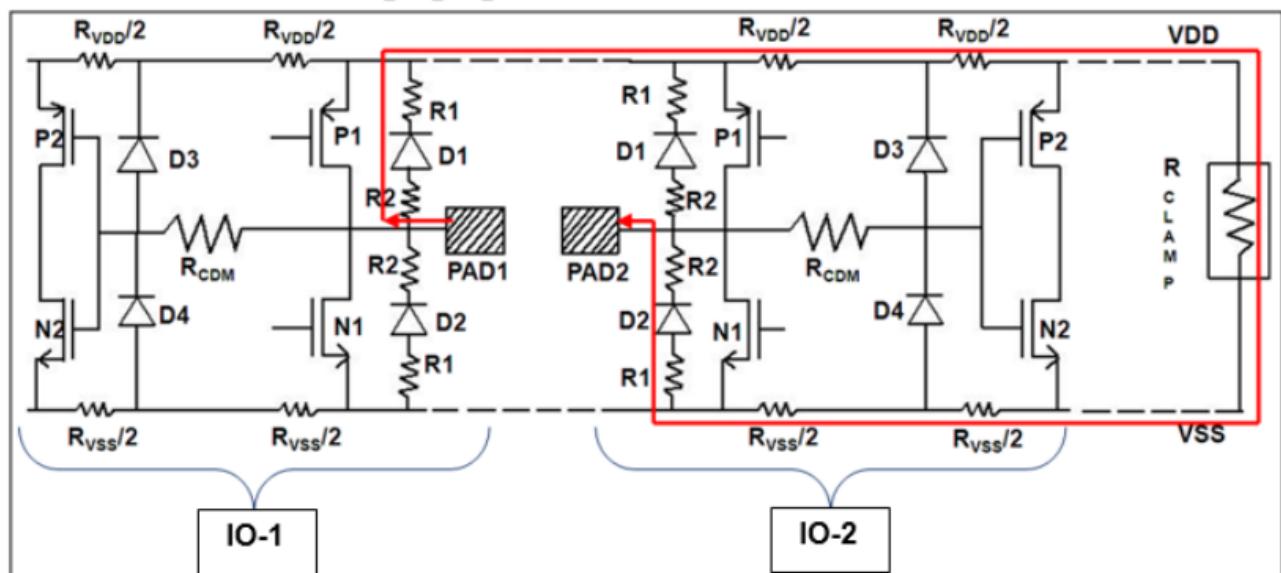


Figure 4-2 Current flow through ESD network during HBM event

Figure 4-3 indicates the current flow direction for a CDM event for which the gate-oxide breakdown parameter for the PMOS transistor (P2 of IO-1) is evaluated. The CDM pulse originates from VDD and a major portion flows to PAD1 through RCLAMP, RVSS, R1, D2, R2. A very small amount of current ($\sim 0.1\%$) takes the path through D4, RCDM from RVSS.

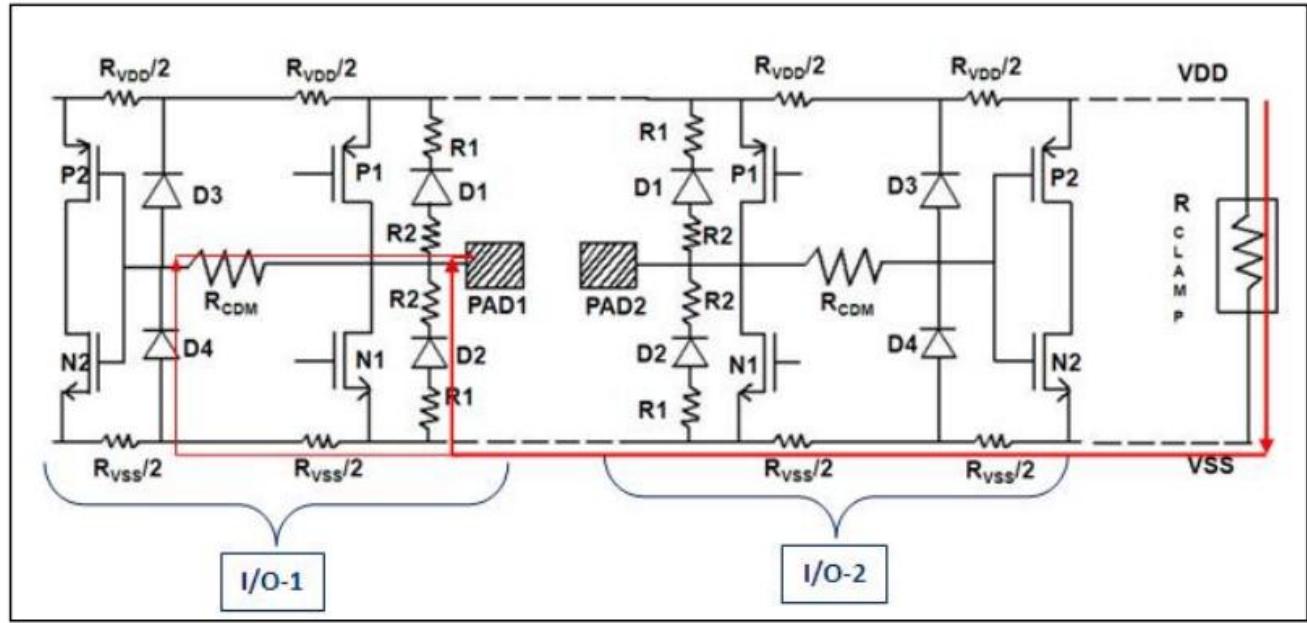
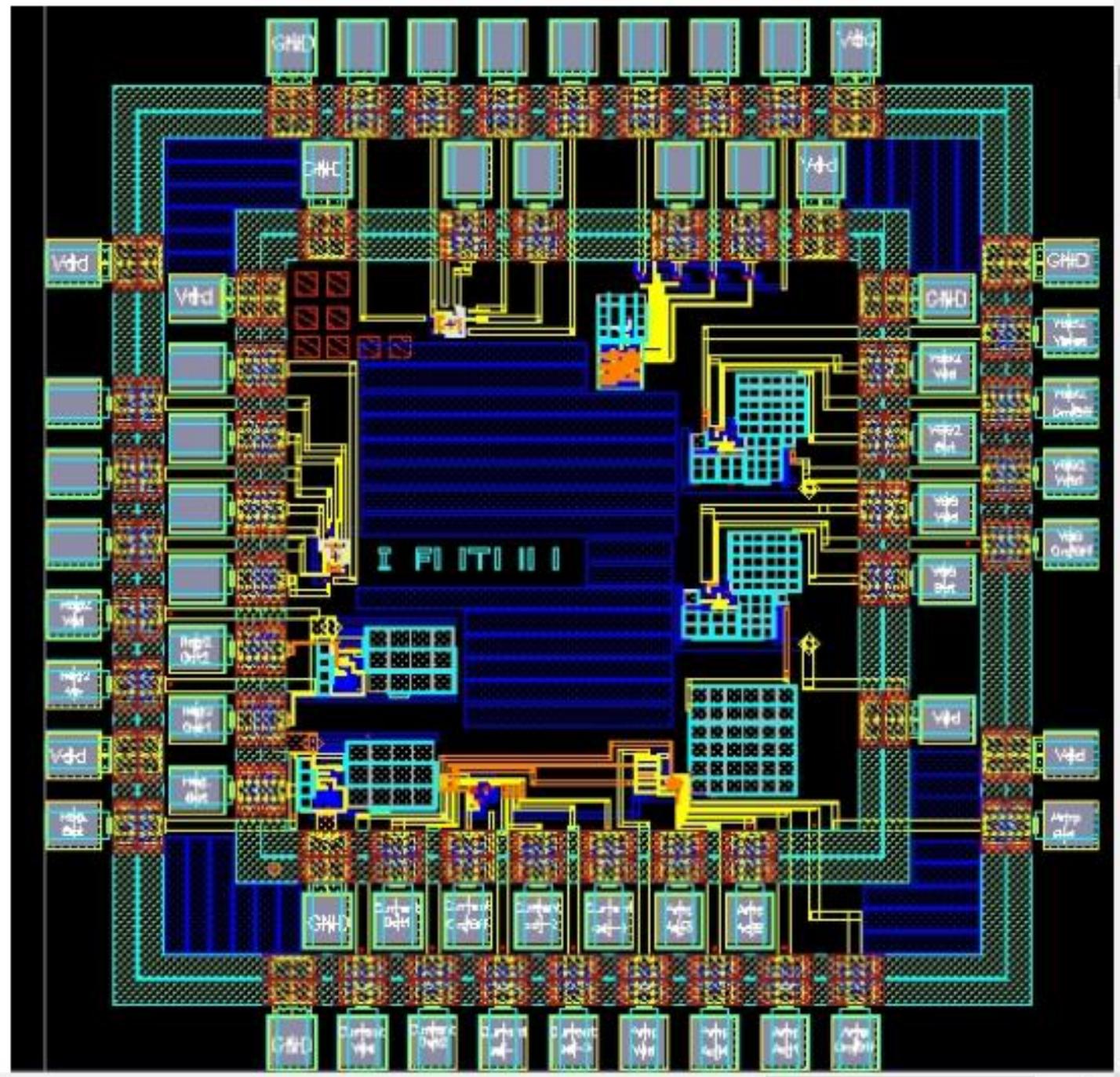


