14EI802

Hall Ticket Number:

IV/IV B. Tech (Regular/Supplementary) DEGREE EXAMINATION

March, 2019 — **Electronics & Instrumentation Engineering**

Eighth Semester — **VLSI Design**

Time: Three Hours — **Maximum :** 60 Marks

Answer Question No.1 compulsorily. (1X12 = 12 Marks)

Answer ONE question from each unit. (4X12=48 Marks)

1 Answer all questions (1X12=12 Marks)

a) What is photolithography in VLSI design.
b) Illustrate latch up condition in CMOS circuits? How to prevent it?
c) Implement OR gate using pass transistor design.
d) What is propagation delay.
e) Define sheet resistance.
f) What are stick diagrams.
g) Explain switch logic.
h) List some architectural issues in VLSI design
i) Distinguish between PAL & PLA.
j) Distinguish function and procedures.
k) What is clocked CMOS logic.
l) Give the general structure of a VHDL program.

UNIT I

2 a) Explain in detail about NMOS fabrication technology. 6M
b) Describe the CMOS inverter &derive its dc characteristics. 6M

(OR)

3 a) Explain in detail about the I_{ds} and V_{ds} relationship 6M
b) Explain the alternative forms of pull-ups for an inverter 6M

UNIT II

4 a) Describe three sources of wiring capacitances. Explain the effect of wiring capacitances on the performance of a VLSI circuit. 6M
b) Draw a layout of 2 input CMOS AND gate indicating all the regions and layers. 6M

(OR)

5 a) Explain about various design rules. 6M
b) Explain how sealing factors influence the design of a device. 6M

UNIT III

6 a) Discuss architectural issues in design process. 6M
b) Explain about design of an ALU subsystem. 6M

(OR)

7 a) Compare switch logic and gate logic. 6M
b) Give the subsystem design considerations of a four-bit multiplier. 6M

UNIT IV

8 a) Design a full adder using PLA. 6M
b) Explain about ASIC'S 6M

(OR)

9 a) Explain about structural design elements in VHDL. 6M
b) Write a program to implement a 4-bit adder in data flow model. 6M

IV/IV B.Tech Regular/Supplementary Degree Examination. 14EI802
March 2019 VLSI Design.

(2)

Scheme of Evaluation

1. $12 \times 1 = 12M$

2 (a) Six steps - $6 \times 1 = 6M$

(b) circuit diagram - 2M
characteristic - 2M
Derivation - 2M

3. (a) Exp. for T_{sd} - 2M

I_{ds} vs V_{ds} for non saturated region - 3M

I_{ds} vs V_{ds} for saturated region - 1M

(b) Any two forms of pull up - $2 \times 3M$

4 (a) Fringing fields - 2M
Interlayer capacitances - 2M
peripheral capacitance - 2M

(b) CMOS methodology - 1M
CMOS AND gate circuit diagram - 3M
Layout diagram - 2M

5 (a) Any Six rules - $6 \times 1 = 6M$

(b) Any Six scaling factors - $6 \times 1 = 6M$

6 (a) Any Six issues - $6 \times 1 = 6M$

(b) Block diagram - 3M
Explanation - 3M

7 (a) Any three Comparisons - $3 \times 2 = 6M$

(b) Design Consideration - 3M
Design of multiplier - 3M

8 (a) Truth table of Full adder - 2M
Sum and Carry expressions - 1M
Realisation using PLA - 3M

(b) classification of ASICs - 1M
Description of each - $5 \times 1 = 5M$

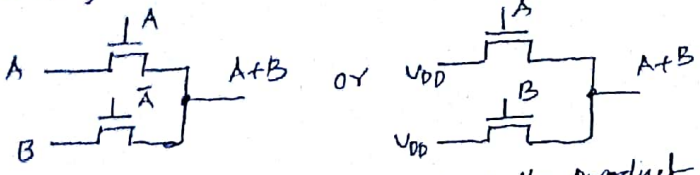
9 (a) Description of Structural Design - 2M
Syntax of VHDL Component declaration - 2M
Syntax of VHDL Component statement - 2M

(b) Package declaration - 1M
Entity declaration - 2M
Architecture body - 3M

1. (a) photolithography is the technology to create a pattern on the silicon wafer using an ultraviolet (UV) ray of light.

(b) In cmos circuits, the parasitic NPN and PNP bipolar junction transistors (BJT) form a thyristor or PNPN structure. When the thyristor is triggered by noise/glitch, the BJTs become on, and it leads to a formation of a stable circuit between the power and ground lines, known as cmos latch up. Prevention techniques are (i) use of p+ and n+ gaurd rings around nmos and pmos connected to ground and V_{DD} respectively. (ii) keeping sufficient spacing between nmos and pmos transistors.

(c)



(d) The delay unit τ is defined as the product of sheet resistance, R_s and standard unit gate capacitance, $\square C_g$ i.e., $\tau = R_s \square C_g$

(e) Sheet resistance, $R_s = \rho/t$ ohm per square where ρ = resistivity, t = thickness

(f) A stick diagram is a simple way of representing the layout by using thick lines with their interconnections.

(g) Switch logic is the one which uses pass transistors & transmission gates.

(h) Requirement of a logical and systematic approach.

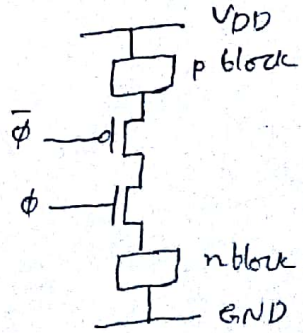
(i)

PAL	PLA
1. Programmable AND, fixed OR	(1) Both AND, OR are programmable.
2. cheaper than PLA	(2) costlier than PAL.
3. Moderate flexible	(3) Extremely flexible.
4. It can implement SOP with limited number of terms	(4) It can implement SOP with any number of terms.

(j) Functions :- Functions are used to describe frequently used sequential algorithms that return a single value. This value is returned to the calling program using a return statement.

Procedures :- These are used to partition large behavioral descriptions. A procedure can return zero & more values using parameters of mode out and inout.

(k) The arrangement of a clocked cmos logic is shown below. ϕ is the clock signal.



The logic in clocked cmos is evaluated only during the on period of the clock.

(l) VHDL program bike structure is shown below.

