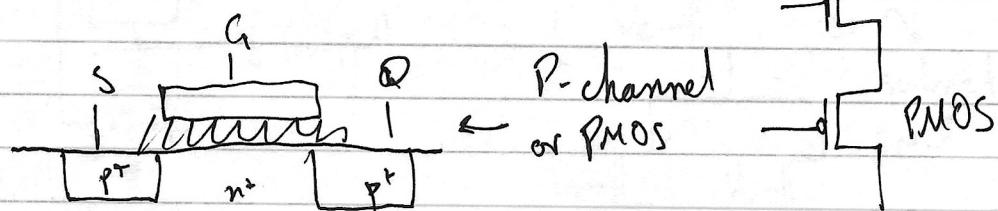
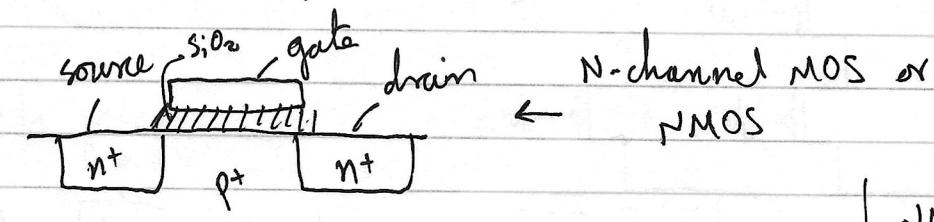


LECTURE 01

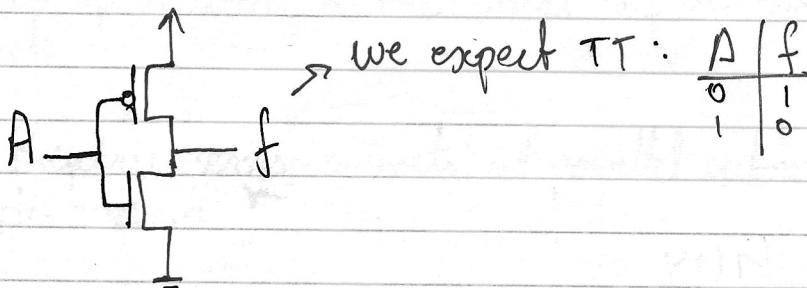
Sept 03

CMOS Logic Circuits

- Built out of MOS transistors



CMOS Inverter



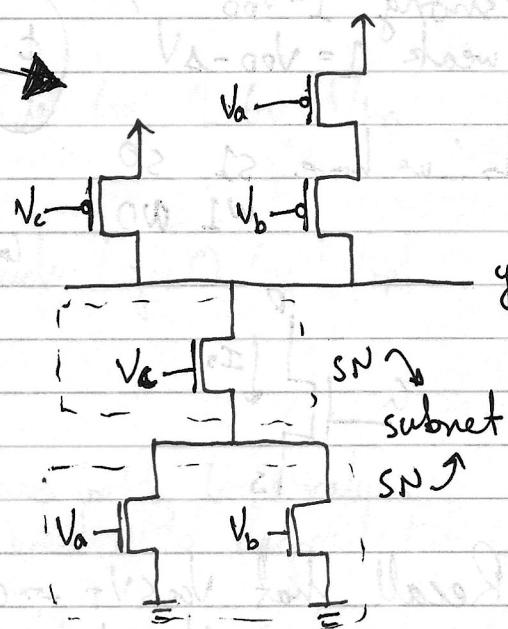
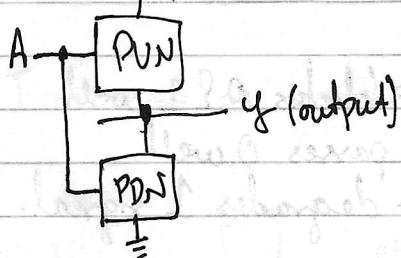
LECTURE 02

Sept 04

Today's topics: CMOS Logic, Pass Transistors, Tristate Gate

E.g. $y = \overline{(a+b)}c$

- In general:



- Pull-up network is the dual of the pull-down network

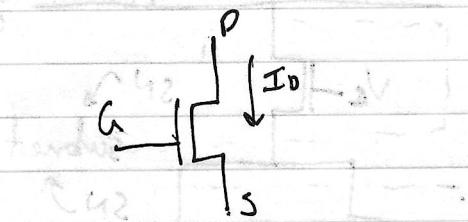
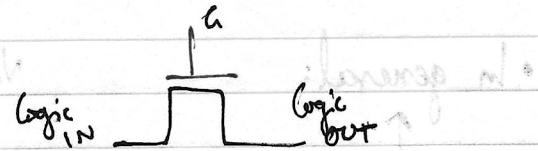
↳ Replace series subnets or parallel subnets & vice-versa.

