

Lambda Based Design Rules.

Introduction: to design rules.

The physical mask layout of any circuit to be manufactured using a particular process must conform to a set of geometric constraints or rules which are generally called layout design rules.

These rules usually specify the minimum allowable line widths for physical objects on chip such as metal and polysilicon separations, interconnects of diffusion areas, minimum feature dimensions, & minimum allowable separations b/w two such features.

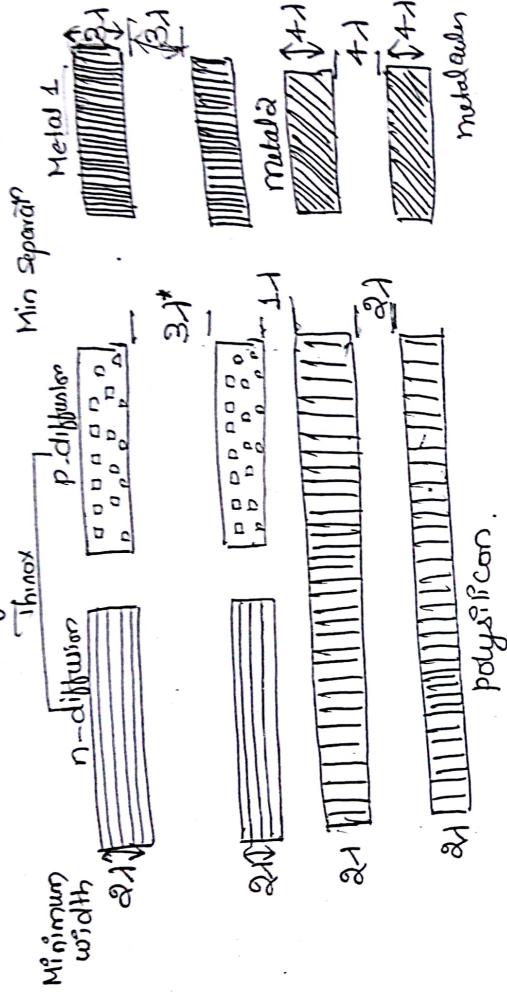
* The main objective of design rules is to achieve, for any circuit to be manufactured with a particular process, a high overall yield & reliability while using the smallest possible silicon area.

The design rules are usually described in two ways.

1. Micro rules: in which the layout constraints such as minimum feature sizes & minimum allowable feature separations are stated in terms of absolute dimensions in micrometre.

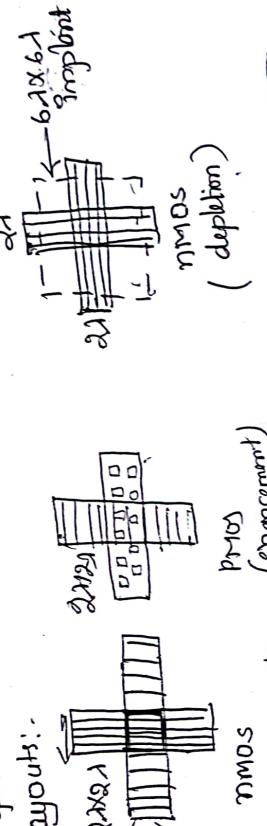
2. Lambda rules, which specify the layout constraints in terms of a single parameter (λ) & thus allow linear, proportional scaling of all geometrical constraints. Under these rules, all dimensions in all layers will be dimensioned in λ units & after that ' λ ' is given a value acceptable to the feature size of

> design rules also specify line widths, separations &
extensions in term of λ . These rules are drawn
indicated in fig. Eg: If $\lambda = 1\mu m$ + width $= 2\mu m$

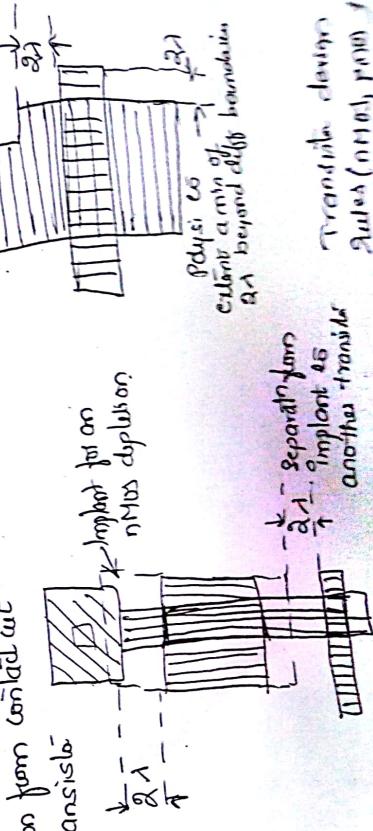


Design rules for wires (nmos + pmos)

Separations & extensions associated with transistors



Separation from contacts
to transistors



Contact Cuts:

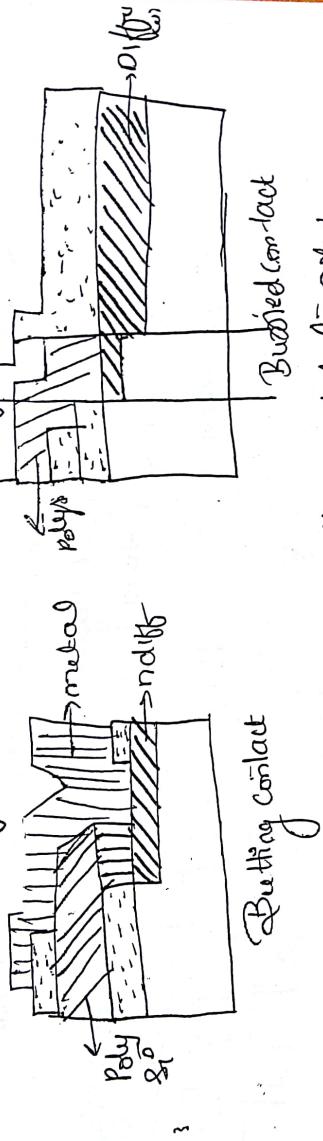
• Electrical connection can be established between layers by special structures called contact cuts.