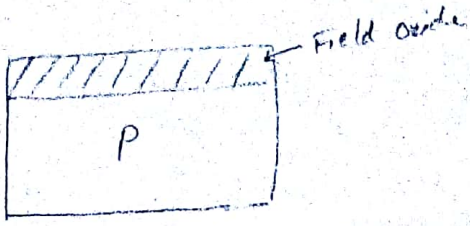
Entity declaration

Architecture definition

2(a)

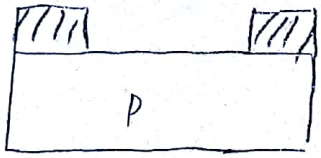
(5)
(4)

Fabrication of n-MOSFET -

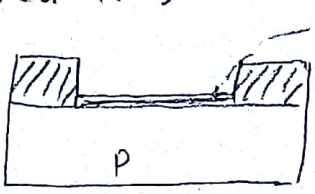


(a) Field oxide

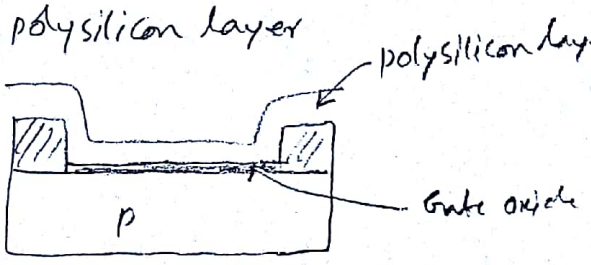
→ open a window in the field oxide



→ Subject to gate oxidation

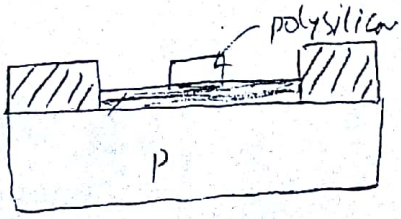


→ deposit polysilicon layer

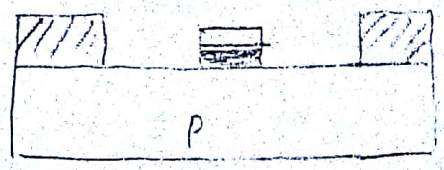


By chemical vapour deposition (CVD), a layer of polysilicon is deposited on the top of silicon

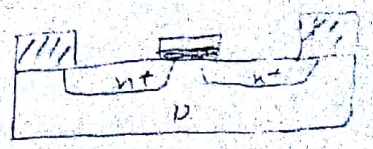
→ Retain the polysilicon only over the gate and remove it from the rest of the region by etching process



→ Open windows for source and drain. A quick dip in hydrofluoric acid will remove thin oxide



→ Realise source and drain



2(b)

CMOS Inverter: -

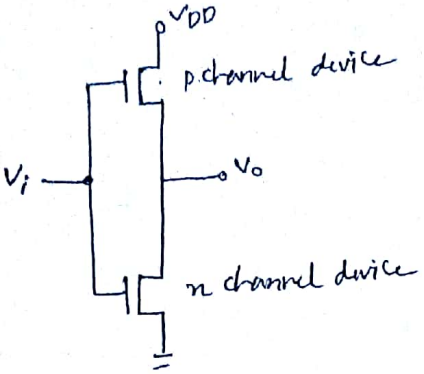


Fig: CMOS Inverter

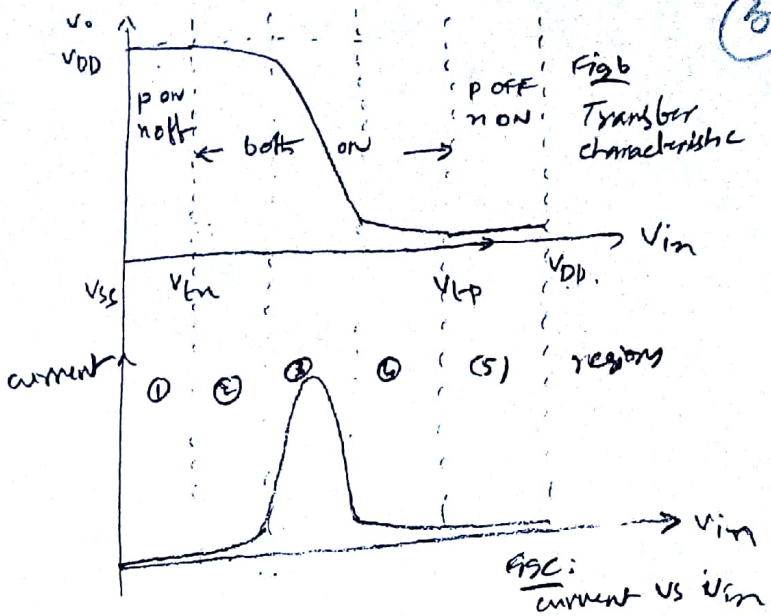


Fig b Transfer characteristic

Fig c: current vs Vin

The current voltage relationships for the mos transistor are

$$I_{ds} = K\frac{W}{L}\left[(V_{gs}-V_t)V_{ds} - \frac{V_{ds}^2}{2}\right] \quad ; \text{ in the resistive region}$$

$$I_{ds} = K\frac{W}{L}\frac{(V_{gs}-V_t)^2}{2} \quad ; \text{ in the saturation region}$$

In both cases the factor K is a technology dependent parameter such that $K = \frac{\epsilon_{ins}\epsilon_0\mu}{D}$

The factor $\frac{W}{L}$ is contributed by geometry.

It is common practice to write $\beta = K\frac{W}{L}$

β may be applied to both nmos and pmos transistors as follows:

$$\beta_n = \frac{\epsilon_{ins}\epsilon_0\mu_n}{D}\frac{W_n}{L_n}$$

$$\beta_p = \frac{\epsilon_{ins}\epsilon_0\mu_p}{D}\frac{W_p}{L_p}$$

where W_n and L_n, W_p and L_p are the n and p transistors dimensions respectively

CMOS inverter has five distinct regions of operation which are shown in the above figure.

In region 1 for which V_{in} = logic 0, the p transistor is fully turned on and the n transistor is fully turned off. Thus no current flows through the inverter and the output is directly connected to V_{DD}.

In region 5, V_{in} = logic 1, n transistor is fully on while the p-transistor is fully off. Again, no current flows and a good logic 0 appears at the output.

In region 2 the input voltage has increased to a level which just exceeds the threshold voltage of the n-transistor. The n transistor conducts and has a large voltage between source and drain; so it is in saturation. The p transistor is also conducting but with only a small voltage across it, it operates in the unsaturated resistive region. A small current now flows through the inverter from V_{DD} to V_{SS}.

Region 4 is similar to region 2 but with the roles of p and n transistors reversed. However, the current magnitudes in regions 2 and 4 are small and most of the energy consumed in switching from one-state to the other is due to the larger current which flows in region 3.

Region 3 is the region in which the inverter exhibits gain and in which both transistors are in saturation.