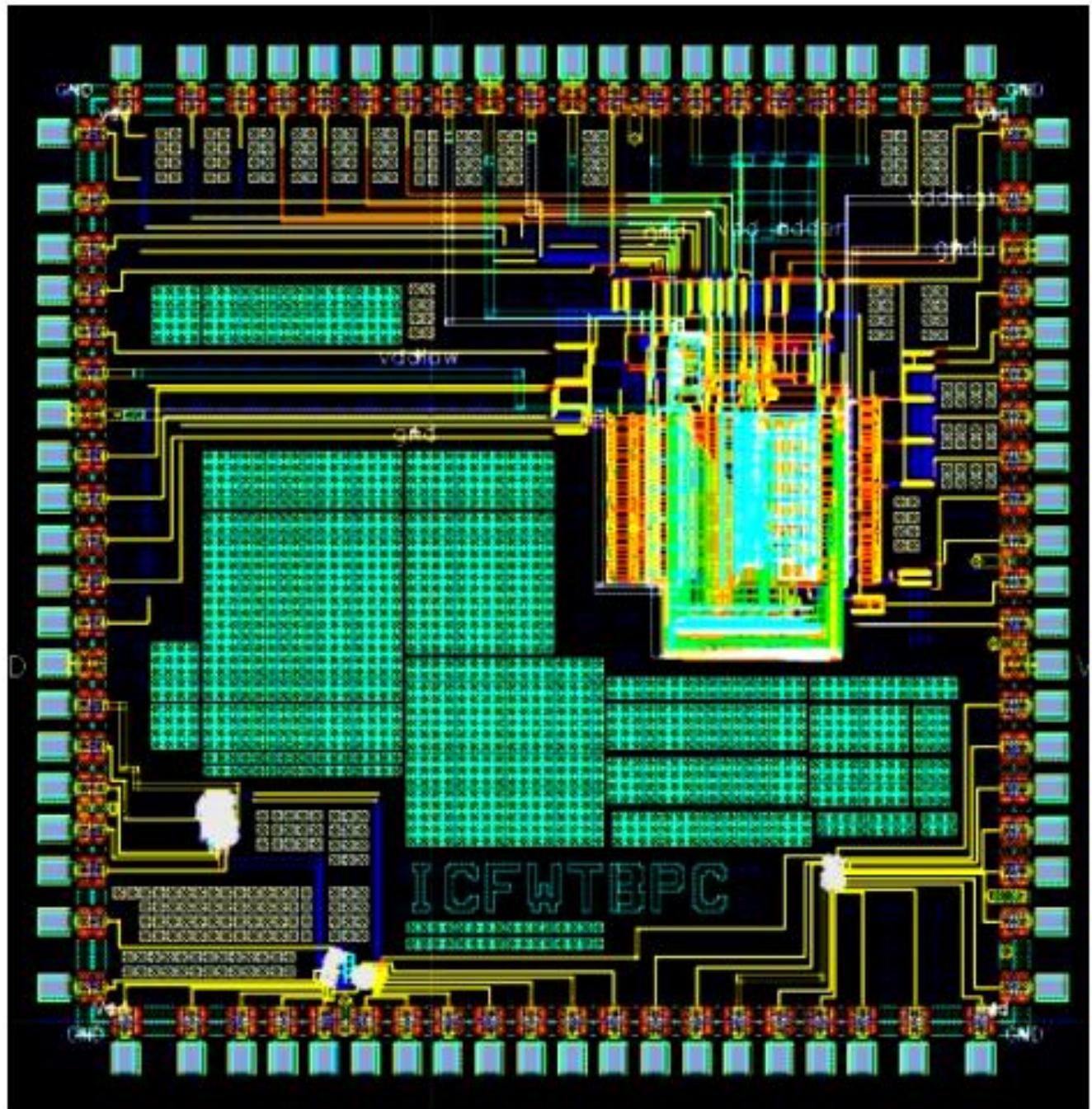
Figure 4-3 Current flow through ESD network during CDM event

Video link for ESD:

1. <https://youtu.be/hWxh3HVbcsg>
2. <https://youtu.be/plZlwt1Mlqw>

Layout view



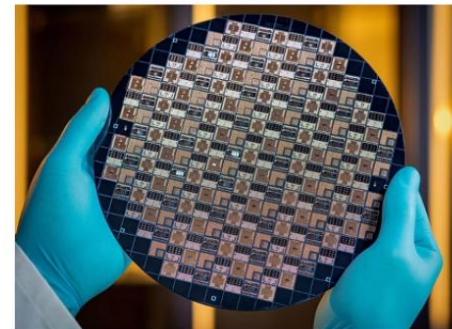
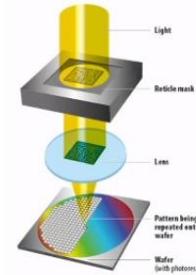


29. EXAMPLE OF ANALOG CIRCUITS

- Operational Amplifier : This includes differential input pair,bias circuit and current mirror.
- DAC: This includes constant current source, amplifier using external Rset to adjust full rang current and bias circuit.
- ADC: this includes comparator, amplifier, sample/hold switch, switching capacitor and reference voltage resistor ladder.
- PLL: this includes VCO (delay stage) and charge pump (current mirror and buffer/opamp)
- Bandgap: BJT, current mirror, bias circuit, differential amplifier and ratioed resistor

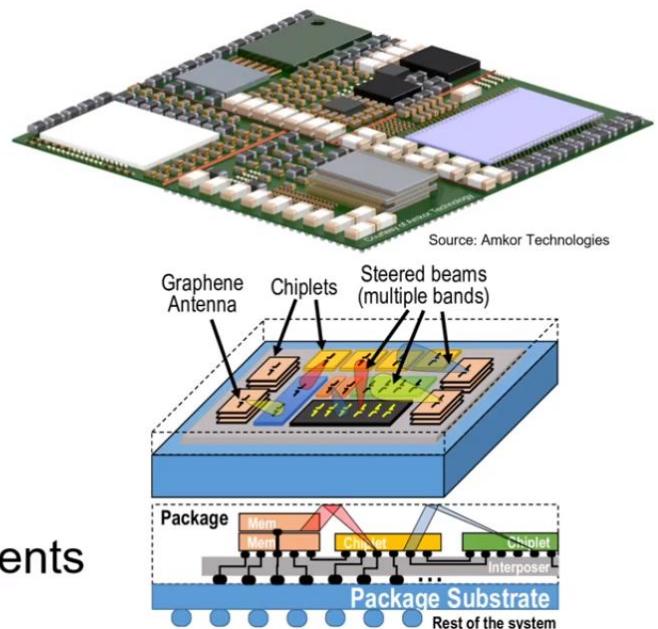
30. WAFER LEVEL PACKAGING (WLP)

- ❖ Creates package while part of a wafer
 - Photolithography
 - RDL and solder balls built onto the wafer
 - Dicing happens once the package is finished
- ❖ Fan-In and Fan-Out
- ❖ Small device applications
- ❖ Much cheaper and reliable alternative to gold wire bond



What is System in Package?

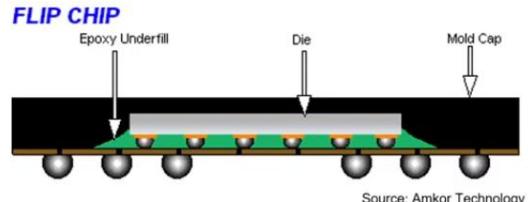
- ❖ Group of ICs mounted on a common substrate
- ❖ Heterogeneous integration
 - More than Moore
- ❖ Uses Interposer and TSV technology
 - High density interconnections
- ❖ Usually bare-die, flip chips.
- ❖ Can contain passive components embedded in the substrate



Components of SIP - Interconnections

❖ Flip Chip technology

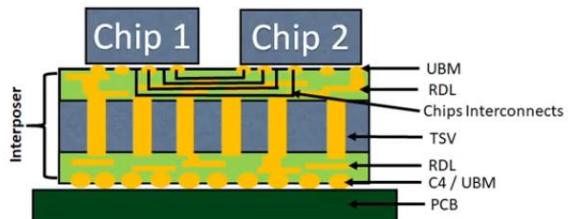
- Die flipped upside down
- Solder bumps connect directly to substrate



Source: Amkor Technology

❖ Interposer Layer

- Through Silicon Vias (TSV)
- Redistribution Layer (RDL)
- Under Bump Metallization (UBM)

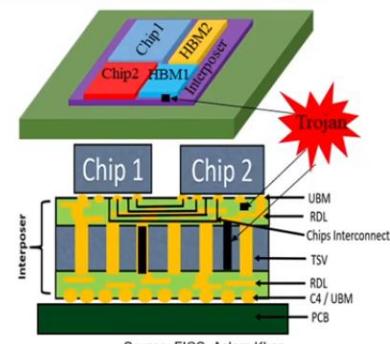


Source: FICS, Aslam Khan

Security for Sip

❖ Security threats in different steps

- Interposer, individual dies, passive components
- Unknown devices, changes in photolithography, recipe, materials...

