Co-2: Understand and analyze delay, design rule and circuit layout of CMOS.

Learning Outcomes (LO)

Students will able to learn

LO-1 : Various CMOS Designs rules(Lambda design rules).

LO-2 : CMOS Circuit Layout.

LO-3: CMOS Delay (Propagation delay, linear delay, RC Delay)

LO-4 : Interconnects in CMOS.

Lo-5 : Power in CMOS.

(1)

Learning Objectives (1 0')

- Few definitions
- Delay (parasitic, RC, Linear)
- Interconnects in CMOS
- Power in CMOS

DELAY :- Delay in CMOS circuits can be of two types
- > Propagation delay
- > RC Delay

PROPAGATION DELAY :-

Suppose that we have a CMOS Inverter whose Output is Connected to some next stage circuits. To Speed the performance of our Circuit we apply a Step Voltage at the input.

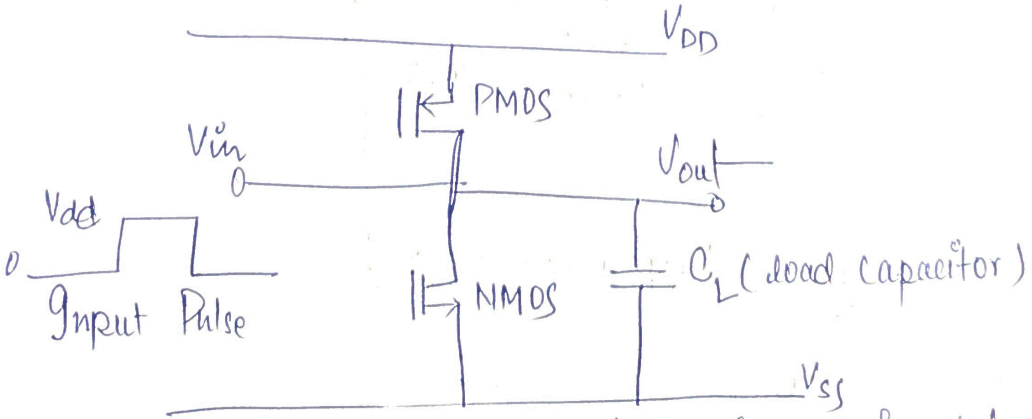


Fig:1 Capacitive load Connected to the Output terminal of CMOS Inverter.

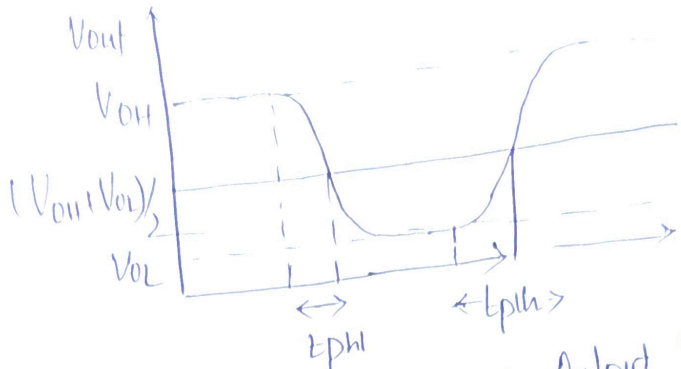


Fig: 2 Plot of the Output Voltage w.r.t time for a step input signal

In above plot we have two time intervals marked by tphl & tplh. here P stands for propagation delay.

HL → for high to low.

LH → for low to high.

⇒ Propagation delay for high to low is given by tphl & defined as time required for the Output to fall from VOH to (VOH + VOL)/2.

⇒ Propagation delay from low to high is given by tplh and defined as the time required for the Output to fall from VOL to (VOH + VOL)/2.

Hence the Propagation delay. (tp)

$$t_p = \frac{t_{phl} + t_{plh}}{2}$$

Factors Effecting Propagation delay:

Threshold Voltage:

With decrease in ~~Propag~~ threshold

Voltage the propagation delay increases.

Width of MOSFETS :-

Propagation delay decreases as we increases the values of MOSFET.

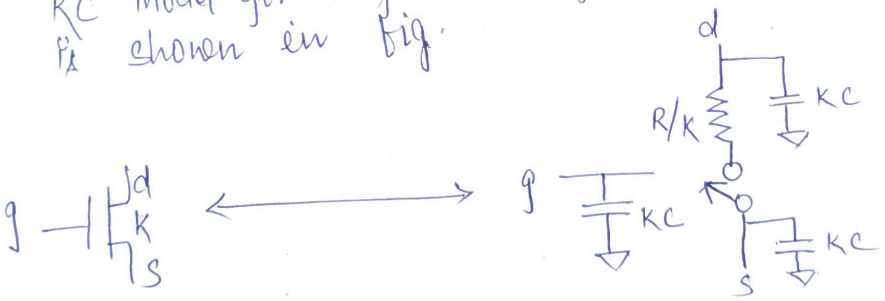
Supply Voltage :- The propagation delay has an inverse relation with the supply voltage. Thus increasing the supply voltage propagation delay increases.