		PUN	
		OFF	ON
PDN	OFF	Z	1
	ON	0	crowbar X
		high impedance state	short circuit from Vdd to GND AVOID

Pass Transistors would actually be $I \times V_{TN}$

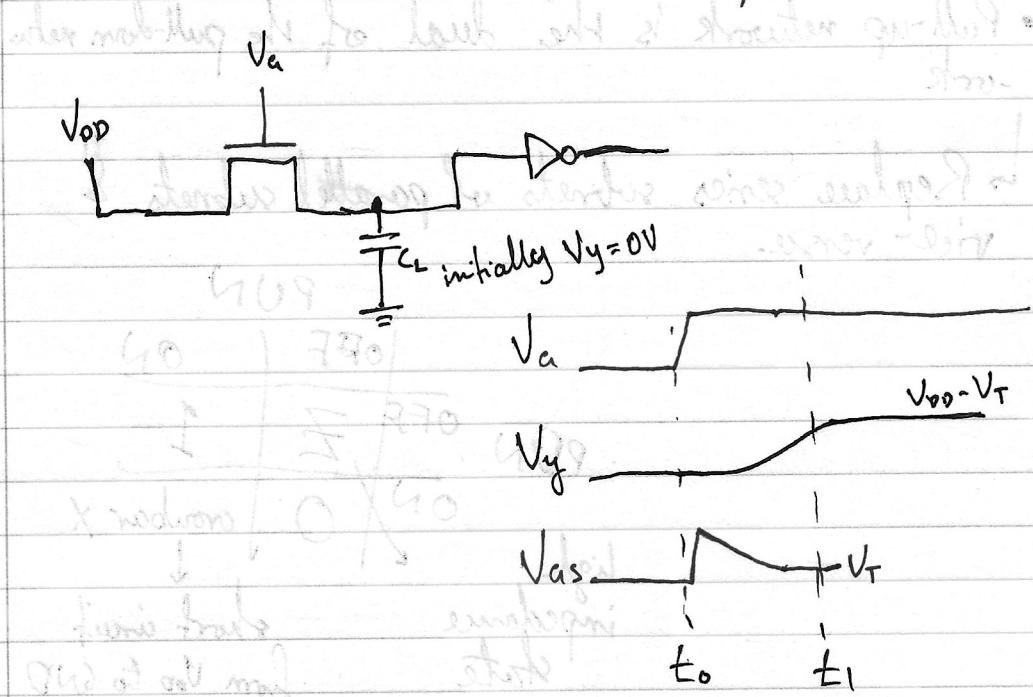
- strong $I = V_{DD}$
- weak $I = V_{DD} - \Delta V$

↳ we have $S_1 S_0$
 $w_1 w_0$

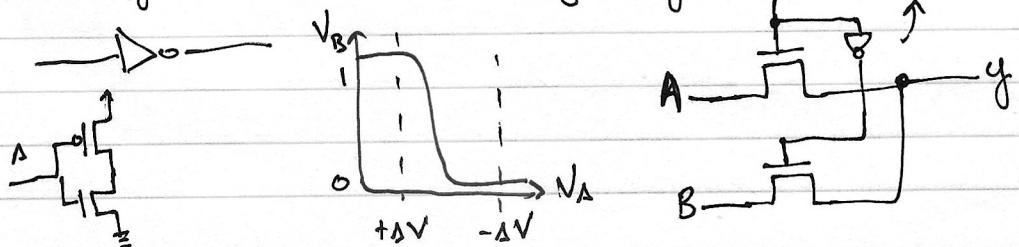


$G=0 \rightarrow$ blocks 0 & 1 well
 $G=1 \rightarrow$ passes 0 well
 $G=1 \rightarrow$ degrades 1 signal.