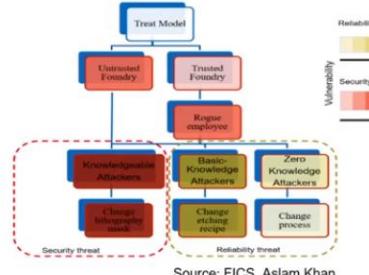


Source: FICS, Aslam Khan

❖ Potential issues causing data to be stolen / lost and device failure

❖ More opportunities for protection

- Sensors, prevention, firewalls...
- More potential ways to monitor operation

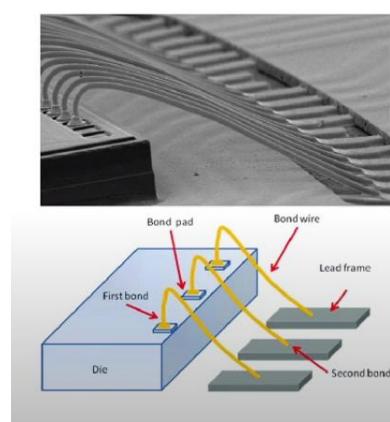


Source: FICS, Aslam Khan

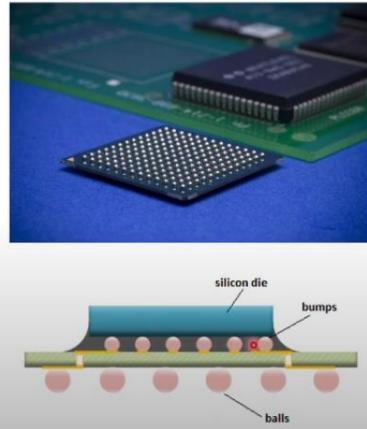
9

Types

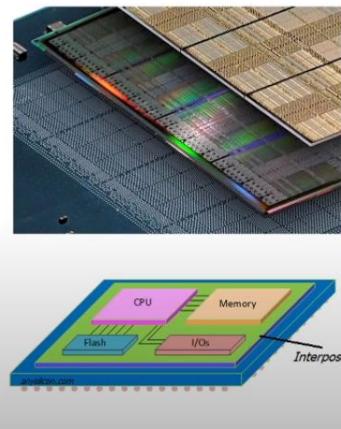
Wire Bond



Flip Chip

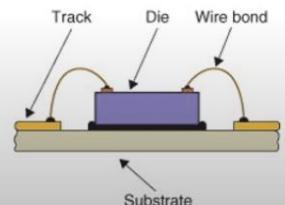
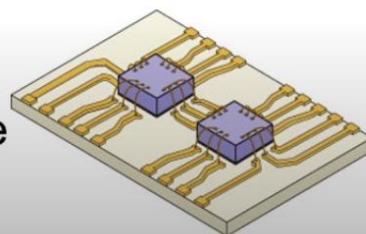
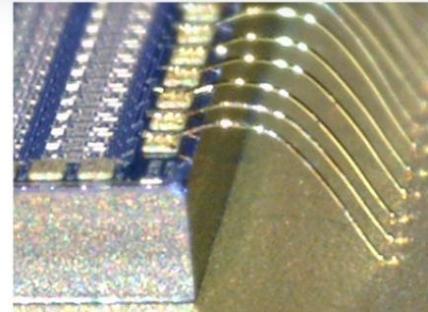


Interposer



Wire Bond

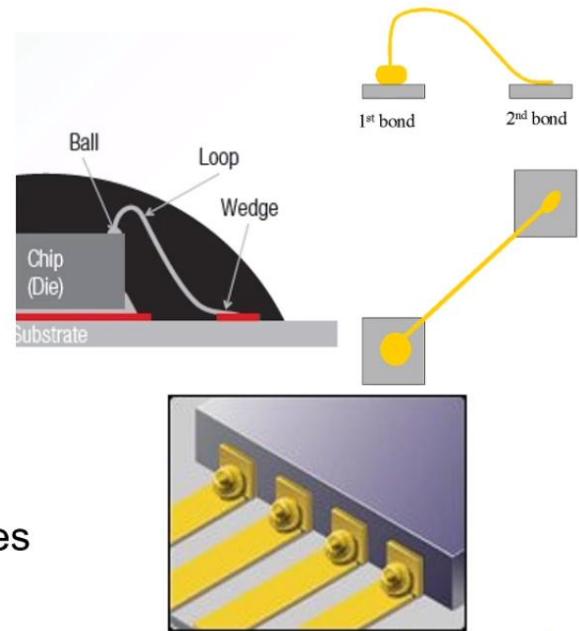
- ❖ Interconnection between die and packaging using wires (Au, Cu, Al, Ag)
- ❖ Most cost effective and flexible
- ❖ Simple implementation, high reliability
- ❖ Uses pressure, heat, and ultrasonic energy to weld the metals



Bonding Techniques

❖ Ball Bonding

- Sphere formed at the end of wire
- No directionality



❖ Wedge Bonding

- Wire pressed against pad
- Directionality – aligning tools

❖ Stud / Bump Bonding

- ❖ Bump on surface, no wire
- ❖ Can connect perpendicular surfaces

❖ Thermosonic Bonding

Flip Chip

- ❖ Connects semiconductor devices to board using solder bumps deposited on pads
- ❖ Chip is flipped so top side faces down
- ❖ Aligned with the pads on the board
- ❖ Smaller chip and shorter interconnections

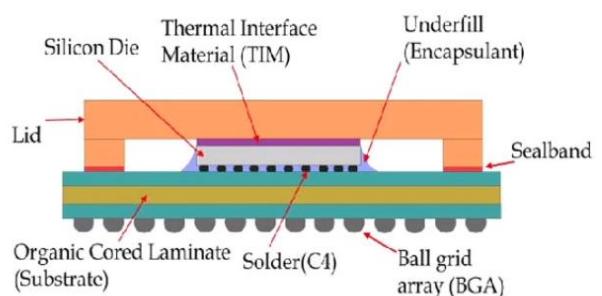
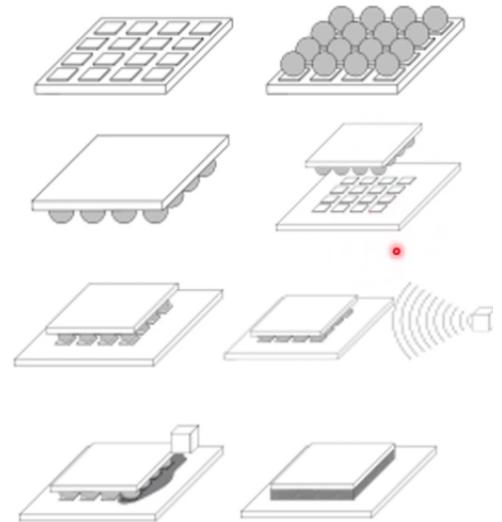


Figure 1. Schematic of a typical flip-chip package assembly.

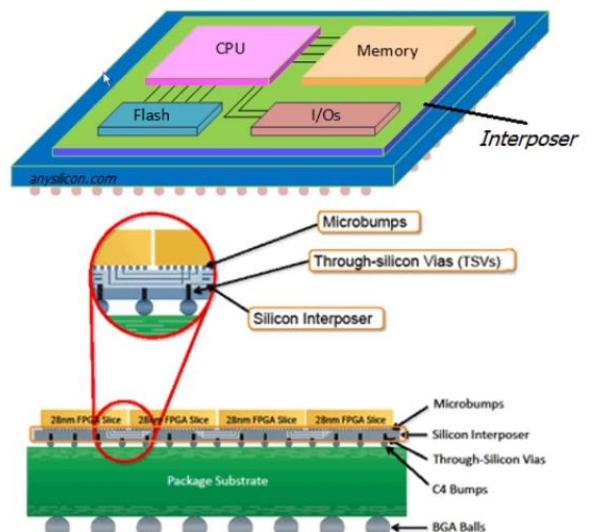
Manufacturing Process

1. IC's are created on the wafer
2. Pads metallized on chip surface
3. Solder dot deposited on pads
4. Chips are cut
5. Chips flipped and aligned
6. Solder remelted to complete connection – Reflow
7. "Underfilled" using isolating adhesive



What is a Silicon Interposer

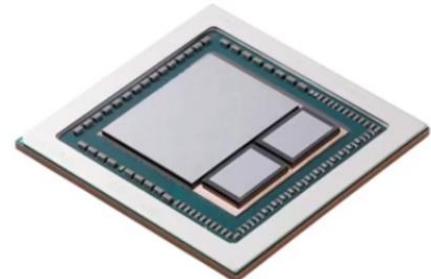
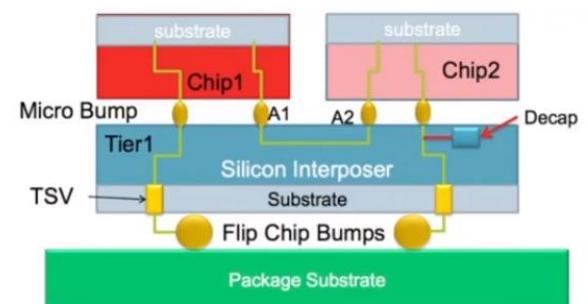
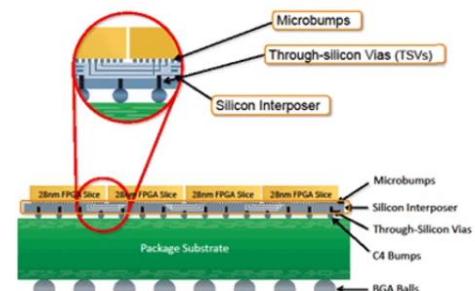
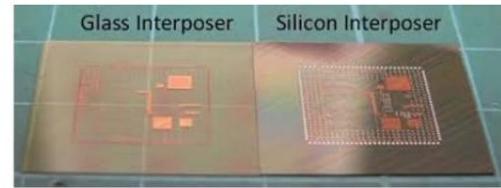
- ❖ Silicon layer between substrate and dies
- ❖ More than just packaging – also serves as a connection
 - Among dies and to the I/Os
- ❖ Uses Through Silicon Vias Technology (TSV)
- ❖ Can be Active or Passive
- ❖ Fabricated by Foundries