Capacitive Load :- By increasing capacitive load propagation delay increases.

✡ RC DELAY :-

* Resistance & capacitance C of MOSFETs are very important to decide the transient response & performance of MOS circuits.

RC model for single nMOS of k times than unit size is shown in fig.

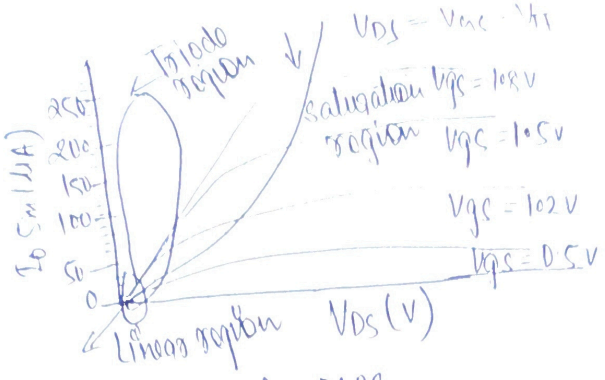


Here R_n is channel resistance of nMOS which is function of biasing voltage as given as

$$R_n = \left(\frac{V_{DS}}{I_D}\right)$$

But for region of nMOS

$V_{GS} > V_{TN}$ & $V_{DS} > 0$ but very small



⇒ But for linear region of nMOS

$V_{GS} > V_{TH}$ & $V_{DS} > 0$ but very small

Than $I_D = \mu_n C_{ox} (W/L)(V_{GS} - V_{TH}) V_{DS}$

hence $\boxed{R_n = 1/[\mu_n C_{ox}(W/L)(V_{GS} - V_{TH})]}$

⇒ Triode region of nMOS

$V_{GS} > V_{TH}$ & $0 < V_{DS} < V_{GS} - V_{TH}$

than I_D is given as

$$\boxed{I_D = \mu_n C_{ox}(W/L)\left[(V_{GS} - V_{TH})V_{DS} - \frac{V_{DS}^2}{2}\right]}$$

hence $\boxed{R_n = 2/[\mu_n C_{ox}(W/L) 2[V_{GS} - V_{TH} - V_{DS}]]}$

⇒ Saturation Region of nMOS is given as

$V_{GS} > V_{TH}$ & $V_{DS} > V_{GS} - V_{TH}$

Than $\boxed{I_D = \frac{1}{2}\mu_n C_{ox}(W/L)[V_{GS} - V_{TH}]^2}$

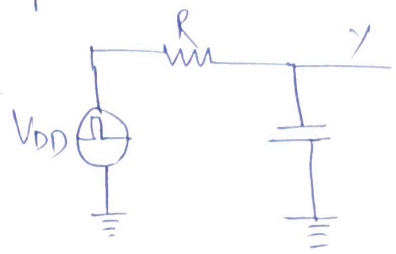
hence $\boxed{R_n = 2V_{DS}/[\mu_n C_{ox}(W/L)(V_{GS} - V_{TH})]^2}$

hence generalized model of channel resistance

$$\boxed{R_n = 1/[B_n(V_{DD} - V_{TH})]}$$

Transient Response from RC Delay Model :-

Once RC delay model is identified it can easily be analysed for the transient response.



The transfer function of 1st Order RC model

$$H(s) = \frac{1}{(1+sRC)}$$

By taking the Inverse Laplace transform of H(s), the step response.

$$V_y(t) = V_{DD}\, e^{-t/\tau} \quad \text{where } \tau = RC$$

Now propagation delay t_{pd} is the time at which Output increased. ie V_y reaches $V_{DD}/2$

Hence

$$V_{DD}/2 = V_{DD} e^{-t_{pd}/RC}$$

It gives $t_{pd} = RC \ln \longrightarrow t_{pd} \propto RC$

Hence delay increases by Increasing both the values of R & C.

Hence the RC delay models in CMOS circuit is used to model the timing delay.

Width, Spacing, Layer, Shielding

Widening a wire reduces resistance but increases capacitance (but less proportionally) → RC delay product improves

→ LAYERS

Metal 1 : Interconnects within cells

Metal 2/3 : Interconnects between cells within units

metal 4/5 : Interconnects between units Critical signals.