The currents in each device must be same since the transistors are in series.

$$I_{dsp} = -I_{dsn}$$

$$\frac{\beta_p}{2}\left(V_{in} - V_{DD} - V_{tp}\right)^2 = -\frac{\beta_n}{2}\left(V_{in} - V_{tn}\right)^2$$

$$V_{in} - V_{DD} - V_{tp} = -\sqrt{\frac{\beta_n}{\beta_p}}\left(V_{in} - V_{tn}\right)$$

$$V_{in} - V_{DD} - V_{tp} = -\sqrt{\frac{\beta_n}{\beta_p}}\,V_{in} + \sqrt{\frac{\beta_n}{\beta_p}}\,V_{tn}$$

$$\left(1 + \sqrt{\frac{\beta_n}{\beta_p}}\right)V_{in} = V_{DD} + V_{tp} + V_{tn}\sqrt{\frac{\beta_n}{\beta_p}}$$

$$V_{in} = \frac{V_{DD} + V_{tp} + V_{tn}\sqrt{\beta_n/\beta_p}}{1 + \sqrt{\beta_n/\beta_p}}$$

In region 3, since both transistors are in saturation, they act as current sources so that the equivalent circuit in this region is two current sources in series between V_{DD} and V_{SS} with the output voltage coming from their common point.

the region is inherently unstable in consequence and the changeover from one logic level to the other is rapid.

If $\beta_n = \beta_p$ and if $V_{tn} = -V_{tp}$, then from the above equation

$$V_{in} = 0.5\, V_{DD}$$

This implies that the changeover between logic levels is symmetrically disposed about the point at which

$$V_{in} = V_{out} = 0.5\, V_{DD}$$

3(a)

Drain to source current I_{DS} versus voltage V_{ds} relationship

Gate voltage induces charge in the channel region between source and drain. The electric field created by the voltage V_{ds}, moves this charge from source to drain.

current from drain to source, $I_{ds} = \dfrac{\text{charge induced in channel } (Q_c)}{\text{Electron transit time } (\tau)}$ → (1)

Transit time, $\tau_{sd} = \dfrac{\text{length of the channel } (L)}{\text{velocity } (v)}$

but velocity, $v = \mu E_{ds}$

where μ = electron & hole mobility (surface)

E_{ds} = electric field (drain to source)

Now $E_{ds} = \dfrac{V_{ds}}{L}$

$\therefore v = \mu \dfrac{V_{ds}}{L}$

Thus, $\tau_{sd} = \dfrac{L}{\mu \frac{V_{ds}}{L}} = \dfrac{L^2}{\mu V_{ds}}$ → (2)

The non saturated region:-

Charge induced in channel due to gate voltage is due to the voltage difference between the gate and the channel, V_{gs} (assuming substrate connected to source).

The voltage along the channel varies linearly with distance x from the source due to IR drop in the channel and assuming that the device is not saturated then the average value is $\dfrac{V_{ds}}{2}$.

Furthermore, the effective gate voltage $V_g = V_{gs} - V_t$ where V_t = threshold voltage.

Note that charge/unit area $= E_g \epsilon_{ins} \epsilon_0$

where E_g = average electric field gate to channel

ϵ_{ins} = relative permittivity of insulator between gate and channel.

ϵ_0 = permittivity of free space

Thus induced charge,

$Q_c = E_g \epsilon_{ins} \epsilon_0 WL$ where W = width of the channel

L = length of the channel

Now $E_g = \dfrac{\left[(V_{gs} - V_t) - \frac{V_{ds}}{2}\right]}{D}$ where D = oxide thickness

Thus $Q_c = WL \epsilon_{ins} \epsilon_0 \dfrac{\left[(V_{gs} - V_t) - \frac{V_{ds}}{2}\right]}{D}$ → (3)

using equations (2) and (3), eqn (1) can be written as

$$I_{ds} = WL \epsilon_{ins} \epsilon_0 \frac{\left[(V_{gs} - V_t) - \frac{V_{ds}}{2}\right]}{D} \Big/ \frac{L^2}{\mu V_{ds}}$$

$$I_{ds} = \frac{\epsilon_{ins}\epsilon_0\mu}{D}\frac{W}{L}\left[(V_{gs}-V_t) - \frac{V_{ds}}{2}\right]V_{ds}$$

Let $K = \frac{\epsilon_{ins}\epsilon_0\mu}{D}$

$$\therefore I_{ds} = K\frac{W}{L}\left[(V_{gs}-V_t)V_{ds} - \frac{V_{ds}^2}{2}\right] \rightarrow (4)$$

$$= \beta\left[(V_{gs}-V_t)V_{ds} - \frac{V_{ds}^2}{2}\right] \text{ where } \beta = K\frac{W}{L}$$

The gate/channel capacitance is given by $C_g = \frac{\epsilon_{ins}\epsilon_0 WL}{D}$

∴ The above relation shown be shown as

$$I_{ds} = \frac{C_g\mu}{L^2}\left[(V_{gs}-V_t)V_{ds} - \frac{V_{ds}^2}{2}\right] \rightarrow 4(a)$$

The Saturated region :-

Saturation begins when $V_{ds} = V_{gs} - V_t$ since at this point the IR drop in the channel equals the effective gate to channel voltage at the drain and we may assume that the current remains fairly constant as V_{ds} increases further. Thus

$$I_{ds} = K\frac{W}{L}\left[(V_{gs}-V_t)V_{ds} - \frac{V_{ds}^2}{2}\right]$$

$$= K\frac{W}{L}\left[(V_{gs}-V_t)^2 - \frac{(V_{gs}-V_t)^2}{2}\right]$$

$$= K\frac{W}{L}\frac{(V_{gs}-V_t)^2}{2} \rightarrow (5)$$

or we may write

$$I_{ds} = \frac{\beta}{2}(V_{gs}-V_t)^2$$

&

$$I_{ds} = \frac{C_g\mu}{2L^2}(V_{gs}-V_t)^2$$

The expressions derived for I_{ds} hold for both enhancement and depletion mode devices.

3(b) (10)

Alternate forms of pull up :-

There are four possible arrangements

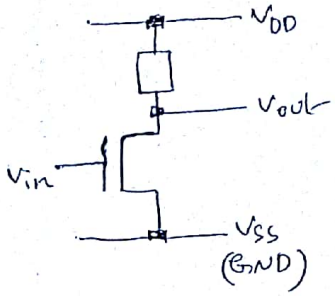
(i) Load resistance R_L

(ii) nMOS depletion mode transistor pull up

(iii) nMOS enhancement mode pull up

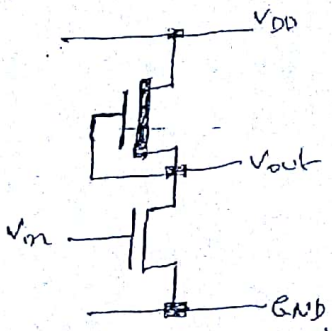
(iv) Complementary transistor pull up (cMOS)

(i) Load resistance R_L :-



This arrangement is not often used because of the large space requirement of resistors produced in a silicon substrate.