Interposer

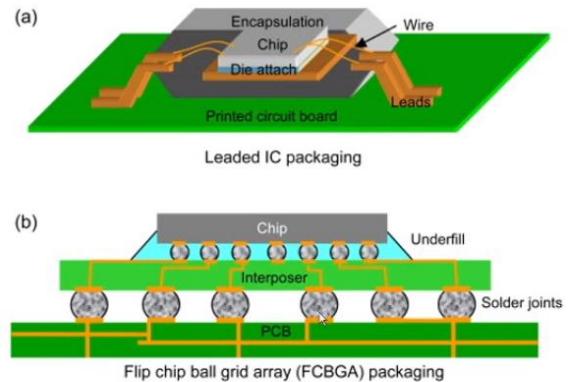
- ❖ Layer between the substrate and the dies
- ❖ More than just packaging – also serves as a connection
- ❖ Silicon, glass, or polymer
- ❖ Shorter interconnections, higher bandwidth, better performance
- ❖ Active or passive interposer
- ❖ Makes use of flip-chip

- ❖ Silicon interposer between SiP substrate and dies
 - Dies are flip-chipped
- ❖ Uses TSV and RDL to connect the dies to one another
 - ❖ Also connects them to the I/Os
- ❖ Horizontal integration on the interposer
- ❖ Not just a steppingstone for 3D



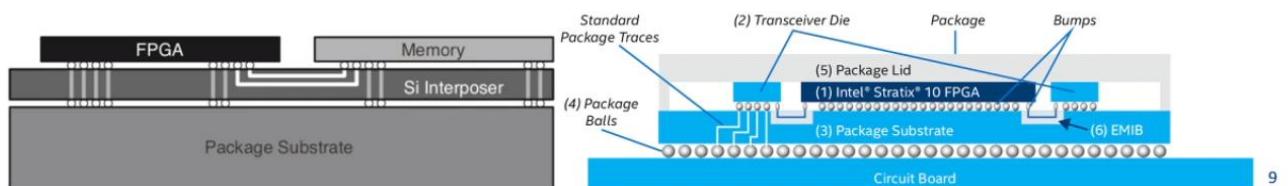
Advantages of the Silicon Interposer

- ❖ High wiring density
- ❖ CTE matched to the silicon die
 - Interposer to chip specific
- ❖ Excellent electrical and thermal performance, lower power.
 - Shorter interconnection chip – substrate
- ❖ Lower cost of active devices due to partitioning large die
- ❖ Possibility of integrating passive devices into the substrate



Interconnection Types

- ❖ Through Silicon / Glass / Polymer Vias
 - Vertical copper pathway through the interposer
 - Has a Redistribution Layer (RDL) on the top and bottom
- ❖ Bridge
 - In-package interconnect for multi-die packages
 - EMIB – Embedded Multi-die Interconnect Bridge (Intel)
 - Potential of becoming an independent component



Overview

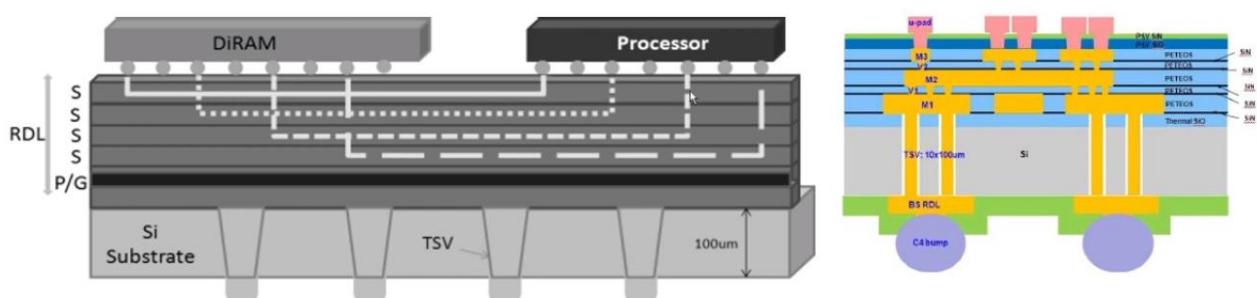
	Wire Bond	Flip-Chip	Interposer
Cost	Low	Medium	High
I/Os and Bandwidth	Low	Medium	High
Chip Size	Requires more space	Requires less space	Requires the least space

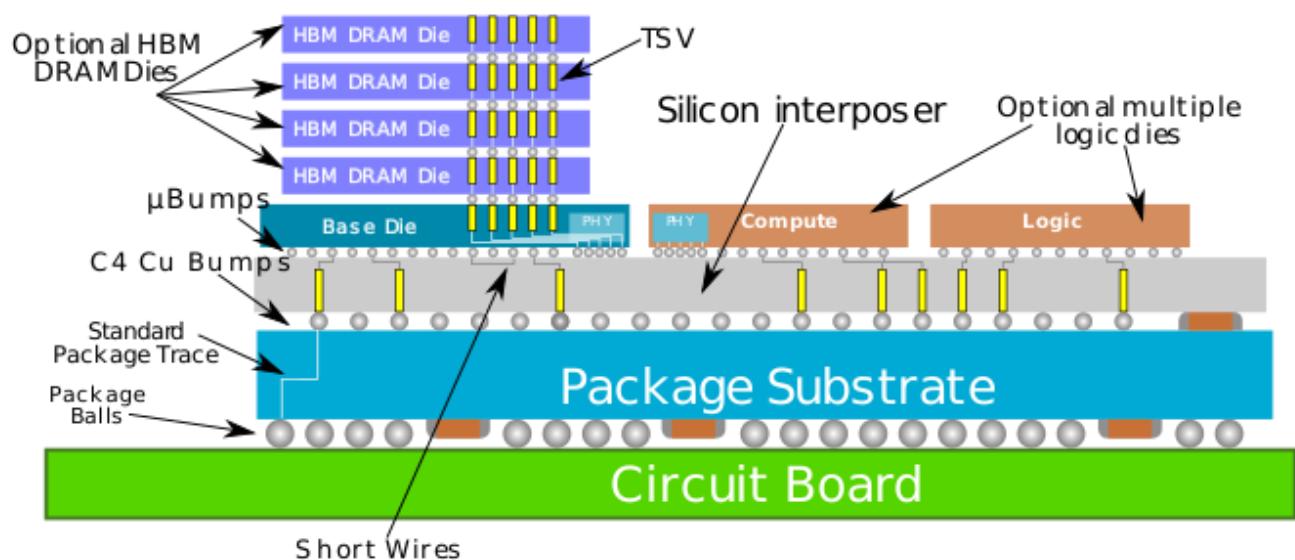
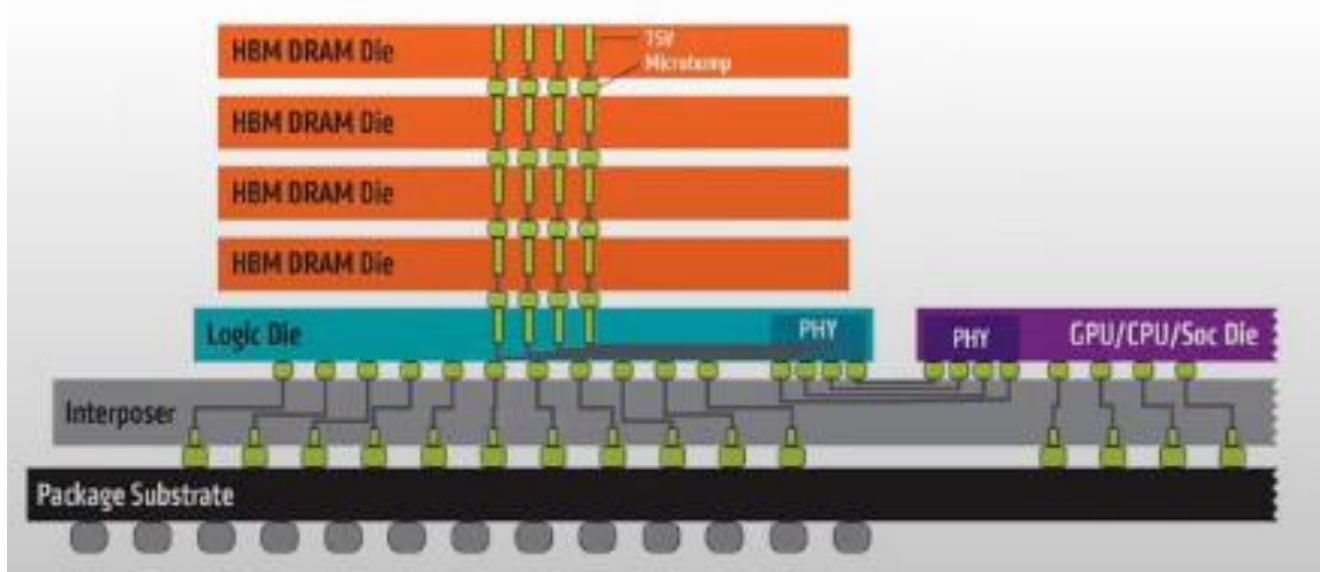