- Recall that $V_{AS} \leq V_T \rightarrow$ cutoff
 $V_{AS} \geq V_T \rightarrow$ saturation/triode



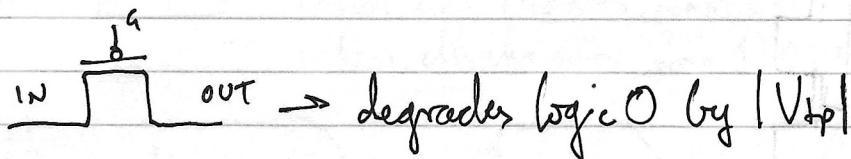
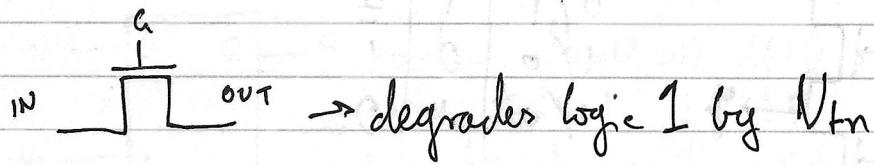
Restoring vs Non-Restoring Logic



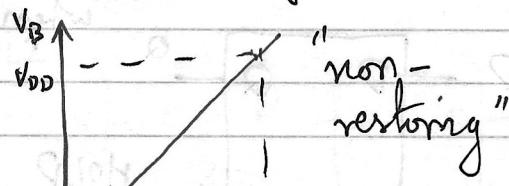
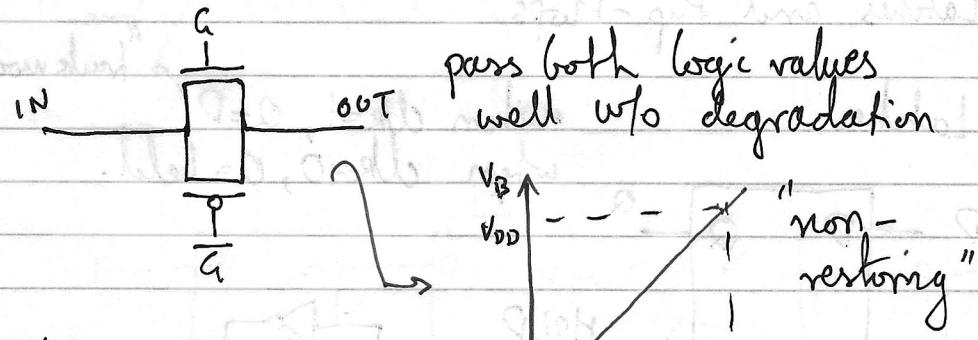
LECTURE 03