(ii) nMOS depletion mode transistor pull up :-

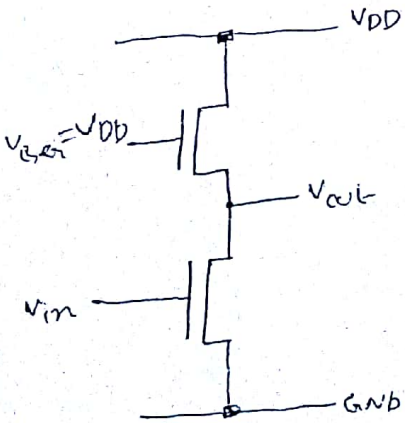


→ Dissipation is high since rail to rail current flows when V_{in} = logical 1.

→ Switching of output from 1 to 0 begins when V_{in} exceeds V_t of pull down device.

→ When switching the output from 1 to 0, the pull up device is non saturated initially and this presents lower resistance through which to charge capacitive loads.

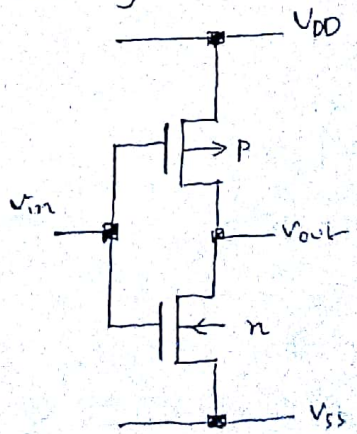
(iii) nMOS Enhancement mode pull up :-



→ Dissipation is high since current flows when V_{in} = logical 1

→ V_{out} can never reach V_{DD} (logical 1) if $V_{Gen} = V_{DD}$ as is normally the case

(iv) Complementary transistor pull up (cmos)



→ No current flow either for logical 0 or for logical 1 inputs

→ Full logical 1 and 0 levels are presented at the output.

→ For devices of similar dimensions the p-channel is slower than the n-channel device.

(d) Wiring Capacitances :-

Earlier, the area capacitances associated with the layers to substrate and beam gate to channel are considered.

However, there are other significant sources of capacitance which contribute to the overall wiring capacitance.

Three such sources are

(i) Fringing Fields :- The fringe field is the peripheral magnetic field outside of the magnetic core.

Capacitance due to fringing field effects can be a major component of the overall capacitance of interconnect wires.

For fine line metallization, the value of fringing field capacitance (C_{ff}) can be of the same order as that of the area capacitance.

Thus, C_{ff} should be taken into account if accurate prediction of performance is needed.

$$C_{ff} = \epsilon_{SiO_2} \epsilon_0 \, l \left[\frac{\pi}{\ln\left\{1 + \frac{2d}{t}\left(1 + \sqrt{1 + \frac{t}{d}}\right)\right\}} - \frac{t}{4d} \right]$$

Where l = wire length
t = thickness of wire
d = wire to substrate separation

Then, total wire capacitance

$$C_W = C_{area} + C_{ff}$$

(ii) Interlayer Capacitances :-

Quite obviously the parallel plate effects are present between one layer and another.

For example, for a given area, metal to polysilicon capacitance must be higher than metal to substrate.

The reason for not taking such effects into account for simple calculations is that the effects occur only when layers cross or when one layer underlies another, and in consequence interlayer capacitance is highly dependent on lay out. However, for regular structures it is readily calculated and contributes significantly to the accuracy of circuit modeling and delay calculation.

iii) Peripheral Capacitance :-

The source and drain n-diffusion regions (n-active regions for orbit process) form junctions with the p-substrate or p-well at well defined and uniform depths;

Similarly for p-diffusion (p-active) regions in n-substrates or n-wells

For diffusion regions, each diode thus formed has associated with it a peripheral (side-wall) capacitance in picofarads per unit length, which, in total, can be considerably greater than the area capacitance of the diffusion region to substrate;

The smaller the source to drain area, the greater becomes the relative value of the peripheral capacitance.

For orbit processes, the n-active and p-active regions are formed by impurity implant at the surface of the silicon and thus, having negligible depth, they have negligible peripheral capacitance.

However, for n- and p-regions formed by a diffusion process, the peripheral capacitance is important and becomes particularly so as we shrink the device dimensions.

In order to calculate the total diffusion capacitance, we must add the contributions of area and their peripheral components.

$$C_{total} = C_{area} + C_{periph}$$

Table : Typical values of for diffusion capacitances

Diffusion Capacitance	Typical values		
	8 μm	2 μm	1.2 μm
Area C (C_{area})	1.0×10^{-4} pF/μm²	1.75×10^{-4} pF/μm²	3.75×10^{-4} pF/μm²
Periphery (C_{periph})	8.0×10^{-4} pF/μm²	negligible *	negligible *

* Assuming implanted regions of negligible depth.

4(b)

CMOS Design Methodology :-

It has three steps

1. Given the Boolean expression, take its complement.
2. Design Pull down network (PDN) by realizing
   - AND terms using series connected nMOSFETs
   - OR terms using parallel connected nMOSFETs.
3. Design Pull up network just reverse (& dual) of the PDN.
   - AND terms using parallel connected pMOSFETs
   - OR terms using series connected pMOSFETs.