12

RDL- Redistribution Layer

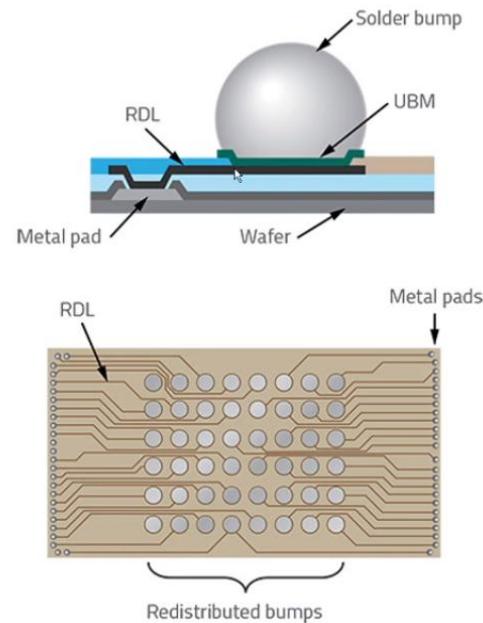
- ❖ Horizontal pathway along the interposer
 - Connects dies to one another
- ❖ Connects the solder bumps to the TSV by rerouting
- ❖ Copper lines etched into SiO₂ layers and polished





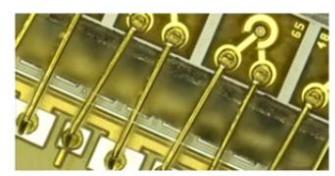
UBM – Under Bump Metallization

- ❖ Thin pad connecting the solder bump and the copper in the RDL
- ❖ Serves as barrier to stop diffusion
- ❖ Acts as a mechanical connection
- ❖ Nickel-based UBM widely used

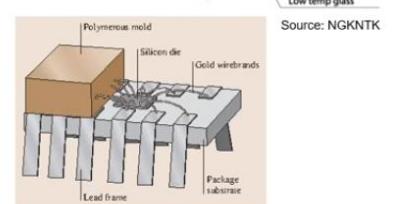
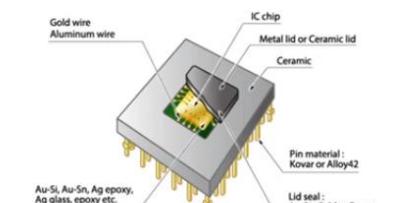


Package Materials

- ❖ Leadframe – Alloy 42, Inconel, Cu
- ❖ Encapsulant – Epoxy, Silicon, Polyimide
- ❖ Die – Si, SiC
- ❖ Die Attach – Polyimides
- ❖ Substrate – Organic laminate materials
- ❖ Interposer – Si, Glass, Organics
- ❖ Dielectrics - Oxides, Semiconductors, Polymers, Ceramics, Composites...
- ❖ Interconnects – Cu, Au, Ag, Al...

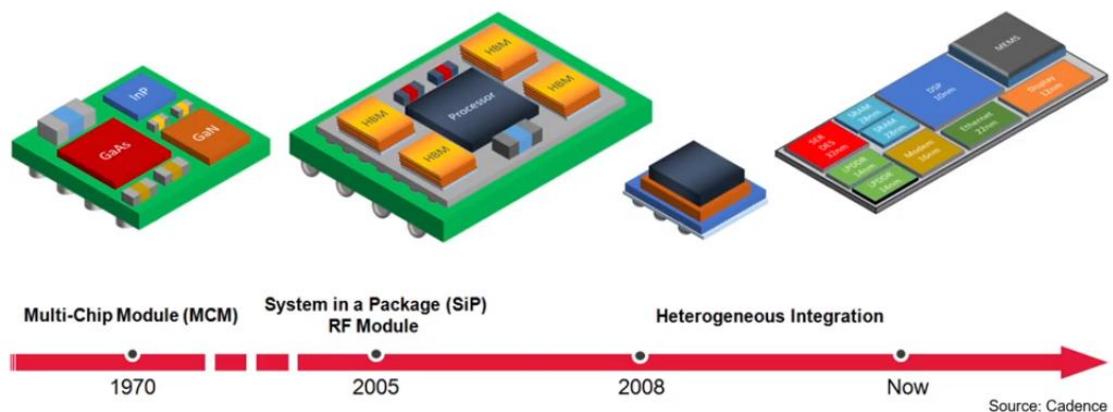
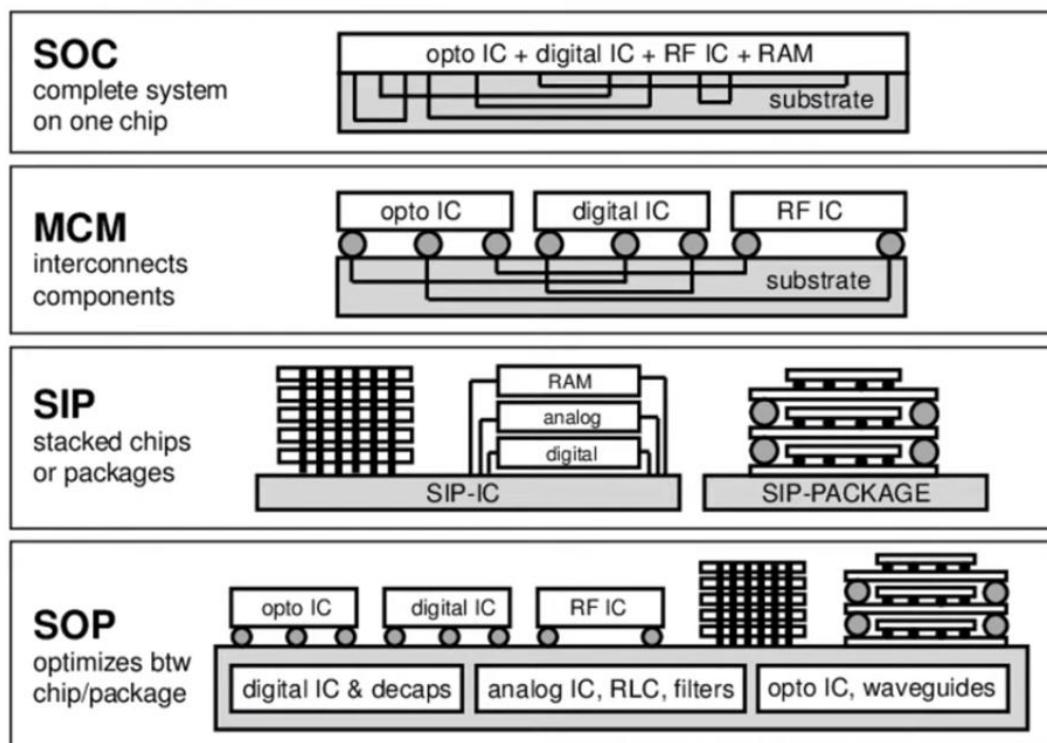


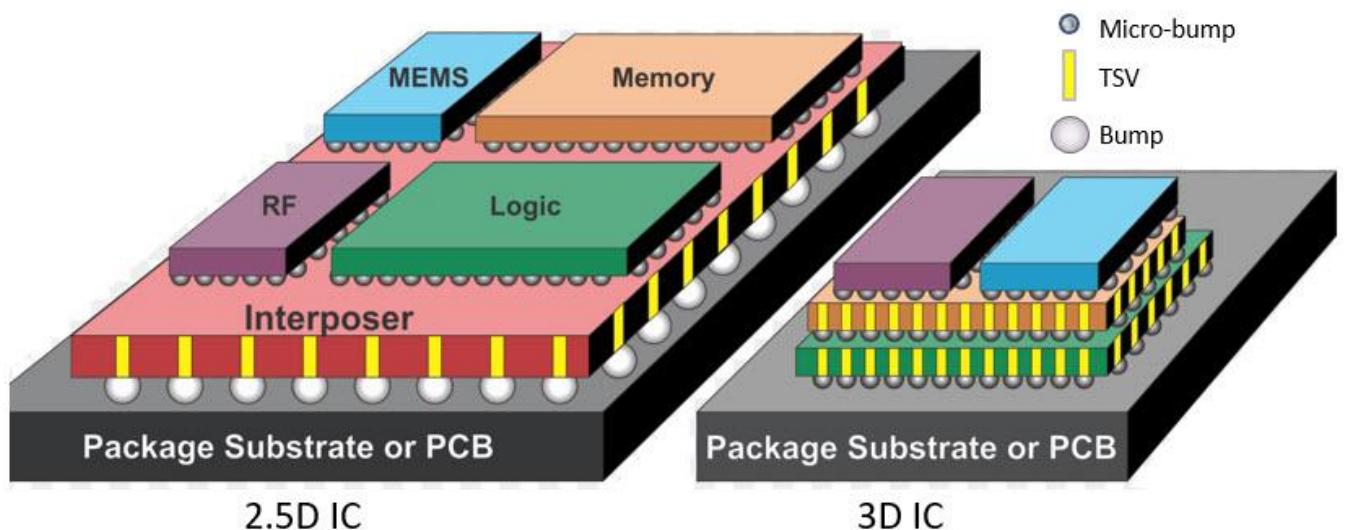
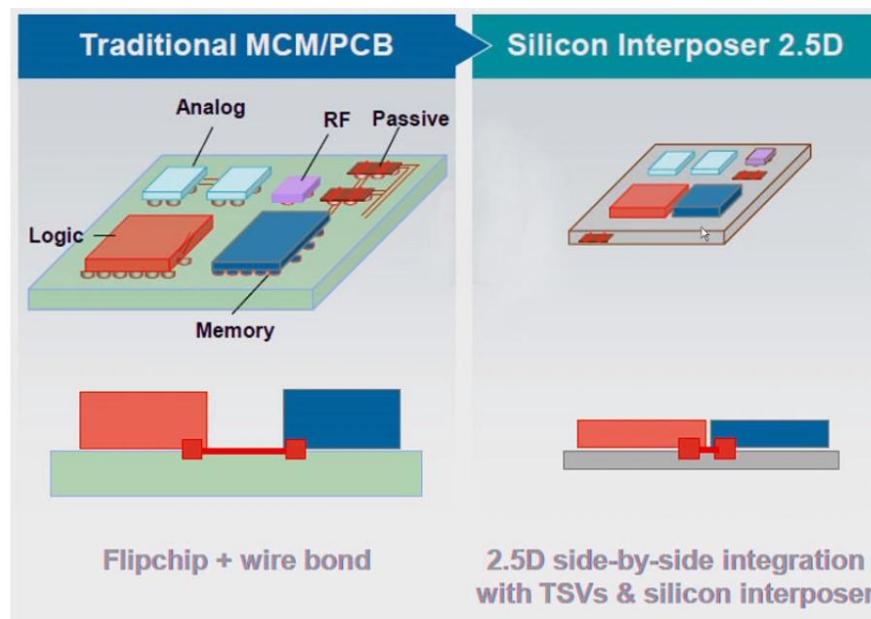
Source: Palomar Technologies