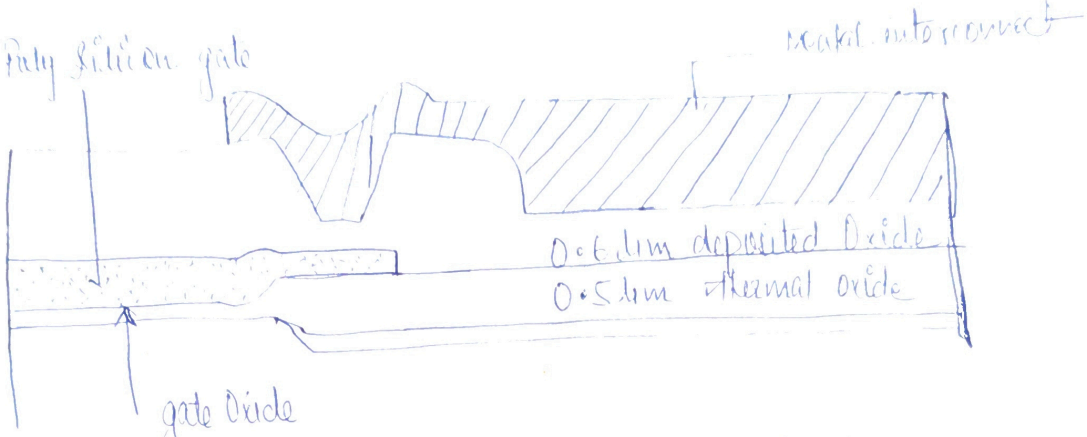
metal 6 : I/O pads, clock, power, ground.

AMI 0.6 um process has 3 metal layers
M.L for within - cell

PROPAGATION DELAY

INTERCONNECT :

* "Wires" consist of metal lines connecting the output of the inverter to the input of the next stage.



The P+ Layer (i.e heavily doped with acceptors) under the thick thermal oxide (500nm = 0.5mm) and deposited oxide (600nm = 0.6mm) depleted only slightly when positive voltages appear on the metal lines; so the capacitance is approximately the Oxide Capacitance.

$$C_{wire} = C_{thickox} (W_m^* L_m)$$

where Oxide thickness = 500nm + 600nm = 1.1 μm

In integrated Circuit (ICs), interconnects are structures that connect two or more circuit elements (such as transistors) together electrically.

In fabrication interconnects are formed during the back-end of line after the fabrication of transistor on the substrate.

⇒ Depending upon signal propogation Interconnects are classified as

LOCAL INTERCONNECTS :-

Local Interconnects connects circuit elements that are very close together, such as transistors seprated by ten or other transistor cantiguously.

GLOBAL INTERCONNECTS :-

Global Interconnects can transmit further, such as over large area sub-circuits.

Interconnect Resistance :-

ρ = resistivity ($\Omega * m$)

$R_{\square}$ = sheet resistance ($\Omega/\square$)

$\square$ = count number of square

$\square$ is dimensionless unit(!)

$R = R_{\square} * (\#\text{ of square})$

PROPAGATION DELAY

It is very important to determine the propagation delay in digital circuits as the shorter delay results in higher speed.

The dynamic behaviour of inverter is characterized in terms of the time delay between switching of input voltage (from Low to high & vice-versa) and the corresponding change appearing at the output. Such a delay called propagation delay arise for two reason

⟹ The transistor that implement the switches exhibit finite switching times

⟹ The capacitance that is inevitably present output reaches its required level of V_{OH} or V_{OL}.