• There are three possible approaches for establishing contacts below poly_x & diff_x in nMOS circuit.

i) Poly to metal & then metal to diffusion.

ii) A buried contact-poly in using metal.

iii) A butting contact-poly in using metal.



(i) Polysilicon to the metal & then metal to poly:

- oxide is removed from 2×2 contact cut down to the underlying poly_x wire. Then metal is deposited.
- the oxide etched area to the
- it flows through the oxide etched area to the polysilicon area.
- then polysilicon is deposited on the surface, which acts as conduction path.

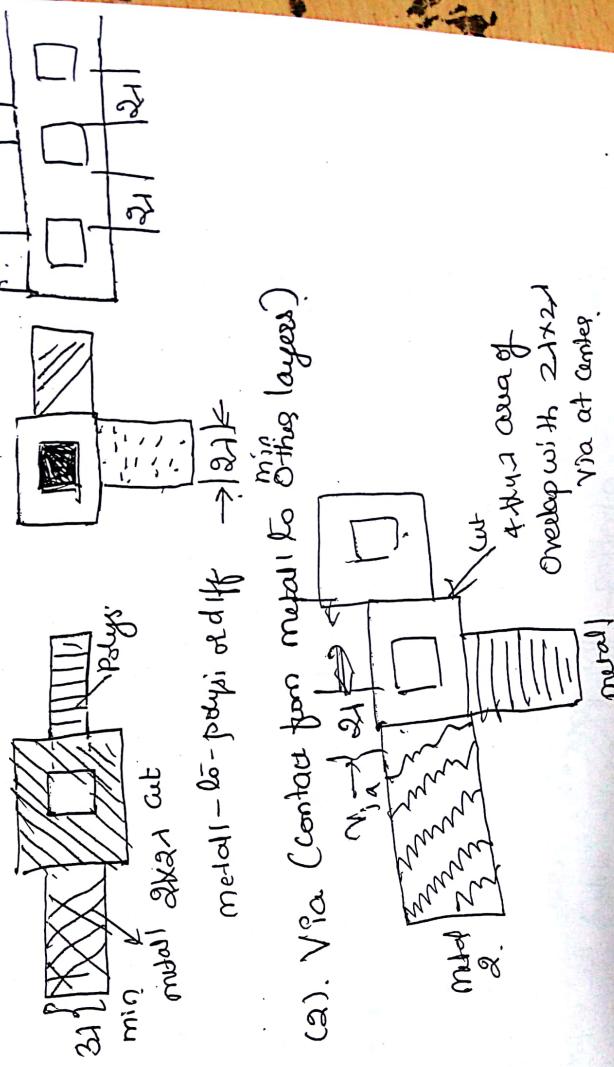
(ii) Buried contact:

Before starting the process, there is oxide layer on the silicon surface, oxide is etched to expose the underlying poly_x wire. Then metal is deposited on the exposed surface. The diffusion is carried out onto the exposed metal. Step (ii) diffusion takes place impurities will diffuse into the contact when diffusion as well as diffusion in polysilicon.

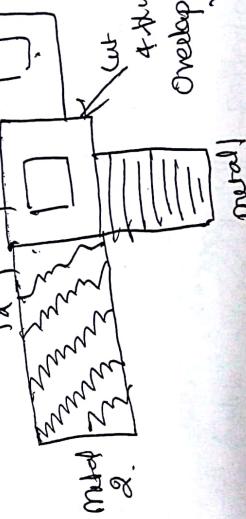
> Buried contacts are smaller than butting

(1) Butting contact:

- > Butting contact is a complex process.
- > A 2×2 contact cut is made down to each layer to be joined. Layers are butted together so that contact cuts become contiguous.
- > The polysilicon & diff. outlines the overlap & thinning under polysilicon acts as an mask in the diffusion process.
- > Polysilicon & diffused layers are butted together.
- > Contact below two layers is made by metal overlay.



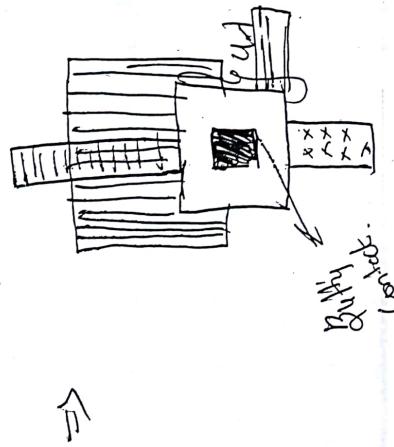
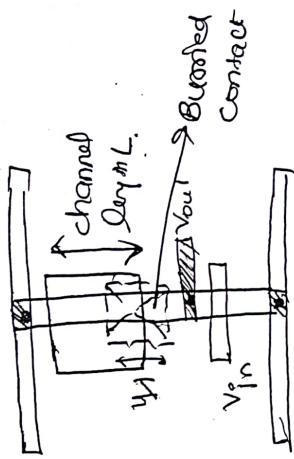
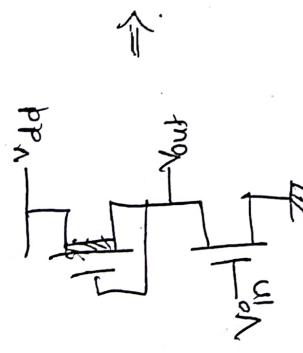
(2) Via Contact from metal to other layers:



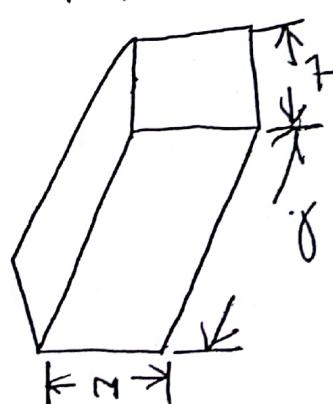
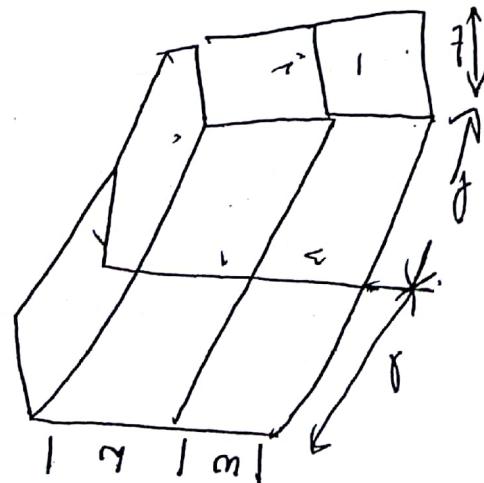
Notes

15B81A04G7

Nmos depletion mode inverter



- Effect of taping tools on capacitance
-
- (i) Parallel plate:- Capacitance of the bottom of the wall
- \rightarrow $C = \epsilon_0 \frac{A}{d}$
- as a conductor over a ground plane.
- On isolated wire over the substrate can be modelled
- as a conductor over a ground plane.
- III plate
- Tapping capacitance
- (ii) Taping:
- Capacitance :-
- * $C = \epsilon_0 \frac{A}{d}$
- * d unique addition of a second wire on the same layer
- * d unique addition of that overlaps (capacitor)
- can exhibit capacitance
-



$$C = \epsilon_0 \frac{A}{d}$$

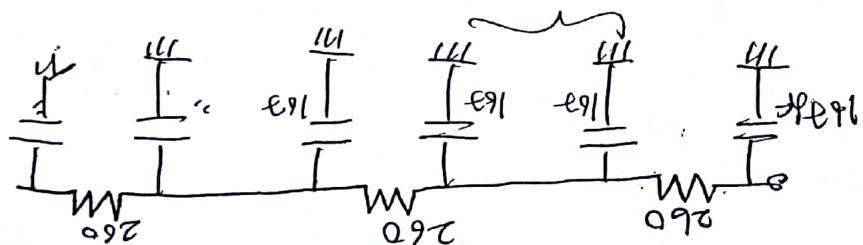
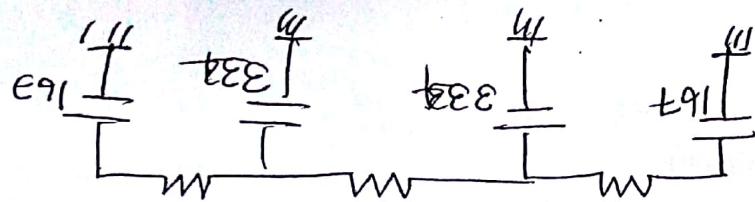
Two conductors with equal separation

$$R = R_0 \times \left(\frac{w}{d}\right)^2$$

+ Rct block

$$R = R_0 \left(\frac{h}{d}\right)^2$$

i) Rectangular box



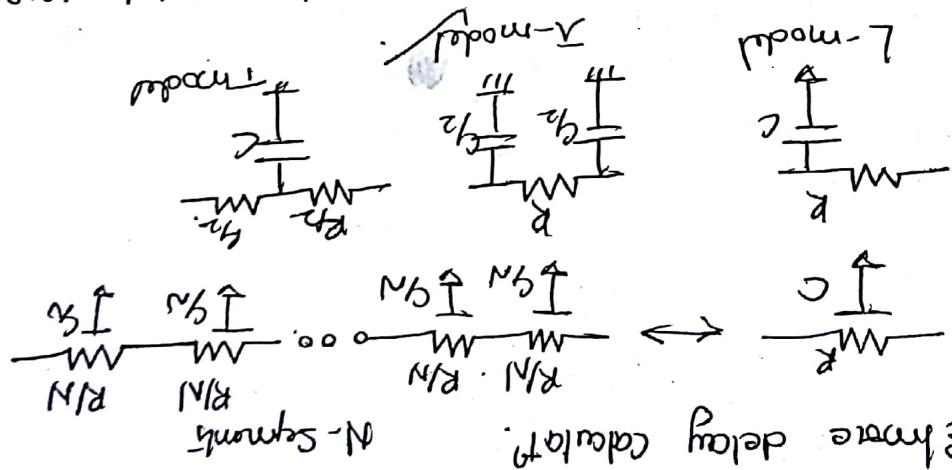
Each T-segment has an equivalent of τ -model

$$\text{Capacity element: } (0.12 \text{ pF/mm}) \cdot (5000 \mu\text{m}) = 1 \text{ pF}$$

$$\text{Total resistance is } 0.05 \Omega/\text{m} \Rightarrow 1525 \times 1 = 78.1 \Omega$$

$$\text{Width is } 5000 \mu\text{m} / 0.32 \mu\text{m} = 15625 \text{ segments in length}$$

Consider a summing, 0.32 mm wide metal wire in 18000 segments. The total resistance of 0.05 Ω/m + capacitance is 0.12 pF/mm . Construct a 8-segment τ -model of the lossless process.



Computable distributed RC delay or thought time.