Design of two input AND gate using CMOS Technology

If A and B are inputs and Y is the output, $Y = AB$

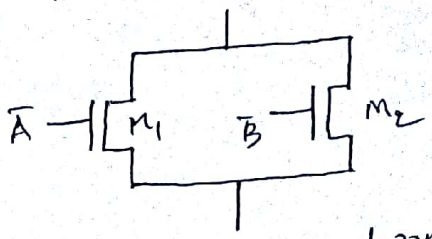
As per cmos methodology

1. Take Complement of Y.

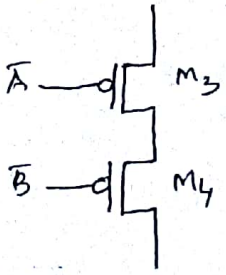
$$\bar{Y} = \overline{AB}$$

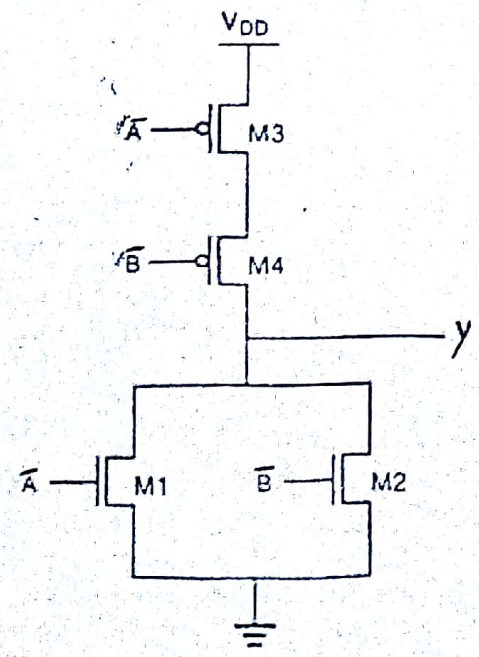
$$= \bar{A} + \bar{B}$$

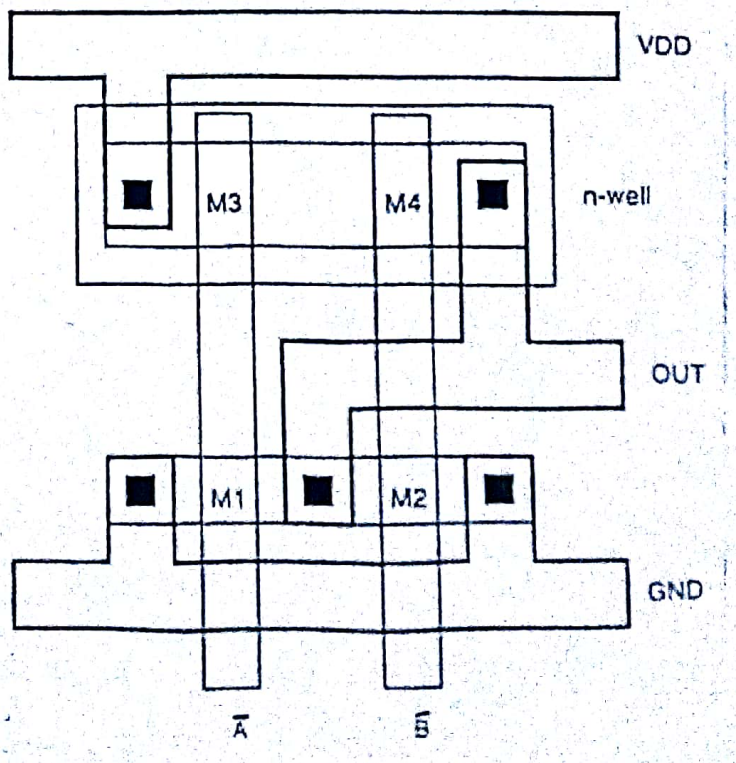
2. Design the Pull down Network.
   There is only one OR term.
   Use parallel connected nMOSFETs



3. Design the pull up network.
   There is only one OR term.
   Use series connected pMOSFETs.







λ based Design Rules:-

Features of λ-based Design rules:

→ λ is the size of a minimum feature

→ All the dimensions are specified in integer multiple of λ

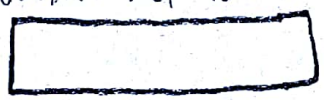
→ Specifying λ particularizes the scalable rules

→ parasitics are generally not specified in λ units.

→ These rules specify geometry of masks, which will provide reasonable yields.

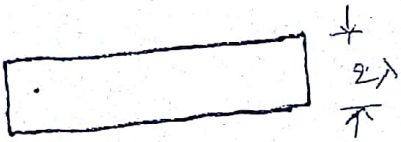
Guidelines for using λ-based Design rules:

Diffusion not lower than 2λ



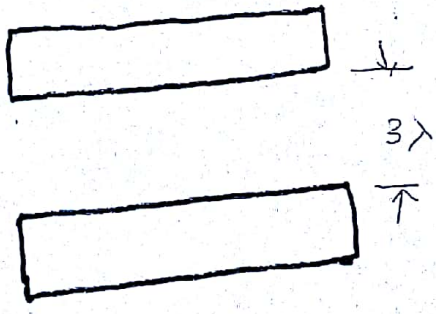
; minimum width of diffusion is 2λ

poly not lower than 2λ



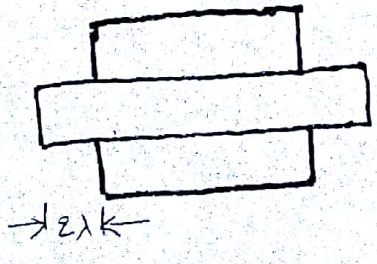
; minimum width of poly is 2λ

Diffusion and diffusion not lower than 3λ



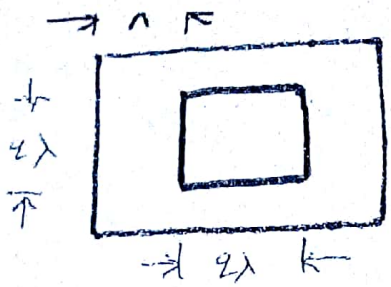
, minimum distance between two diffusion layers 3λ

poly cross diffusion



It is necessary for the poly to completely cross active, otherwise the transistor that has been created crossing of diffusion and poly, will be shorted by diffused part of source and drain.

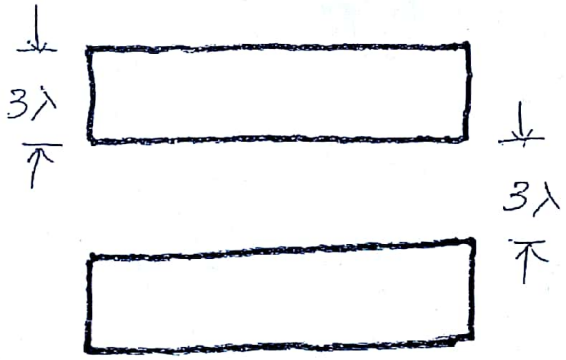
Contact Cut on Metal



Contact window will be of 2λ by 2λ that is minimum feature size while metal deposition is of 4λ by 4λ for reliable contact.

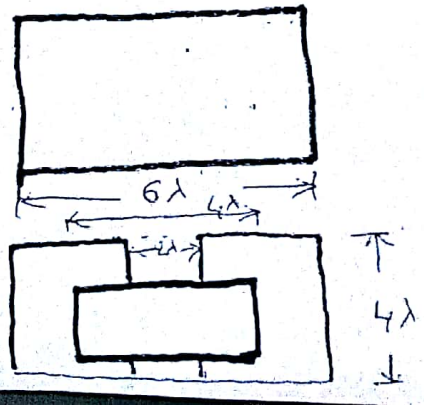
On Metal:

Two metal wires have 3λ distance between them to overcome capacitance coupling and high frequency coupling. Metal wires width can be as large as possible to decrease resistance.

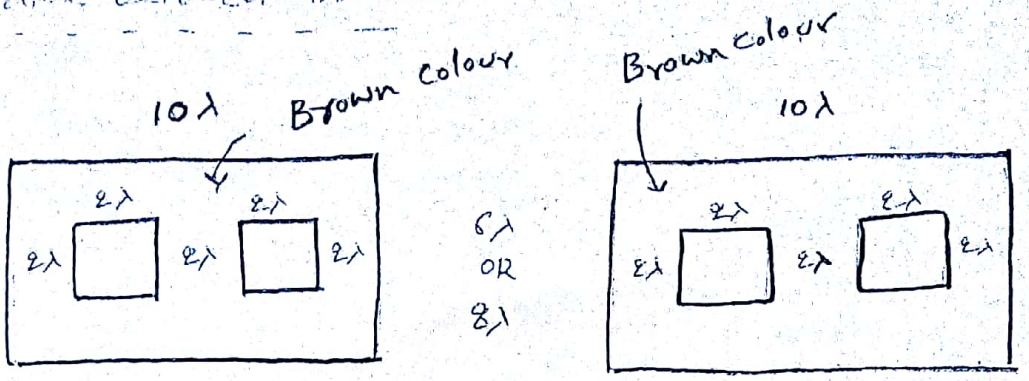


Buttering Contact:-

Buttering contact is used to make poly and silicon contact. Window's original width is 4λ, but on overlapping width is 2λ. So actual contact area is 6λ by 4λ.