3

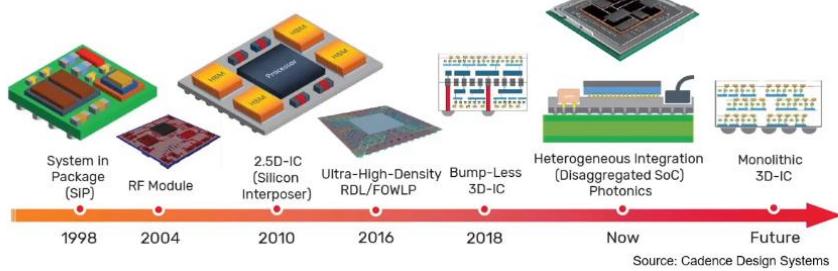
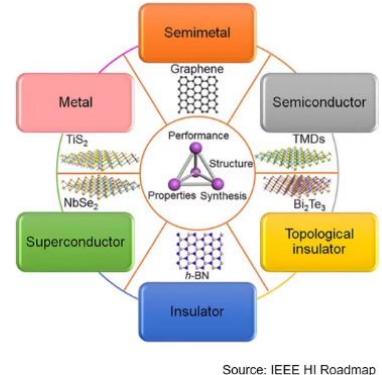
Summary





Future

- ❖ 2D materials for nano applications
- ❖ AI can learn to design materials
- ❖ New technology for better switches, transistors, conductors, bonding...
- ❖ Monolithic 3D ICs
- ❖ Higher density, efficiency...



Video link:

1. <https://youtu.be/WMQtD4hDHak>

2. <https://youtu.be/igIm170dkms>

31. INDUSTRY EXAMPLE

Block diagram:

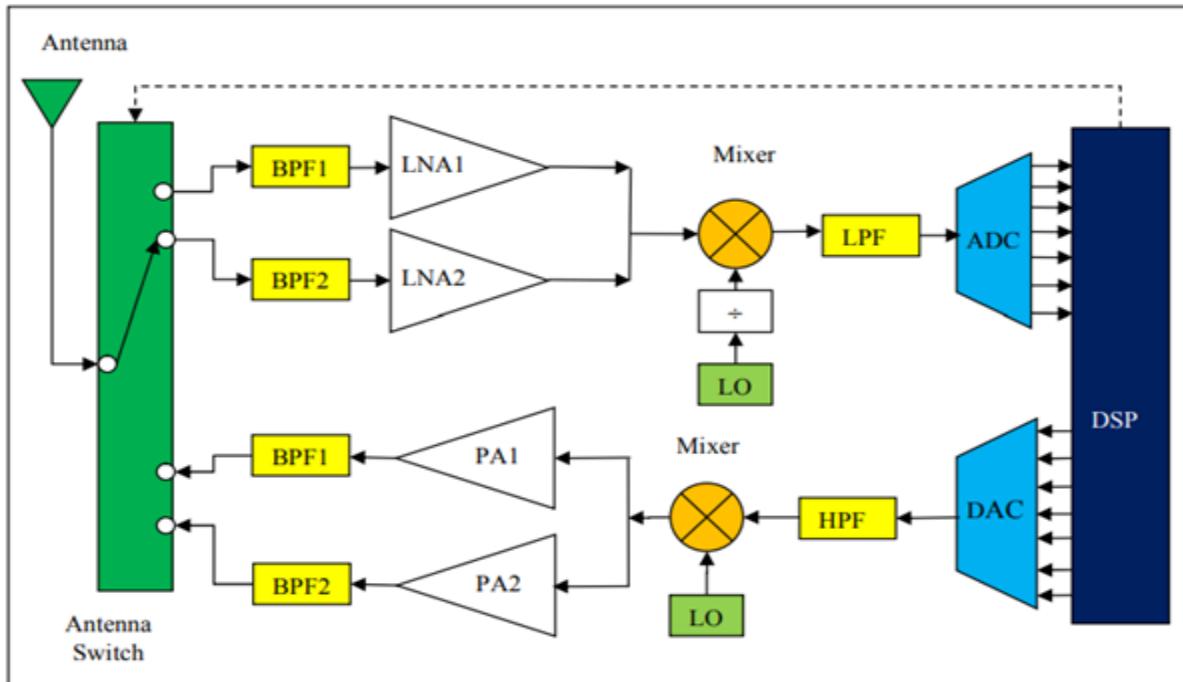
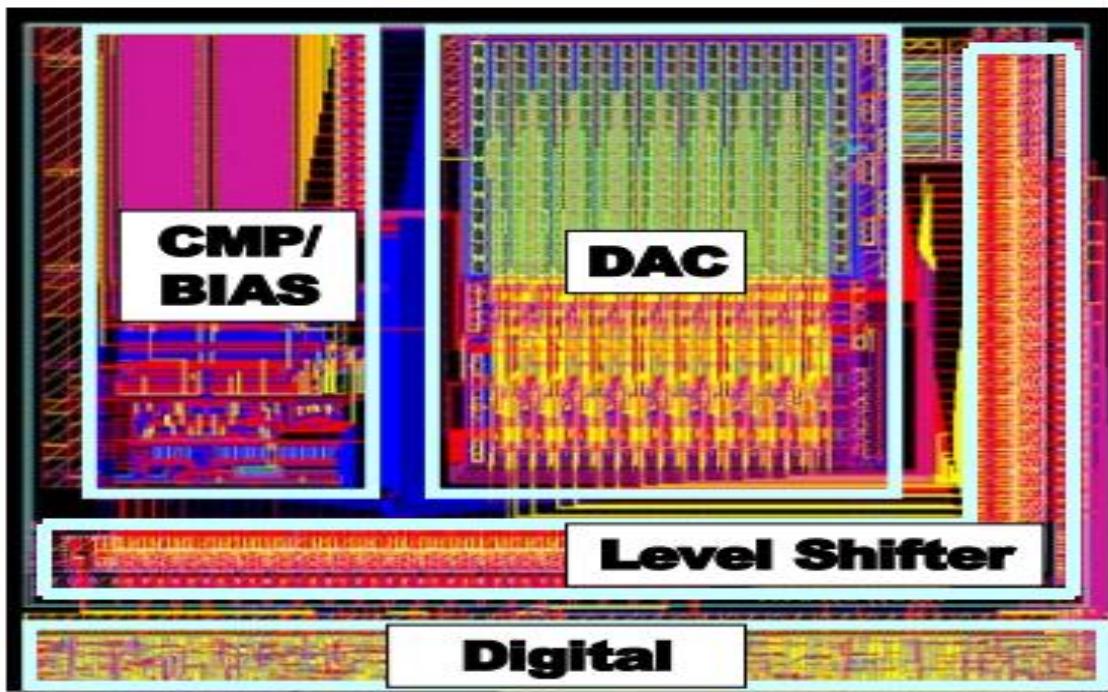


Figure — Functional block diagram of a radio transceiver

Example of Celllevel:



Example of a chip level layout:

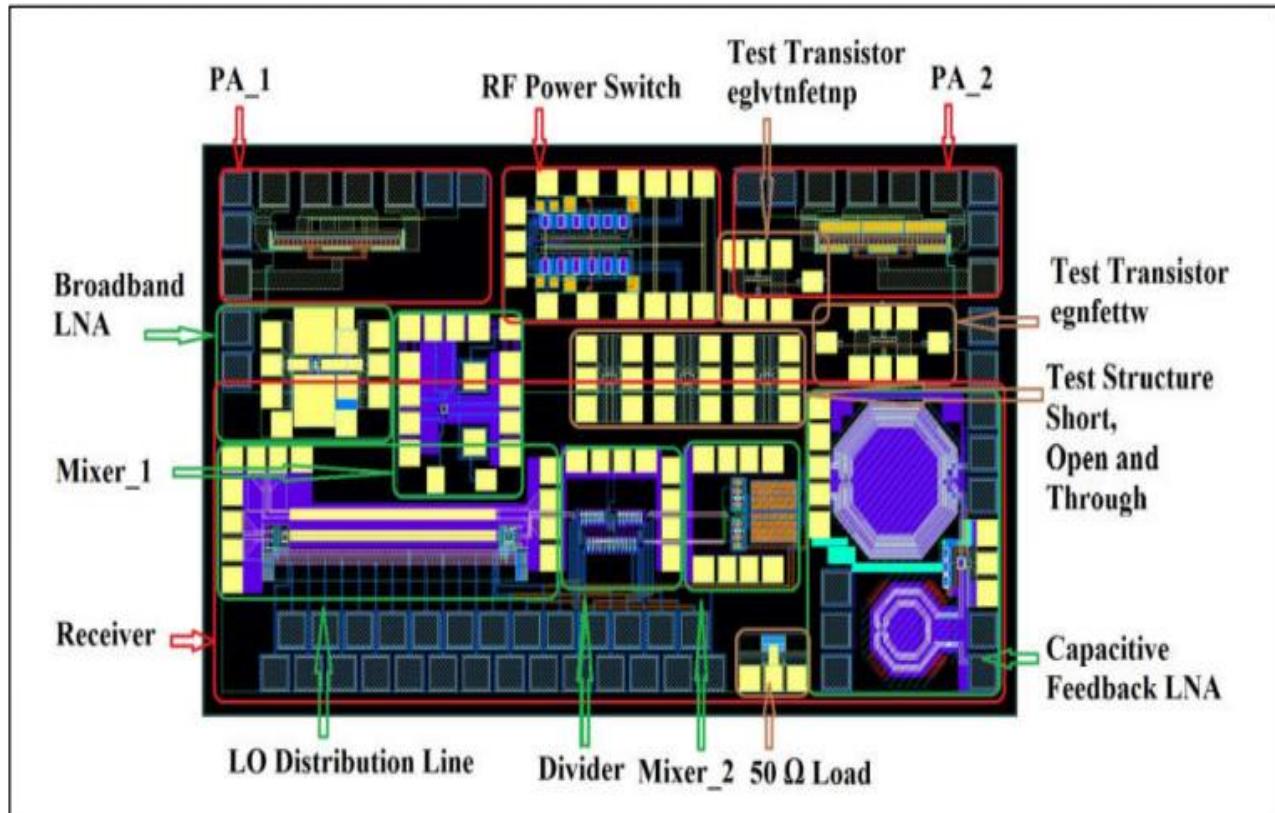


Figure 2-6 Final layout of the test chip

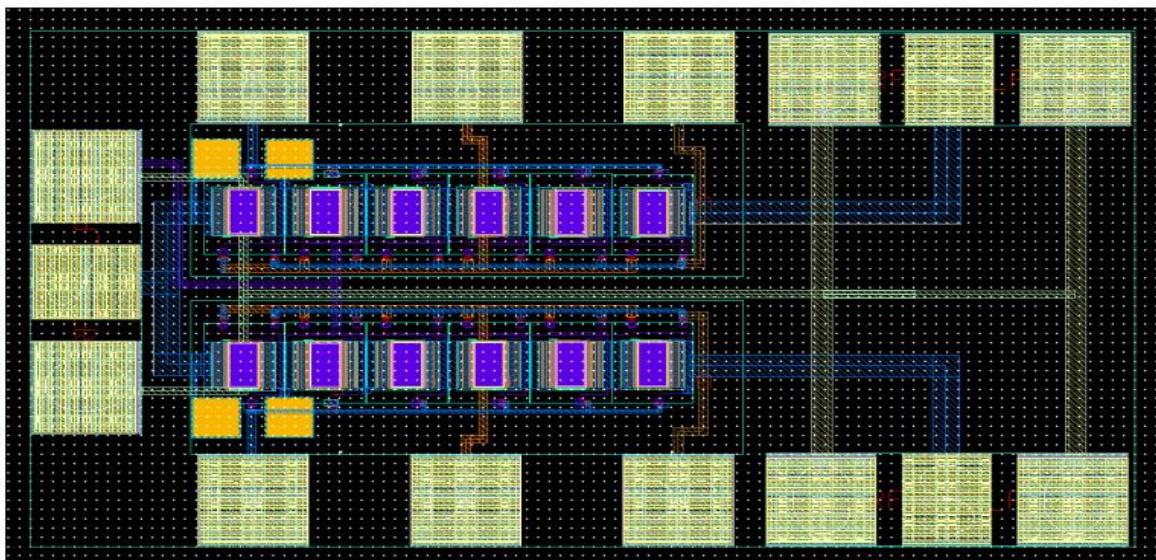
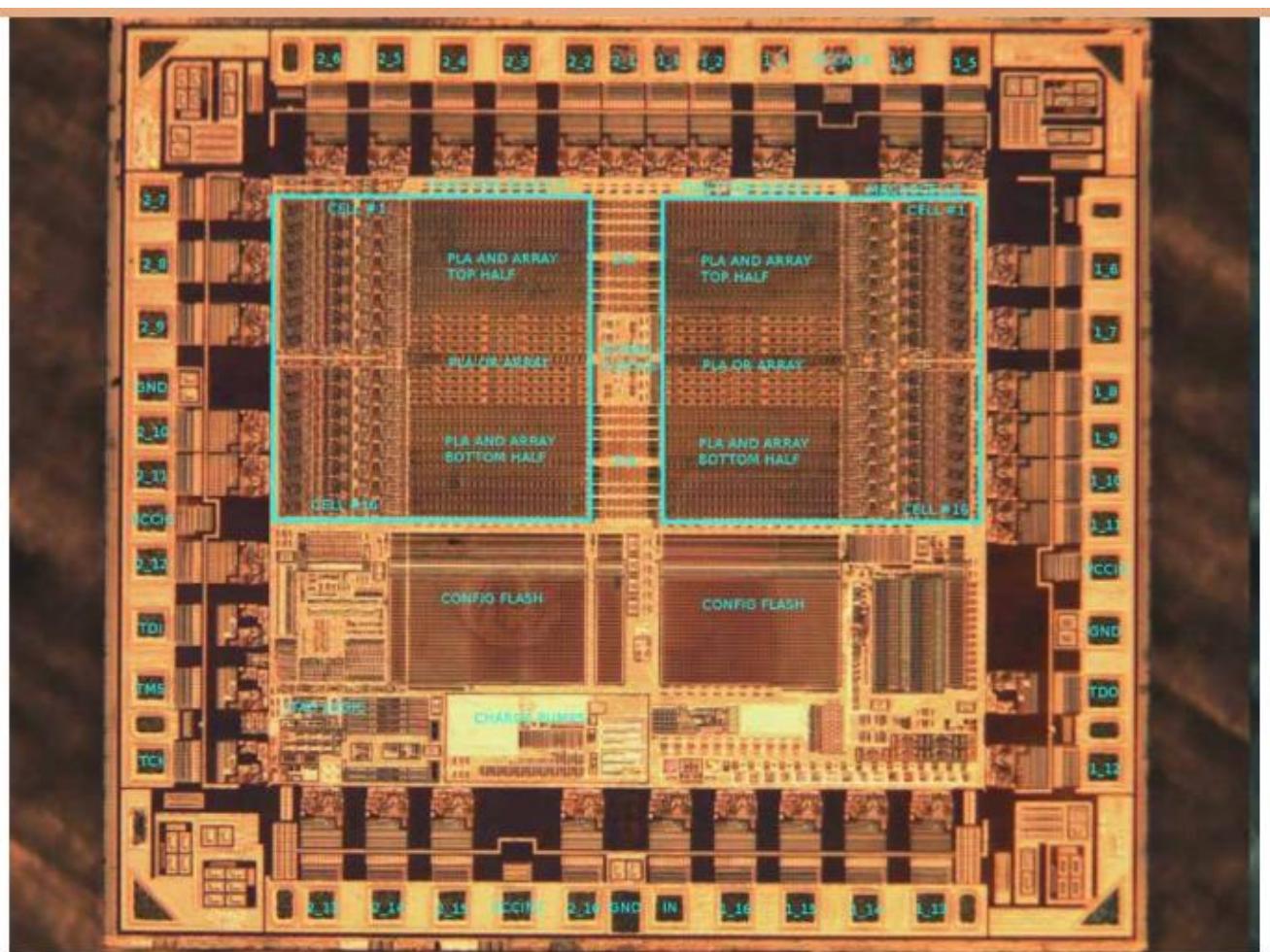


Figure 4-6 Complete layout of the six-stacked RF power switch



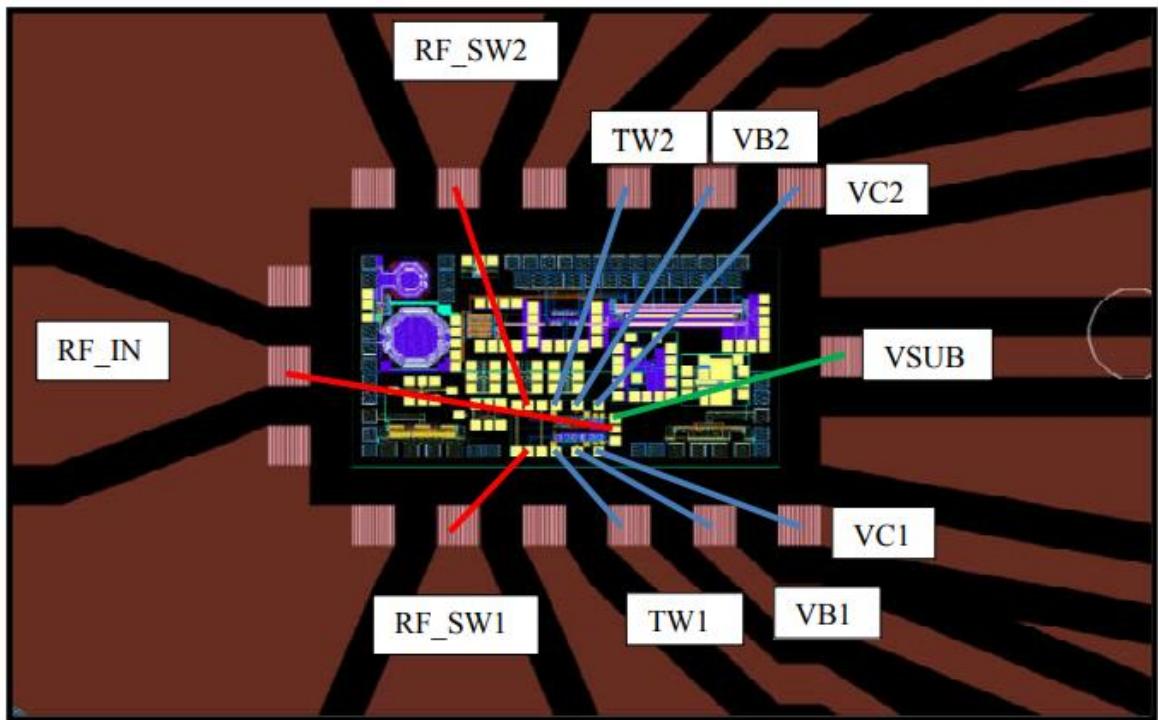


Figure 4-12 Bond-wire diagram of RF power switch in a test PCB (sample 1)