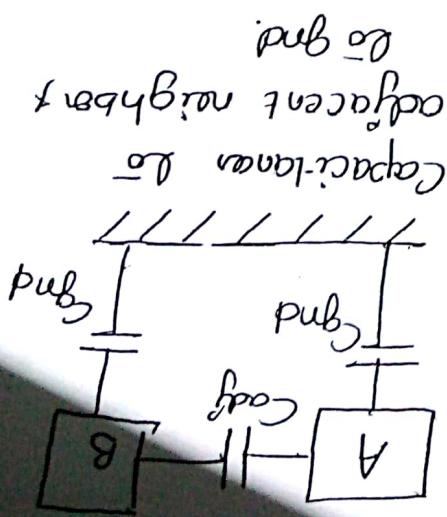
(II) Long wires have significant resistance that

(III) wire capacitance adds loading to each gate

Interconnect increase circuit delay because

* Delay:

When wire A, switch, it finds
to bring its neighbor B along
with it on account of capacitive
coupling, also called cross talk
if B is supposed to switch
of inductively, this may affect the
if B is not supposed to switch,
cross talk comes now in B
as cross talk depends on the ratio of
load/capacitance.

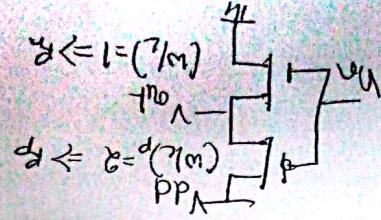


$$\frac{d}{dt} = \frac{d}{dt} - \frac{d}{dt}$$

↓
1 → R_P × 2
↓
2 → R_P × 2

from
PMS is twice
that of C

$$\frac{1}{C} \propto \left(\frac{1}{R_P}\right)$$



$$L \leq 1 = \left(\frac{1}{R_P}\right) \quad \text{and} \quad C = \frac{1}{R_P} \left(\frac{1}{V_L}\right)$$

form of PMS is twice that of C $\therefore I = \frac{1}{R_P} \neq \frac{1}{R_P} = \frac{1}{R_P} \therefore$

$$\frac{1}{C} \propto \frac{1}{R_P} \quad \text{and} \quad \frac{1}{R_P} \propto \frac{1}{C}$$

$$L \propto C \quad \text{and} \quad C \propto L$$

$$L \propto C \quad \text{and} \quad C \propto L$$

$$R_P \propto C \quad \text{and} \quad C \propto R_P$$

∴ $R_P \propto C$ $\therefore R_P \propto \frac{1}{L}$

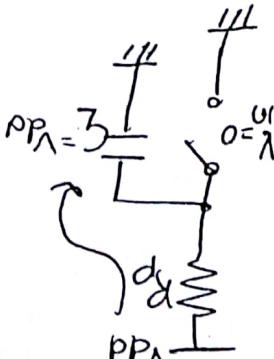
If both current are same, then symmetry

$$\frac{I_p}{R_p} \propto \frac{1}{C} \quad \text{and} \quad \frac{I_n}{R_n} \propto \frac{1}{L}$$

$$I_p = I_n \quad \text{and} \quad R_p = R_n$$

To get equal charging & discharging time $\therefore I_p = I_n$
To get equal charging & discharging time $\therefore I_p = I_n$

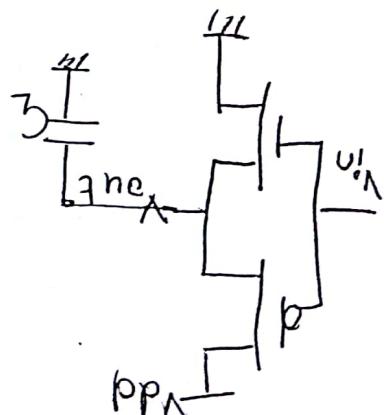
Low - High - Low



$$3 \times C = \frac{1}{R_P} \quad \text{and} \quad 3 \times L = \frac{1}{R_P}$$

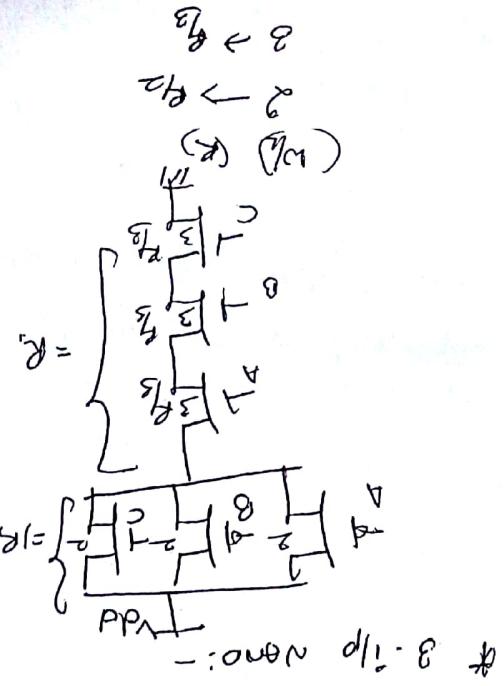
$$\text{charging time: } T_c = R_P \times C$$

$$\text{discharging time: } T_d = R_P \times L$$



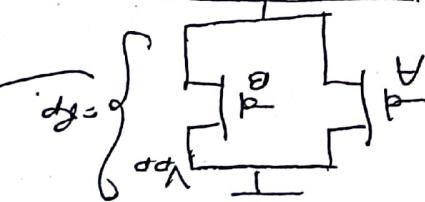
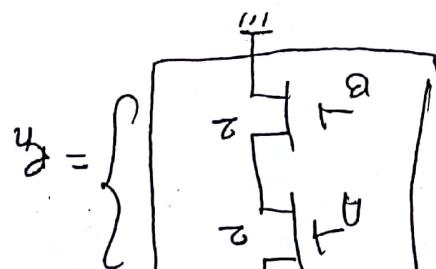
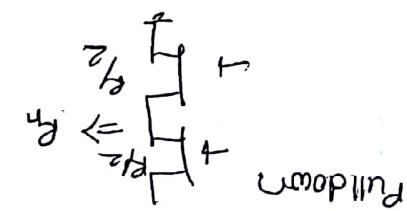
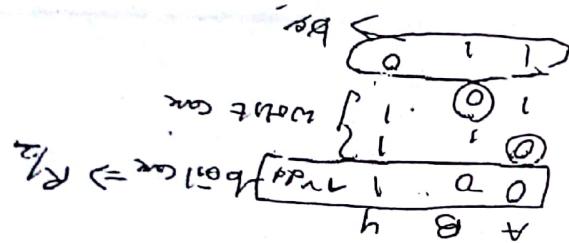
Symmetrical three phases