The distance between two Well



The distance between two wells depends on the well potentials as shown above.

~~The reason for 8λ is that if both wells are at same~~

The reason for 6λ is that if both wells are at same high potential then the depletion region between them may touch each other causing punch-through.

The reason for 6λ is that if both wells are at different potentials then depletion region of one well will be smaller, so both depletion regions will not touch each other so 6λ will be good enough.

The active region has length 10λ which is distributed over the followings

→ 2λ for source diffusion
→ 2λ for drain diffusion
→ 2λ for channel length
→ 2λ for source side encroachment
→ 2λ for drain side encroachment

5(b)

Scaling factors for Device parameters :-

In this section, simple derivations and calculations reveal the effects of scaling.

(1) Gate Area A_g :-

$$A_g = L W$$

where L and W are the channel lengths and widths respectively. Both are scaled by $\frac{1}{\alpha}$.

Thus A_g is scaled by $\frac{1}{\alpha^2}$

(2) Gate Capacitance per unit area C_0 or C_{ox}

$$C_0 = \frac{\epsilon_{ox}}{D}$$

where ϵ_{ox} = permittivity of the gate oxide (thinox) $\left[= \epsilon_{ins} \epsilon_0 \right]$

and D = gate oxide thickness.

Here D is scaled by $\frac{1}{\beta}$

Thus C_0 is scaled by $\frac{1}{(1/\beta)} = \beta$

(3) Gate Capacitance C_g :-

$$C_g = C_0 L W$$

Thus C_g is scaled by $\beta \frac{1}{\alpha^2} = \frac{\beta}{\alpha^2}$

(4) Parasitic Capacitance C_x :-

C_x is proportional to $\frac{A_x}{d}$

where d = depletion width around source or drain

d is scaled by $\frac{1}{\alpha}$

A_x = area of the depletion region around source or drain.

A_x is scaled by $\frac{1}{\alpha^2}$

Thus C_x is scaled by $\frac{1}{\alpha^2} \cdot \frac{1}{(1/\alpha)} = \frac{1}{\alpha}$

(5) Carrier density in channel Q_{on} :-

$$Q_{on} = C_0 V_{gs}$$

where Q_{on} is the average charge per unit area in the channel in the on state. Note that C_0 is scaled by β and V_{gs} is scaled by $\frac{1}{\beta}$

Thus Q_{on} is scaled by 1

(6) Channel Resistance R_{on} :-

$$R_{on} = \frac{L}{W} \frac{1}{Q_{on}\mu}$$

where μ is the carrier mobility in the channel and it assumed constant.

Thus R_{on} is scaled by $\frac{1}{\alpha} \frac{1}{1/\alpha} 1 = 1$

(7) Gate Delay T_d :-

T_d is proportional to $R_{on} C_g$

Thus T_d is scaled by $1 \frac{\beta}{\alpha^2} = \frac{\beta}{\alpha^2}$

(8) Maximum operating frequency f_0

$$f_0 = \frac{W}{L} \frac{\mu C_0 V_{DD}}{C_g}$$

& f_0 is inversely proportional to delay T_d.

Thus f_0 is scaled by $\frac{\alpha^2}{\beta}$

(9) Saturation current I_{dss} :-

$$I_{dss} = \frac{C_0 \mu}{2} \frac{W}{L} (V_{gs} - V_t)^2$$

Noting that both V_{gs} and V_t are scaled by $\frac{1}{\beta}$, we have

I_{dss} scaled by $\beta \left(\frac{1}{\beta}\right)^2 = \frac{1}{\beta}$

(10) Current density J :-

$$J = \frac{I_{DSS}}{A}$$

where A is the cross-sectional area of the channel in the 'on' state which is scaled by $\frac{1}{\alpha^2}$.

So, J is scaled by $\frac{1/\beta}{1/\alpha^2} = \frac{\alpha^2}{\beta}$

(11) Switching energy per gate E_g :-

$$E_g = \frac{1}{2} C_g (V_{DD})^2$$

So, E_g is scaled by $\frac{\beta}{\alpha^2} \frac{1}{\beta^2} = \frac{1}{\alpha^2 \beta}$

6(a)

Architectural issues :-

In the design of a VLSI system, a logical and systematic approach is essential.

The design time tend to rise exponentially with increased complexity.

Guidelines in designing a VLSI System

1. Define the requirements (properly and carefully).
2. Partition the overall architecture into appropriate subsystems
3. ~~Consider Communication paths~~ carefully in order to develop sensible interrelationships between subsystems.
4. Draw a floor plan of how the system is to map onto the silicon
5. Aim for regular structures so that design is largely a matter of replication.
6. Draw suitable (stick & symbolic) diagrams of the leaf-cells of the subsystems.
7. Convert each cell to a layout
8. Carefully and thoroughly carry out a design rule check on each cell.
9. Simulate the performance of each cell/subsystem.

The whole design process will be greatly assisted if considerable care is taken with

(1) the partioni partitioning of the system so that there are clean or clear subsystems with a minimum interdependence and complexity of interconnection between them.

(2) the design simplification within subsystems so that architectures are adopted which allow the exploitation of a cellular design concept. This allows the system to be composed of relatively few standard cells which are replicated to form highly regular structures.

6(b)

(20)

Design of an ALU subsystem

The heart of the ALU is a 4 bit adder circuit. A 4 bit adder must take the sum of two 4 bit numbers.

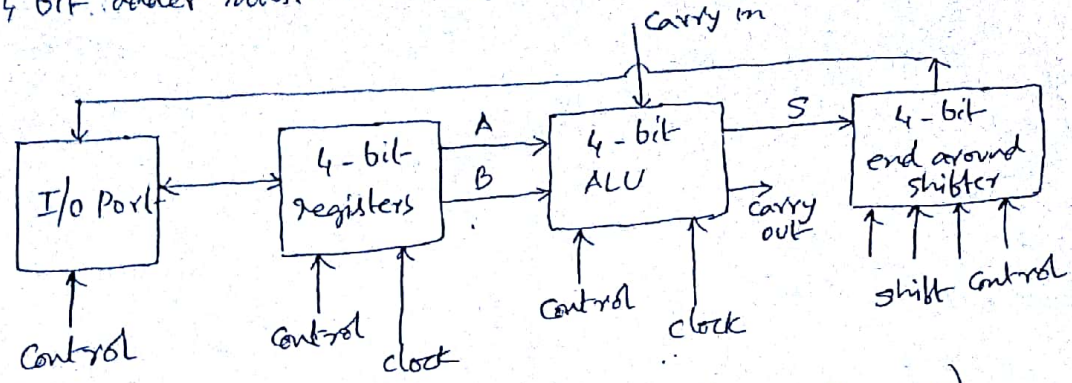


Fig 8.1: 4 bit data path for processor (block diagram)

The shifter circuit accepts and shifts a 4 bit parallel sum from the ALU.