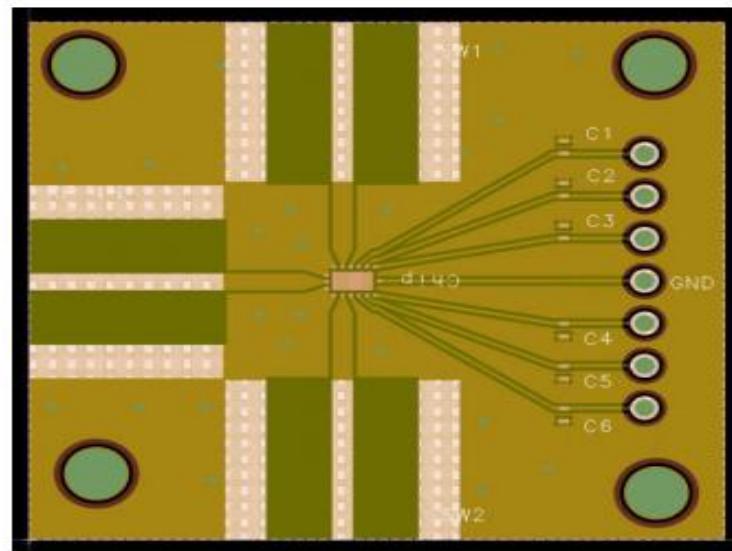


Figure 4-11 Test PCB for the RF power switch

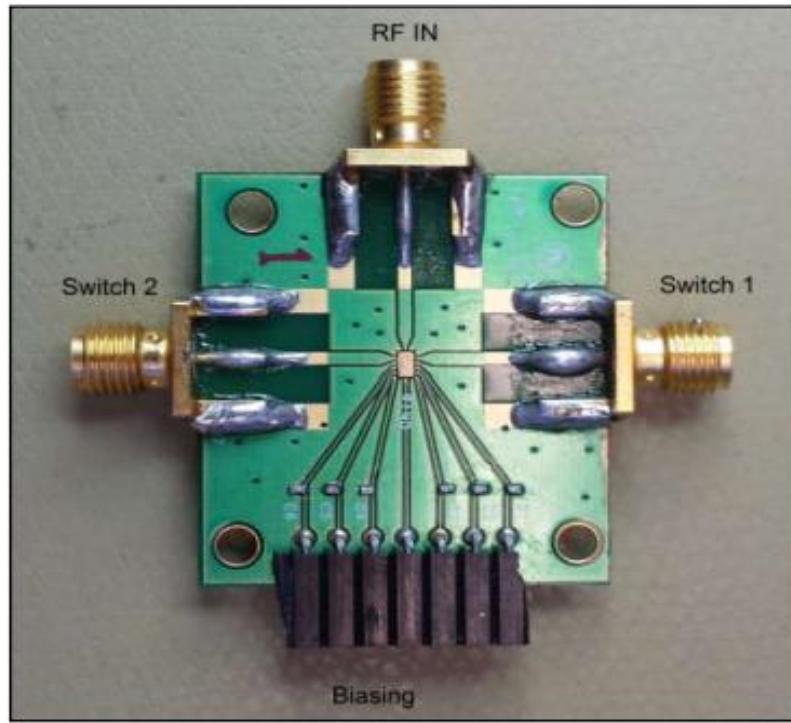


Figure 5-2 Test PCB with mounted chip and other components (sample 1)

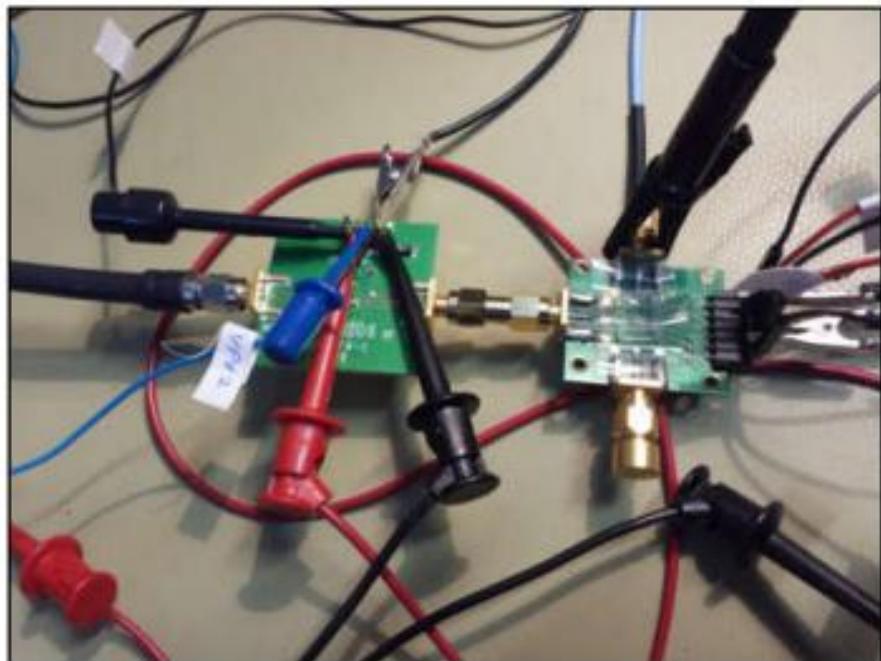
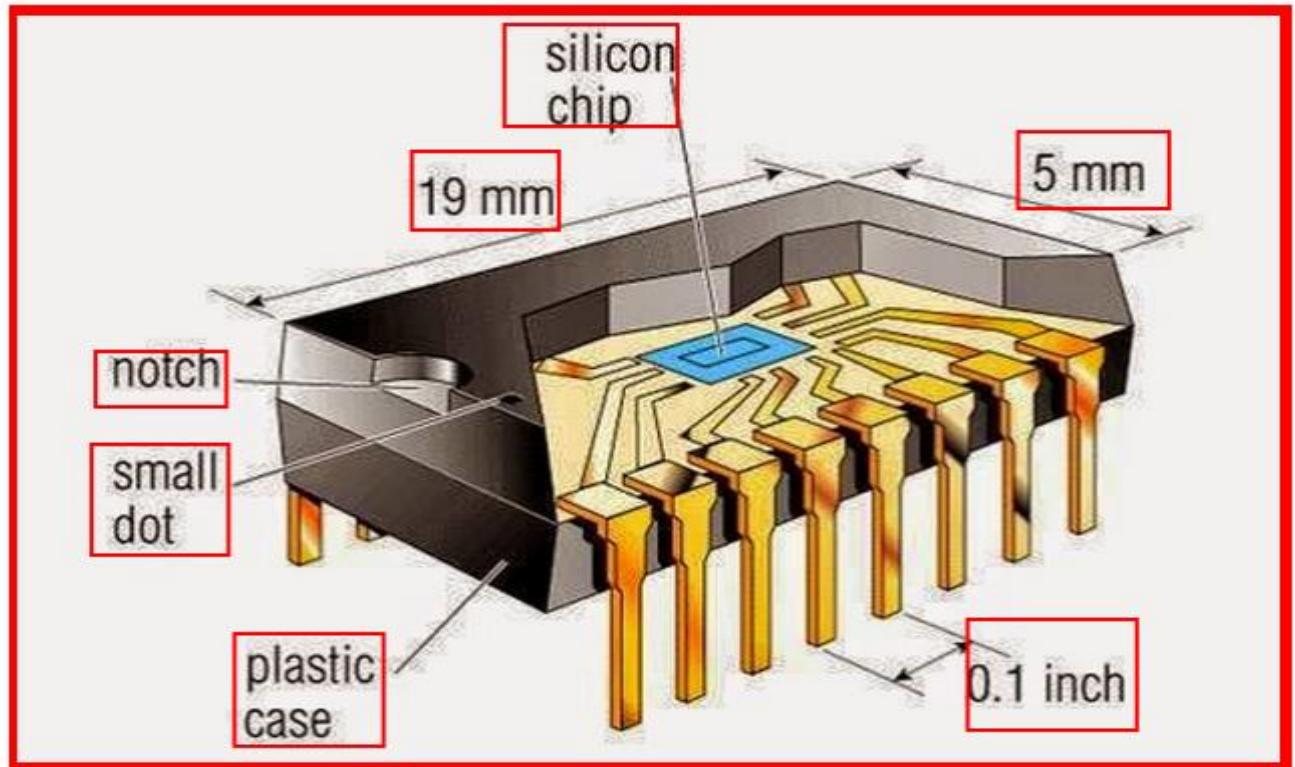


Figure 5-5 Measurement setup for the RF Switch



Linkedin:- <https://www.linkedin.com/in/shubham-kumar-4354b7135/>