* Equal charging / discharging time

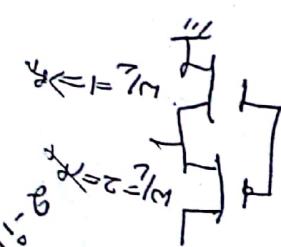


on ap is V_{DD} :: consider only one
out put case if we have only one pmos

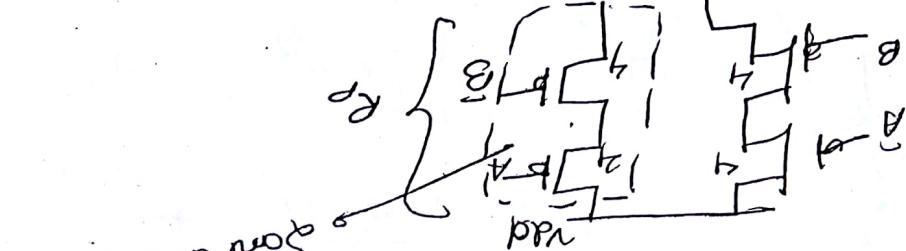
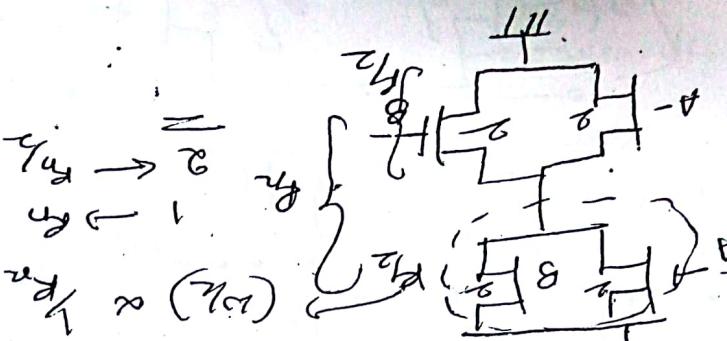
∴ for overall case if we have only one pmos is



Q-11p NAND gate: design *



Summing & Invert



$$Y = \frac{(A+B) \cdot (A+C)}{A \cdot B \cdot C}$$

* 2-11p Ex-OR gate

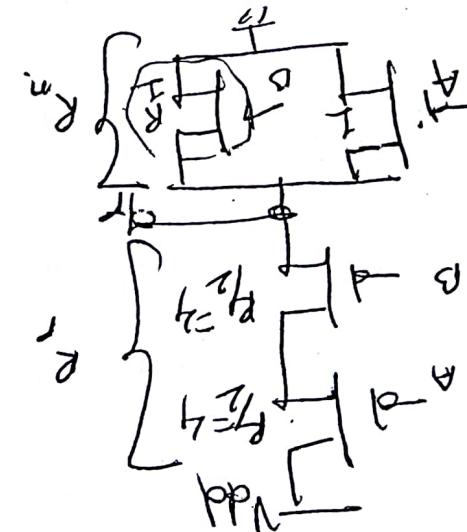
$$F = (A \cdot B) + (A \cdot C)$$

Pull-down: Any one node is pulled down

$\downarrow Q_{12} \rightarrow P_{12}$

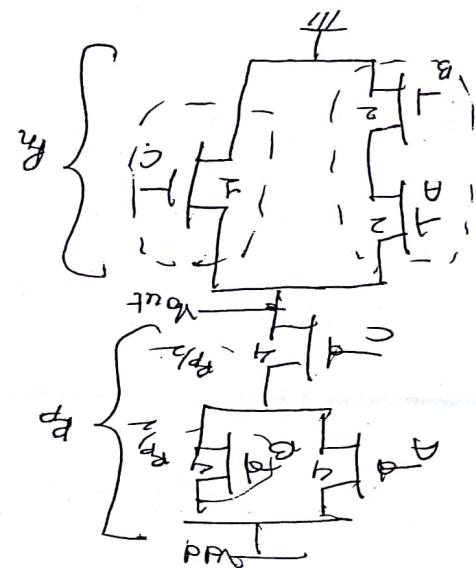
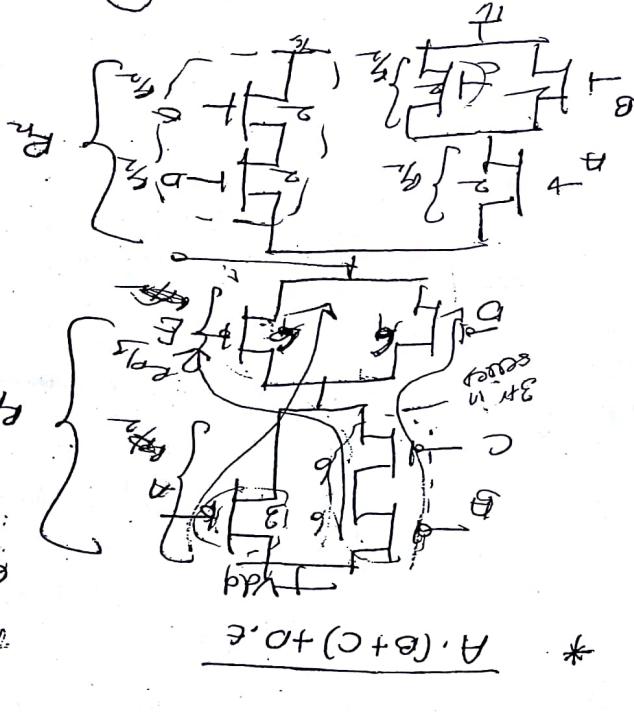
$I \leftarrow 2\mu A$