The sum is to be stored in parallel at the output of the adder from where it may be fed through the shifter and back to the register array.

Thus, a single 4-bit data bus is needed from the adder to the shifter and another 4-bit bus is required from the shifted output back to the register array (since the shifter is merely a switch array with no storage capability).

As far as the input to the adder is concerned, the two 4 bit parallel numbers to be added are to be presented in parallel on two 4 bit buses.

Assume clock phase ϕ_1 as being the phase during which signals are fed along buses to the adder input and during which their sum is stored at the adder output.

Thus clock signals are required by the ALU as shown.

The shifter is unclocked but must be connected to four shift control lines. It is also necessary to provide a 'carry out' signal from the adder and, in the general case, to provide for a possible 'carry in' signal, as indicated in the above fig.

7(a)

Switch logic and gate logic

Features of switch logic are

→ It is based on pass transistors or on transmission gates
→ It is better for small arrays
→ It takes no static current from the supply rails. The power dissipation is low.
→ It is similar to logic arrays based on relay contacts
→ Logic levels propagated through pass transistors are degraded by threshold voltage effects.

→ pass transistors have significant disadvantage in that load presented to the inputs depend on the load connected to the output, and output rise/fall times are a strong function of input rise/fall times. This makes timing analysis difficult. As a result pass transistors tend to be only used in localized environments where the input drivers and output loads are well understood.
→ pass transistor logic often uses fewer transistors.

Gate logic is based on the general arrangement typified by the inverter circuits (the inverter being the simplest gate.

The drawbacks of switch logic can be overcome with gate logic but requires more transistors.

7(b)

Subsystem Design Considerations of a 4 bit multiplier

- Lower unit cost
- Higher reliability
- Lower power dissipation, lower weight and lower volume.
- Better performance - particularly in terms of speed power product.
- Enhanced repeatability
- Reduced design/development periods.

If the subsystems are regular and therefore composed of relatively few actual leaf cells (circuits), then the designer can concentrate on the main problem of VLSI design - the routing of interconnections and communication paths.

Array multiplier

The design of an unsigned array multiplier is based on the simple pen and paper method of multiplication - taking one bit at a time from the multiplier, the partial products are generated by multiplying each bit to the multiplicand.

These partial products are then written in rows, by shifting each row by one bit position to the right.

These rows are then added column-wise. This multiplier has a very regular architecture and is most suitable for VLSI implementation.

The following fig illustrates an unsigned array multiplier.

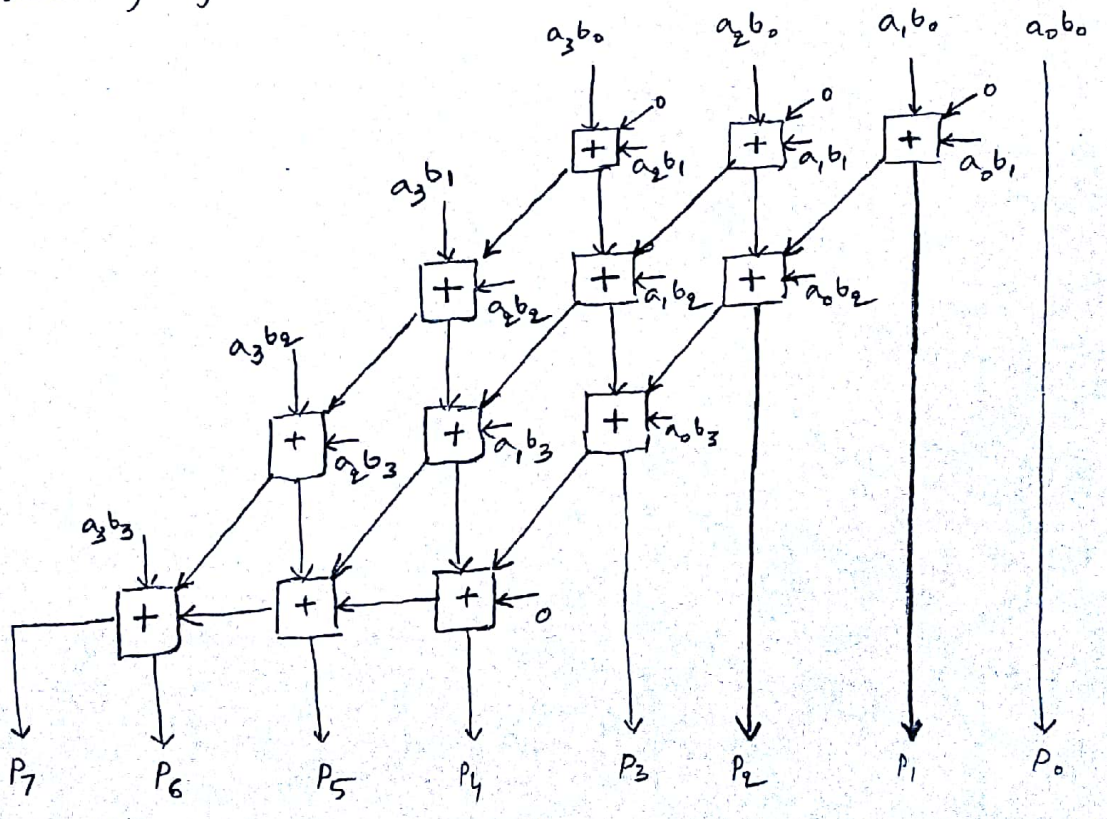


Fig: Unsigned array multiplier

8(a)