Sept 08

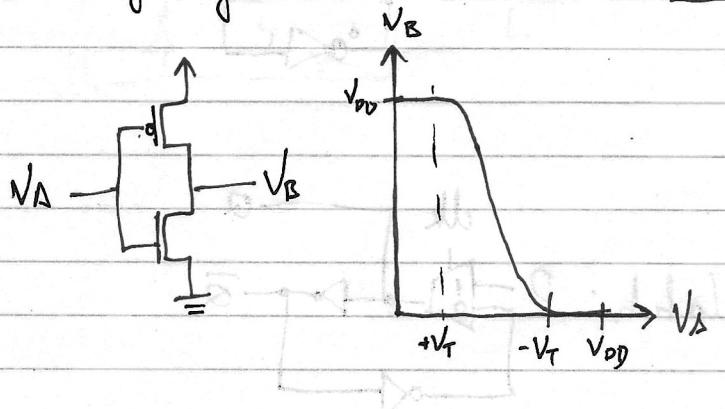
Pass Transistors



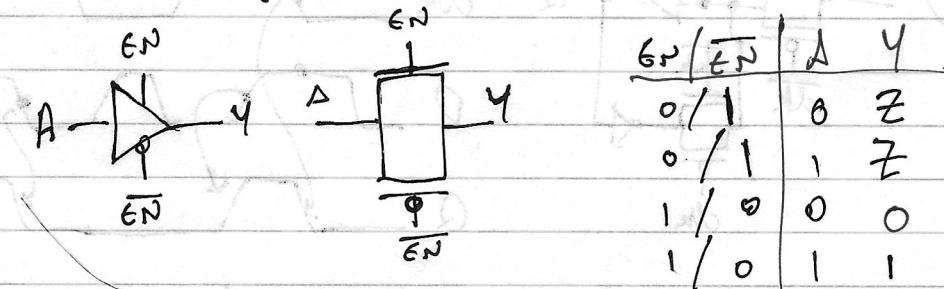
Transmission Gate

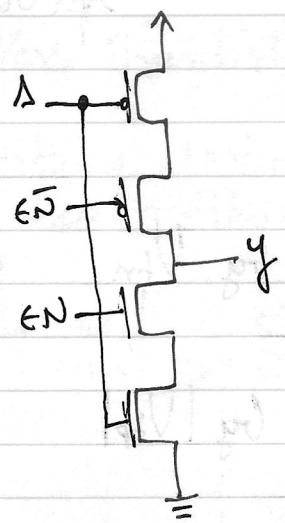


Restoring Logic



Tri-State Logic

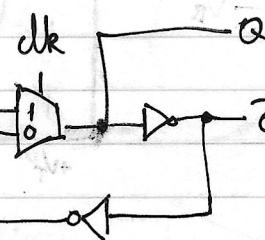
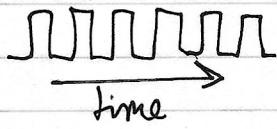
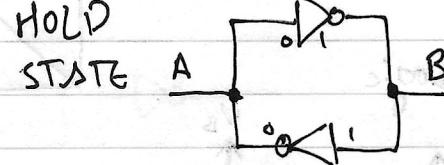
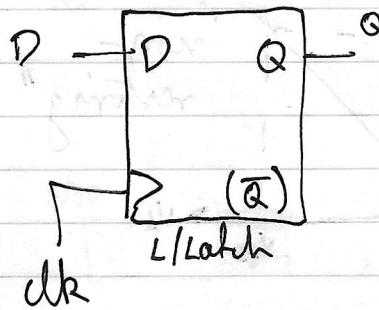




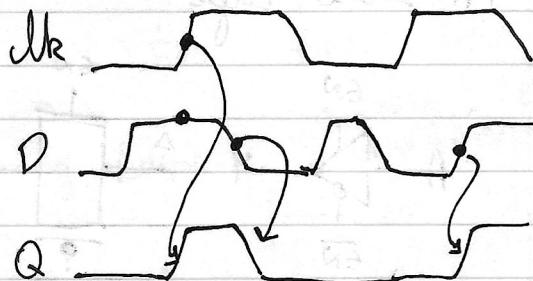
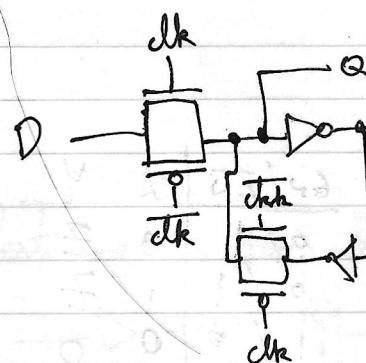
EN/EN	A	Y
0/1	0	Z
0/1	1	Z
1/0	0	1
1/0	1	0

Latches and Slip-Slops

Latches when $clk=1$, $Q=D$ → "latch mode"
when $clk=0$, Q = held.

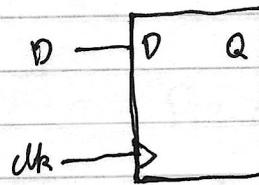


- To create a latch:



— time —

Flip-Flops



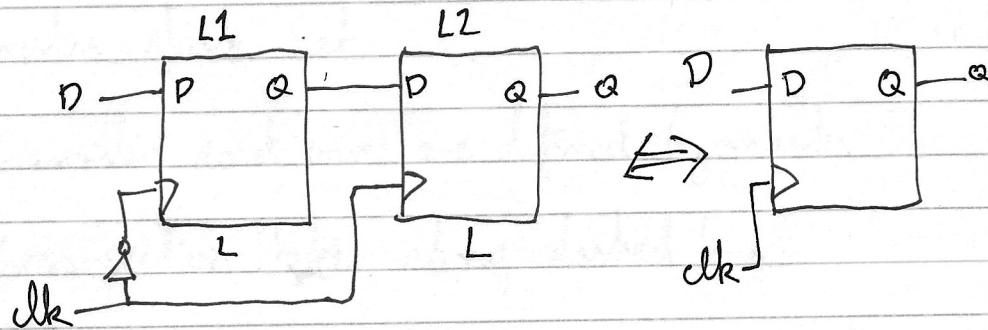
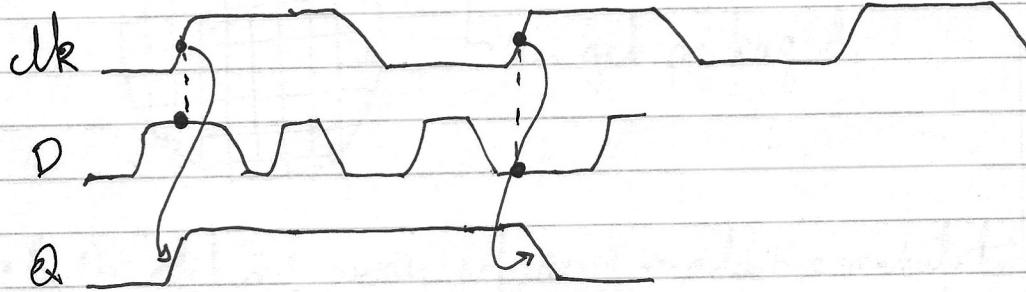
when $\text{clk} = 0$ Q HELD

when $\text{clk} = 1$ Q \neq HELD

when clk changes from 0 \Rightarrow 1 Q = D

when clk changes from 1 \Rightarrow 0, Q HELD

only
"posedge"

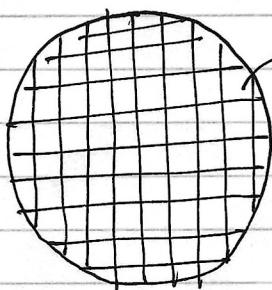


LECTURE 04

Sept 10

- Surface mount \rightarrow silicon die gets packaged into connections w/ pins and/or balls

Die on wafers



Single die

↳ can get packaged.

- SiO₂ etching away w/ acid \rightarrow additive manufacturing

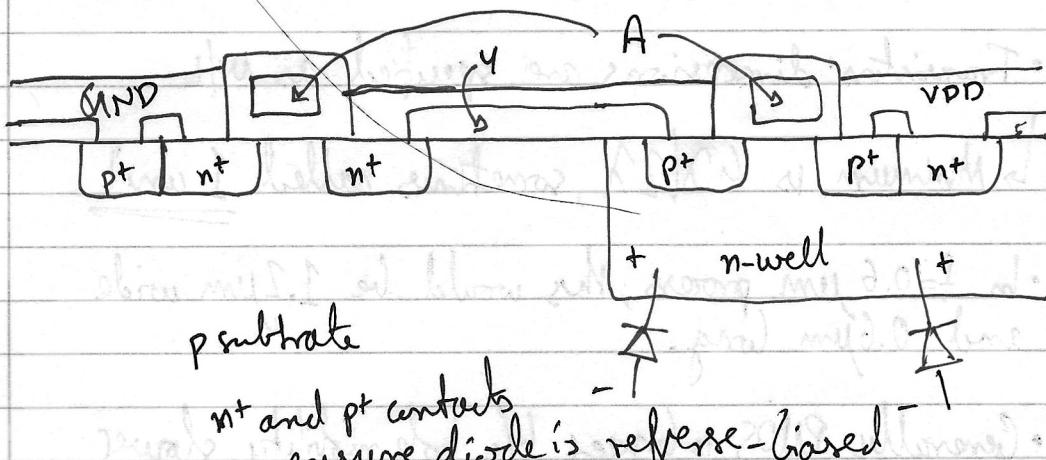
Inverters Mask Set

- Transistors and wires are defined by masks

- Cross-section taken along dashed line

- Large planes and layers of VDD and GND are called utilities

- Metal to light-doped semiconductor forms poor connection called Shottkey Diode



- Six marks in order

↳ ① n-well

↳ ② Poly silicon

↳ ③ n⁺ Diffusion

↳ ④ p⁺ Diffusion

↳ ⑤ Contacts

↳ ⑥ Metal Layer

Layout

- Chips are specified by set of marks

- Feature size f - distance b/w source & drain

↳ Improves ~30% every 3 years or so

- Normalize design rules to feature size

↳ We use $\lambda = f/2$ or $f = 2\lambda$

- Transistors' dimensions are specified as w/L

↳ Minimum is $4\lambda/21$, sometimes called 1 unit

- In $f=0.6 \mu\text{m}$ process, this would be $1.2 \mu\text{m}$ wide and $0.6 \mu\text{m}$ long.