$Q_{12} \leftarrow R_{PQ} - (W/L) \cdot R_{PQ}$

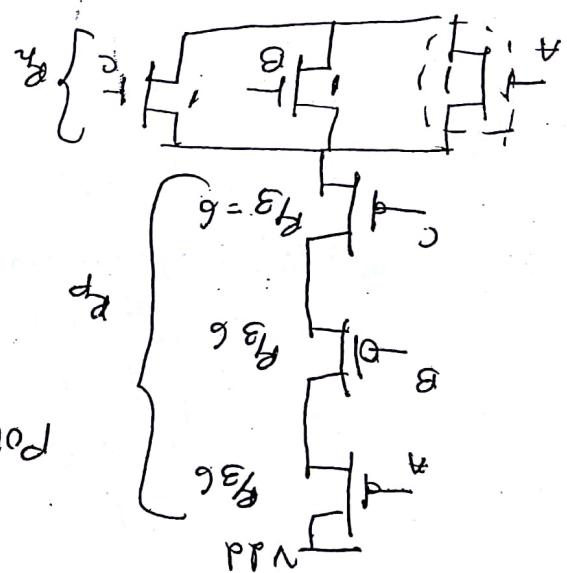


* 2-11p NOR gate

$$\begin{aligned} \frac{d^2y}{dt^2} &= g \\ \frac{d^2y}{dt^2} + g &= 0 \\ \frac{d^2y}{dt^2} &= -g \\ y'' &= -g \end{aligned}$$



$$\begin{aligned} \alpha_1 &\Rightarrow \alpha_2 \\ \alpha_2 &\leq \alpha_3 \\ \alpha_3 &\leq (\%)_{\text{opt}} \end{aligned}$$



* 8-11p NOR Gate

Logic Gate

$$\text{Efficiency} = \frac{\text{Actual Output}}{\text{Theoretical Output}}$$

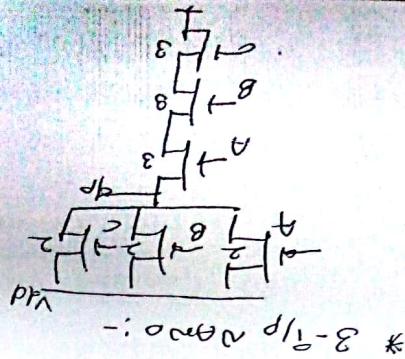
$$b^2(4b_1g) =$$

$$\{u_{M\bar{G}} + u_{M\bar{G}}\} =$$

$$\{d_m + u_m\} b_j = \underline{\underline{u}}$$

6. 100% - E

$$\frac{E_0}{E} = \frac{3\omega_0^2}{4\omega_0^2} = \frac{3}{4} \quad \therefore$$



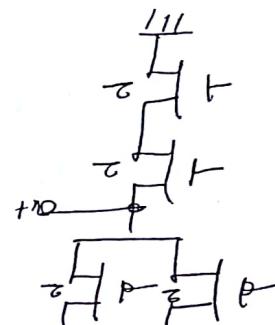
-rob over if-uf *

$$f(\Sigma) = \log \det f_1 + \dots + \log \det f_n$$

$\log_{10} m_F =$

$$= \zeta_2(\alpha_{n+2m})$$

$$y_1 p \cos \phi - C_0 = C_0 + C_1 y_1$$



ppr
: o n e o n d / 1 - 8 p

$F = \frac{m}{r^2} P$ tells why

$$\frac{C_{\alpha} \beta_{\alpha}}{C_{\alpha} \gamma_{\alpha}} = \frac{\beta_{\alpha}}{\gamma_{\alpha}}$$

$$(a_1 + a_2) b =$$

$$(d_{21} + u_{21}) \beta = 0$$

CMOS Inverters \Rightarrow High capacitance of output is ...

निवेदन लोगिका की

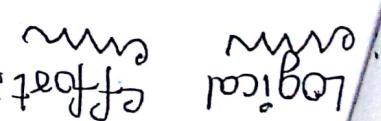
$$\text{goal} / \rightarrow = \beta L$$

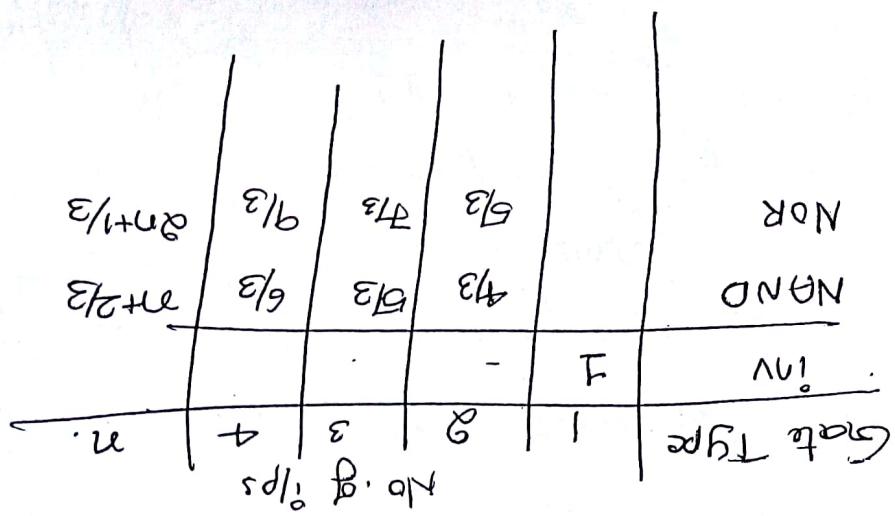
9 " 24 or
24 " 1)

(20)

inventories that can deliver the same off website.