Full Adder using PLA

Truth table for Full Adder

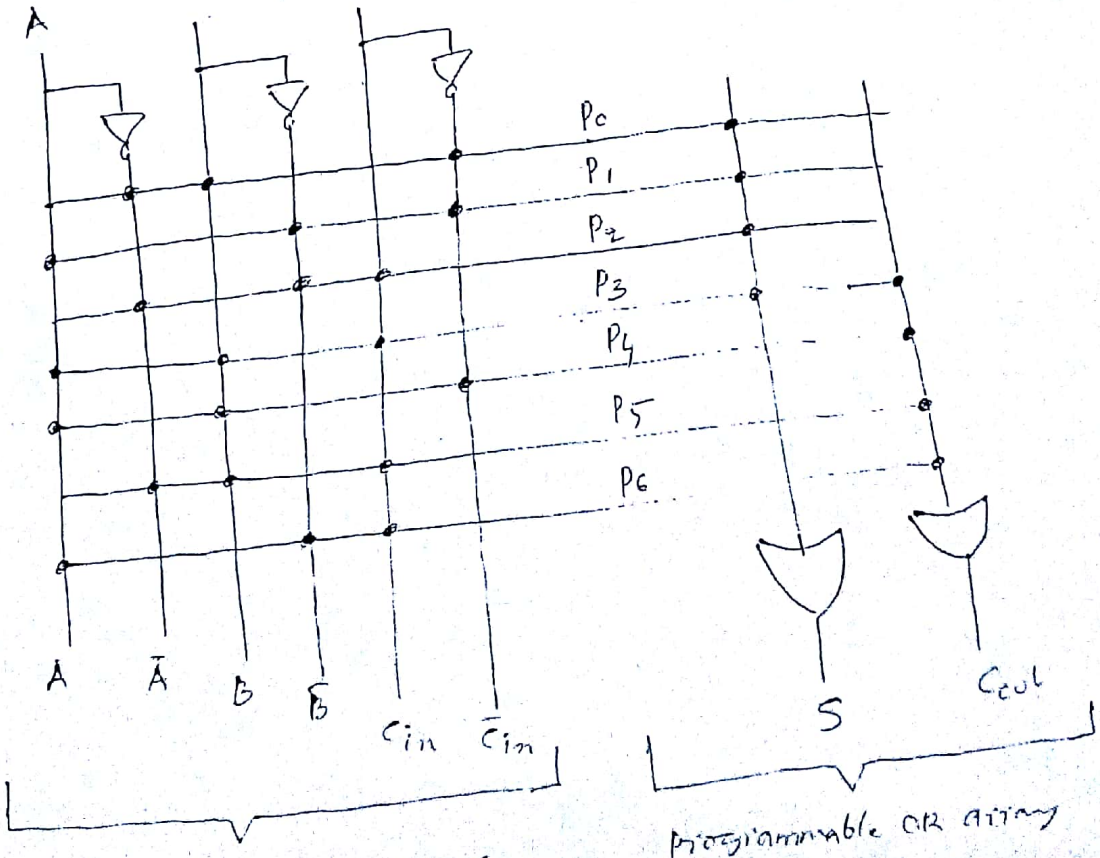
S = Sum
C = Carry

Inputs			outputs	
A	B	C_{in}	S	C_{out}
0	0	0	0	0
0	1	0	1	0
1	0	0	1	0
1	1	0	0	1
0	0	1	1	0
0	1	1	0	1
1	0	1	0	1
1	1	1	1	1

From the truth table

$$S = \underbrace{\bar{A}B\bar{C}_{in}}_{P_0} + \underbrace{A\bar{B}\bar{C}_{in}}_{P_1} + \underbrace{\bar{A}\bar{B}C_{in}}_{P_2} + \underbrace{ABC_{in}}_{P_3}$$

$$C_{out} = \underbrace{AB\bar{C}_{in}}_{P_4} + \underbrace{\bar{A}BC_{in}}_{P_5} + \underbrace{A\bar{B}C_{in}}_{P_6} + \underbrace{ABC_{in}}_{P_3}$$



8(b)

Full form of ASIC is Application Specific Integrated Circuit.

Types of ASICs :-

(i) Full custom ASICs
(ii) Standard cell based ASICs
(iii) Gate array based ASICs
(iv) Programmable logic devices
(v) Field Programmable Gate Arrays

In a full custom ASIC, an engineer designs some or all of the logic cells, circuits or layout specifically for one ASIC.

Standard cell based ASIC uses predesignated logic cells (AND gates, OR gates, multiplexers and flip flops for example) known as standard cells.

In a gate array based ASIC, the transistors are predefined on the silicon wafer. Both cell based and gate array ASICs use predefined cells, but there is a difference - we can change the transistor sizes in a standard cell to optimize speed and performance, but the device sizes in a gate array are fixed.

Programmable logic devices (PLDs) are standard ICs that are available in standard configurations from a catalog of parts and are sold in very high volume to many different customers. These are subdivided in ROM, programmable array logic (PAL), and programmable logic array (PLA).

A step above the PLD in complexity is the field programmable gate array (FPGA). FPGA is an IC which can be hardware programmed to implement various logic functions.

9(a)

Structural Design elements

The body of an architecture is a series of concurrent statements.

In VHDL, each concurrent statement executes simultaneously with the other concurrent statements in the same architecture body.

Thus, in a VHDL architecture body, if the last statement updates a signal that is used by the first statement, then the simulator will go back to that first statement and update its results according to the signal that just changed.

VHDL has several different concurrent statements, as well as a mechanism for bundling a set of sequential statements to operate as a single concurrent statement.

The most basic of VHDL's concurrent statements is the component statement, whose syntax is shown below

Syntax of a VHDL component statement

```
label : component_name port map (signal1, signal2, - - - , signaln);
label : component_name port map (port1 => signal1, port2 => signal2, ---, portn => signaln);
```

Component declaration is shown below

Syntax of a VHDL component declaration

```
component component_name
        port ( signal_names : mode signal_type;
               signal_names : mode signal_type;
                    :
               signal_names : mode signal_type);
end component;
```

A complex digital system is implemented in a hierarchical order. The whole circuit is partioned into blocks, and the blocks are then partitioned into sub-blocks This process is continued to a level where the leaf level of the last level is merely the basic logic gates. Therefore, any hardware description language supports this hierarchical design concept. In the VHDL, this hierarchical design concept is addressed by the structural modeling style. In this style of modeling, a bigger digital block is implemented by a set of smaller blocks of basic logic gates. The blocks of gates that are used to design the bigger block are called components. The components are connected using signals to implement the bigger block.

9(b)

4 bit Adder using Dataflow modeling


```
library IEEE;
use IEEE.std_logic_1164.all;
use ieee.std_logic_arith.all;
use ieee.std_logic_unsigned.all;

entity adder4bit is
   port (A, B : in std_logic_vector (3 downto 0);
         Cin : in std_logic;
         Cout : out std_logic;
         Sum : out std_logic_vector (3 downto 0));
   generic (n : integer := 4);
end adder4bit;
architecture Dataflow of adder4bit is
   signal result: std_logic_vector (4 downto 0);
   signal cin_s : std_logic_vector (3 downto 0);
begin
   cin_s <= (others => '0');
   result <= ('0' & A) + ('0' & B) + (cin_s & cin);
   Cout <= result(4);
   Sum <= result (3 downto 0);
end Dataflow;
```

Scheme & Solutions
Prepared by
B. V. Kumaraswamy
(B. V. KUMARASWAMY)
EIE Dept.
9441503437

Ch. Ho 27/3/2019
Signature of the HOD

H. Srinivas
External Examiner