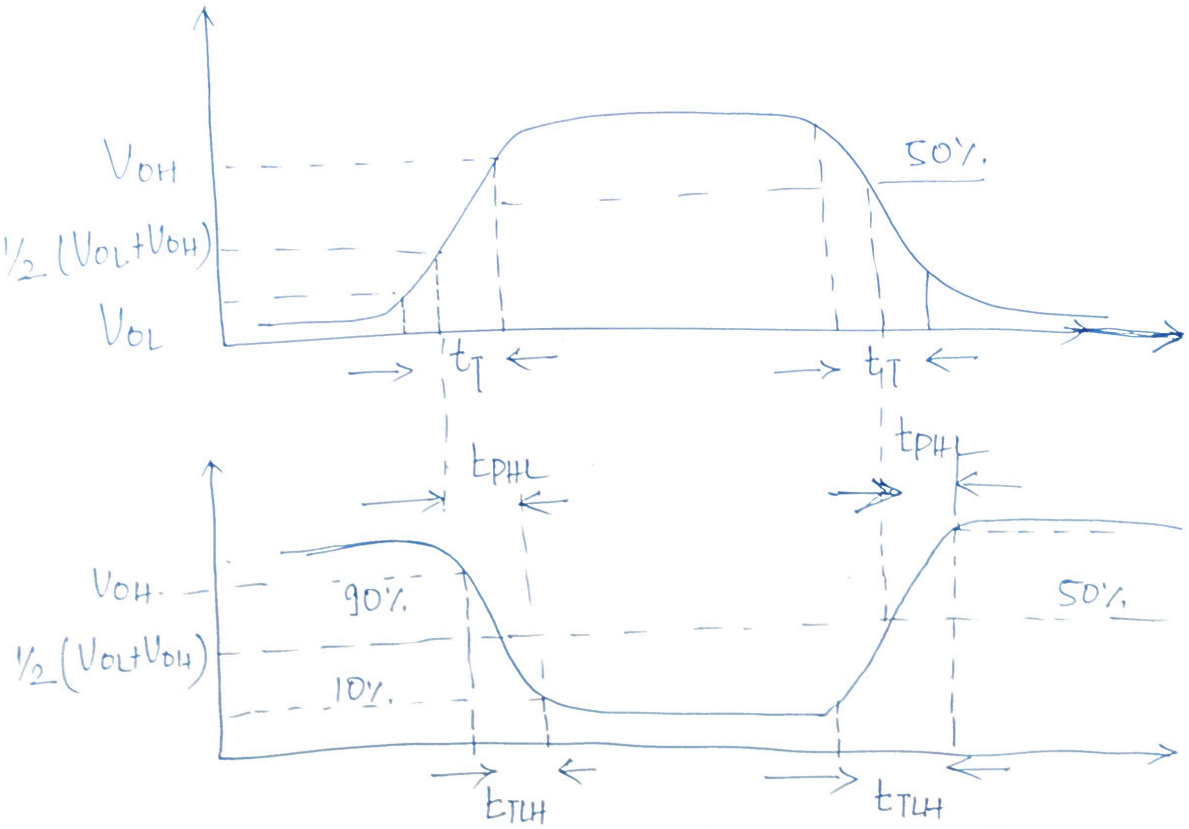


Fig: Propagation delay & transition time of logic Inverter

⟹ There is also a delay time between the input & output waveforms. The usual way to specify the propagation delay is take the average of high to low delays are measured between 50% points of the input & Output waveforms.

TRANSISTOR SIZING :-

⇒ In the designing of CMOS circuit we have to decide W/L ratio for all the devices. These ratio usually are selected to provide the gate with current driving capability in both direction which is equal to that of the basic Inverter.

⇒ In device sizing we find the input combination that result in the lowest Output current and then choose sizes that we will make this current equal to that of the basic Inverter.

⇒ We need to find W/L ratio is base of the resistance is inversely proportional to W/L.

If number of MOSFETs having ratio of $(W/L)_1, (W/L)_2, -$ $(W/L)_n$ then

$$r_{ds} = \frac{1}{K(W/L)(V_{GS}-V_t)}$$

$$r_{ds} = \frac{\text{constant}}{(W/L)}$$

$$\text{constant} = \frac{1}{K(V_{GS}-V_t)}$$

If MOSFET are connected in series than

$$R_{eq} = r_{ds1} + r_{ds2} + \cdots r_{dsn}$$

$$\frac{\text{constant}}{(W/L)_{eq}} = \frac{\text{constant}}{(W/L)_1} + \frac{\text{constant}}{(W/L)_2} + \cdots \frac{\text{constant}}{(W/L)_n}$$

$$\frac{1}{(W/L)_{eq}} = \frac{1}{(W/L)_1} + \frac{1}{(W/L)_2} + \cdots \frac{1}{(W/L)_n}$$

⇒ If MOSFET are connected in parallel then for $(W/L)_{eq}$.

$$\frac{1}{R_{eq}} = \frac{1}{r_{ds_1}} + \frac{1}{r_{ds_2}} + \cdots + \frac{1}{r_{ds_n}}$$

$$\frac{1}{\left[\frac{\text{constant}}{(W/L)_{eq}}\right]} = \frac{1}{\left[\frac{\text{constant}}{(W/L)_1}\right]} + \frac{1}{\left[\frac{\text{constant}}{(W/L)_2}\right]} + \cdots + \frac{1}{\left[\frac{\text{constant}}{(W/L)_n}\right]}$$

$$\Rightarrow \left(\frac{W}{L}\right)_{eq} = \left(\frac{W}{L}\right)_1 + \left(\frac{W}{L}\right)_2 + \cdots \left(\frac{W}{L}\right)_n$$

For consideration of $(W/L)_{eq}$, when all transistors are ON in both cases (series or ||).

⇒ When the transistors are connected in series, and all the transistors are ON.

⇒ When the transistors are connected in parallel, and any one is only ON.