Logical effort of a goal is defined as the ratio of the





$$\frac{3}{1+ab} = 37$$

* $f_{out} = f_{gate}$

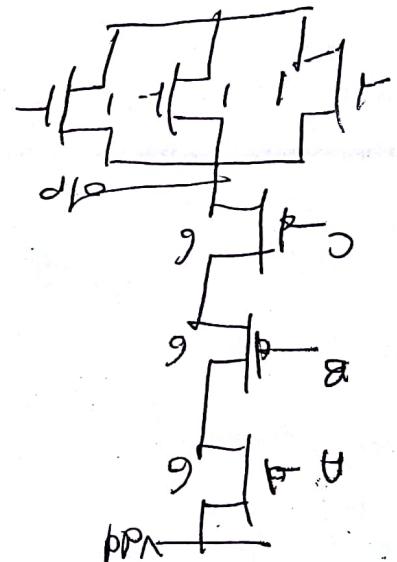
$$LE = \frac{C_{in}}{C_{out}}$$

$$C_{in} = (37) \cdot 4$$

$$C_{out} =$$

$$\{q_{in} + u_n\} N =$$

$$C_{in} = \{q_{in} + u_n\} N$$



3-to-1 NOR Gate

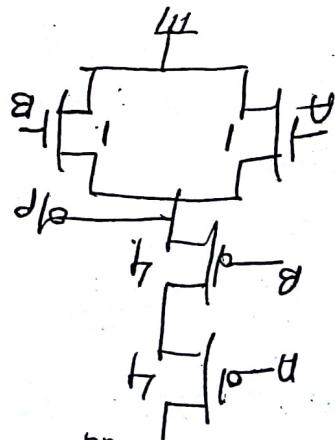
$$\frac{5}{3} = \frac{C_{in}}{C_{out}}$$

$$C_{in} = 3 \cdot C_{out}$$

$$C_{out} =$$

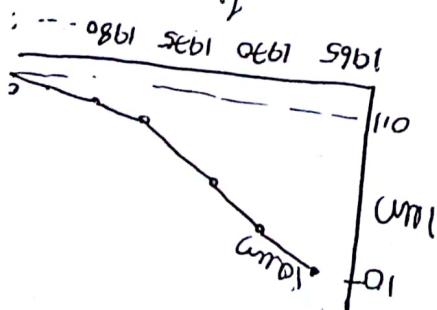
$$\{q_{in} + u_n\} N =$$

$$(C_{in} + u_n) 5 = C_{in} = 5$$



2-to-1 NOR Gate

- Scaling of MOS Circuits
- Justification: Size of the most is shrunk as shown in fig(a) i.e. in 1965
for size is 10 μm, but in 2005 it is reduced to 90 nm. Scaling means adjustment of the dimensions of an electronic device while maintaining the performance of the device.
- VLSI Technology is constantly evolving towards smaller widths + feature size ↓
Reduced feature size generally leads to ↓ better / faster performance ↓
How gallium arsenide ↓
Impact of scaling is characterized in terms of several parameters.
- Minimum feature size
 - Number of gates on each chip (Ld)
 - Power dissipation (P)
 - Maximum operational frequency (f_o) (dissimilarity between Ld and P)
 - Die size
 - Production cost (C_p)
 - Scaling Models
 - Deep model - dimensions + voltage scale together by the same factor
 - Most common model with scalability - only the dimensions
 - Code, voltage scaling second



D is the gate oxide thickness scaled by λ
 where E_s is permittivity of gate oxide (C/m^2) = $\epsilon_0 \cdot \epsilon_r$

$$C_{ox} = E_s / D \quad (\text{C}/\text{m}^2 \text{ or } \text{F}/\text{m}^2)$$

(ii) Gate capacitance per unit area C_{gate}

Thus A_g is scaled by λ^{-2}

W; channel width + both are scaled by λ

where L ; channel length +

$$A_g = \frac{L}{\lambda} \cdot W \quad (i) \text{ Gate Area}$$

∴ Effect of scaling on different parameters of model

vertical dimensions

$1/\lambda$; scaling factor for dimensions, both horizontal

gate oxide thickness D .

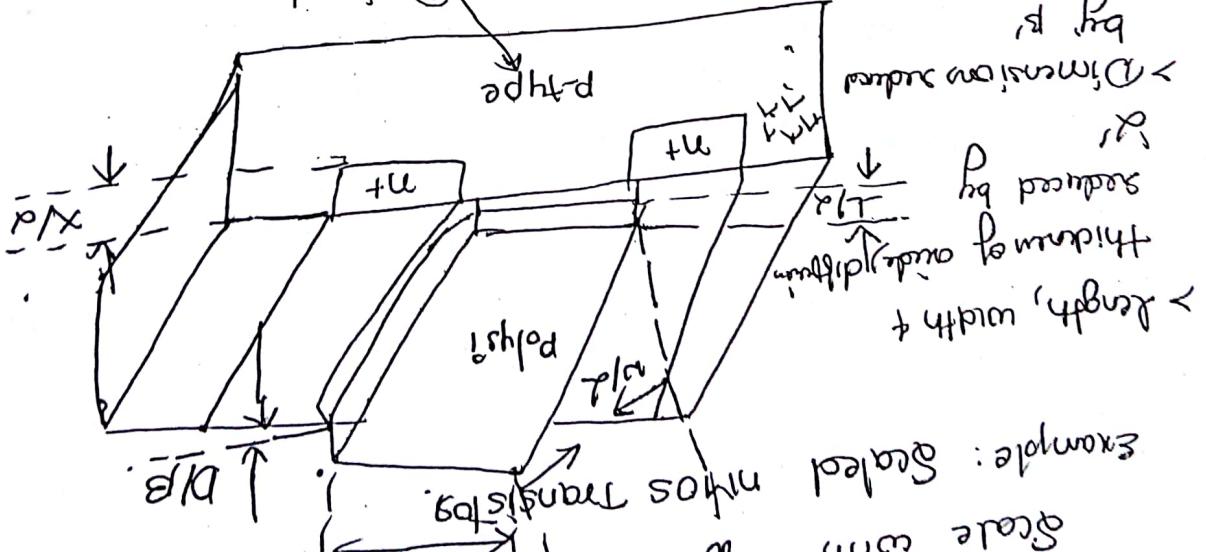
$1/B$; scaling factors for supply voltage V_{DD} +

two scaling factors $1/\lambda$ + $1/B$ are used.

In order to accommodate the scaling model,

scaling factors for device parameters.

Scaling λ



Example: Scaled nmos transistor
 Scale with different factors L/λ , W/λ , D/λ , $n+/p+$
 Model with adiabatic for topology - voltage - width

(i) General scaling

$$\frac{1}{L} \times \frac{1}{\lambda} = R_{on}$$

∴ $L \leftarrow \lambda, w \leftarrow \lambda, \phi_{on} \leftarrow 1$
where $\mu = \text{Channel Capacity per unit length (Gbit/m)}$.

$$R_{on} = \frac{\mu * \mu}{L} * \frac{w}{\lambda}$$

∴ Channel Resistance R_{on} :

Thus ϕ_{on} is scaled by λ

$$\therefore \phi_{on} = B * \lambda^2 = 1$$

C_o is scaled by $B + \lambda^2$ is scaled by λ^2

where ϕ_{on} is the average charge per unit area in the channel in the ion state.

$$\phi_B = C_o * \lambda^2$$

∴ Capacitor density in channel ϕ_B :

$$C_o = \frac{\lambda^2}{\left[\frac{1}{\lambda^2} \right]} \quad \therefore$$

$$-x < \lambda^2 < x$$

Source diagram, scaled by $(L \times W)$

A is the area of the depletion region around

or drawn + scaled by λ^2 .

where R , is the depletion width around source

C_x is the capacitance of A/L .

∴ Parallel plate capacitance C_x :

Thus C_x is scaled by B/λ^2 .

$$-x < \lambda^2 < x$$

$$C_x = C_o * \lambda^2$$

∴ Gate capacitance C_g :

$$B = \frac{C_g}{\lambda^2} \quad \therefore$$

$(\phi_{on}$ is constant)

Thus C_g is scaled by

area

$$E_g = \frac{B}{\alpha^2} \cdot \frac{1}{L} = \frac{B}{\alpha^2}$$

$$E_g = \frac{B}{\alpha^2} \cdot \frac{V_d}{L}$$

$\therefore E_g$ is scaling quantity / Gain \propto

$$\therefore \frac{B}{\alpha^2} = \frac{B}{\alpha^2}$$

$\therefore B$ is scaled by (α^2)

Channel in the "on" state which
where A is cross sectional area of the

$$D_s = I_{DSS}/A$$

\therefore Current density : D_s

$$D_s = \frac{B}{\alpha^2} \cdot \frac{L}{C_s}$$

$$I_{DSS} = \frac{C_s}{2} \times \frac{L}{C_s} \times (V_{GS} - V_t)^2$$

\therefore Saturation Current : I_{DSS}

$$\therefore D_s = \frac{B}{\alpha^2}$$

$$\therefore I_{DSS} \propto B/\alpha^2$$

\therefore If I_{DSS} is inversely proportional to B/α^2

$$\frac{I_{DSS}}{B/\alpha^2} = k$$

\therefore Maximum operating frequency :

$\therefore D_s$ is scaled by B/α^2

$$D_s \propto B/\alpha^2$$

$$I_{DSS} \propto D_s$$

$\therefore T_d$ is proportional to I_{DSS}

\therefore Gate delay T_d

$$\frac{P_d}{P_s} < \left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{1}{\beta_2}} =$$

$$P_d = P_s$$

\Rightarrow Power-Spectral Product

$$P_a = \frac{P_d}{P_s} = \frac{\left[\frac{1}{\alpha_1}\right]}{\left[\frac{1}{\alpha_2}\right]^{\frac{1}{\beta_2}}} = \alpha_2^{\frac{1}{\beta_2}}$$

\Rightarrow Power dissipation per unit area P_a :

$$P_a = \frac{P_d}{P_s}$$

$$P_d = P_s \cdot \alpha_1^{\frac{1}{\beta_2}} \cdot \alpha_2^{\frac{1}{\beta_2}} \quad \therefore P_d = P_s \cdot \alpha_1^{\frac{1}{\beta_2}}$$

$$P_d = P_s \cdot \alpha_1^{\frac{1}{\beta_2}} \cdot \alpha_2^{\frac{1}{\beta_2}} \quad \therefore P_d = P_s \cdot \alpha_1^{\frac{1}{\beta_2}}$$

$$P_d = P_s + P_d^d \quad \text{+ dynamic component } P_d^d$$

Where, the static power component P_d^s ,

$$P_d^s + \text{dynamic component } P_d^d$$

Components of two components: static component

$$P_d^s + P_d^d = P_d$$

power dissipation per gate P_d :