- Generally PMOS is bigger b/c hole mobility slower

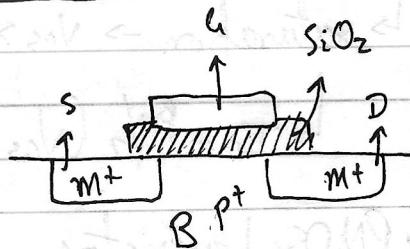
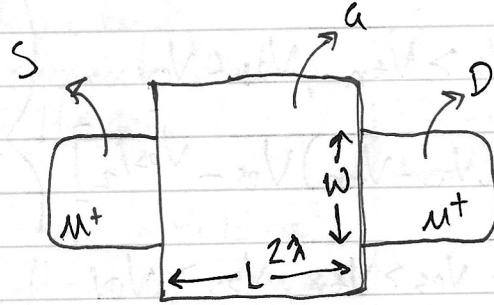
Stick Diagrams

- Help plan layout quickly → no need to be precise or to scale

LECTURE 05

Sept 11

CAPACITANCE $\rightarrow Q = CV$



Threshold Voltage

- We have V_{tp} and $V_{tn} \rightarrow V_t \neq V_T = kT/q$

- $\hookrightarrow V_{tn} \approx 0.2 \rightarrow 0.5V$ depends on fab process.

- Define $\beta_n = \mu_m C_{ox} (W/L)$

- $\hookrightarrow \mu_m$ = mobility of electrons in channels of NMOS

- $\hookrightarrow C_{ox}$ = gate oxide capacitance per unit length

- $\hookrightarrow W/L$ \rightarrow transistor width and length.

- Simplified MOS modelling

- \hookrightarrow NMOS \rightarrow source/drain junction

- \rightarrow source junction for which
is at lower voltage

$V_{tn} > 0$

$I_D > 0$

$V_{os} > 0$

- \rightarrow can change during operation

- 3 regions of operation (by Sedra/Smith)
 - ↳ cutoff $\rightarrow V_{GS} \leq V_{tn}, I_D = 0$
 - ↳ triode $\rightarrow V_{GS} > V_{tn}, V_{DS} \leq V_{OV}$
 - ↳ saturation $\rightarrow V_{GS} > V_{tn}, V_{DS} > V_{OV}$
- For NMOS transistors

- PMOS transistors

- ↳ Source is terminated at the higher potential
- ↳ cutoff \rightarrow

- ↳ triode \rightarrow

$$I_D = \frac{\beta_n}{2} (V_{GS} - V_{tn})^2$$

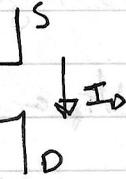
- ↳ saturation \rightarrow

$$I_D =$$

$\epsilon_0 \rightarrow$ permittivity of free space

$\kappa_{ox} \rightarrow$ relative permittivity of SiO_2

$t_{ox} \rightarrow$ thickness of 15 \AA to 100 \AA

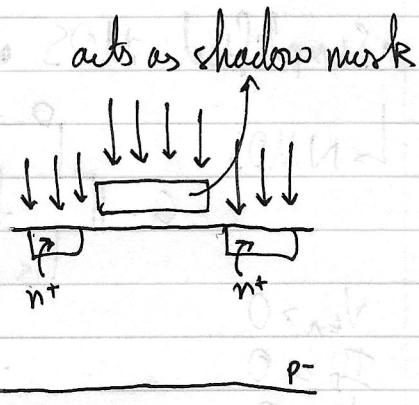


- Oxide Capacitance

- ↳ Should be aware of overlap capacitance

$$C_{ox} = \kappa_{ox} \epsilon_0 \text{ units of } F/m^2$$

$$\approx 3.9 \leftarrow \frac{1}{\kappa_{ox}} \text{ or } F/\mu\text{m}^2$$



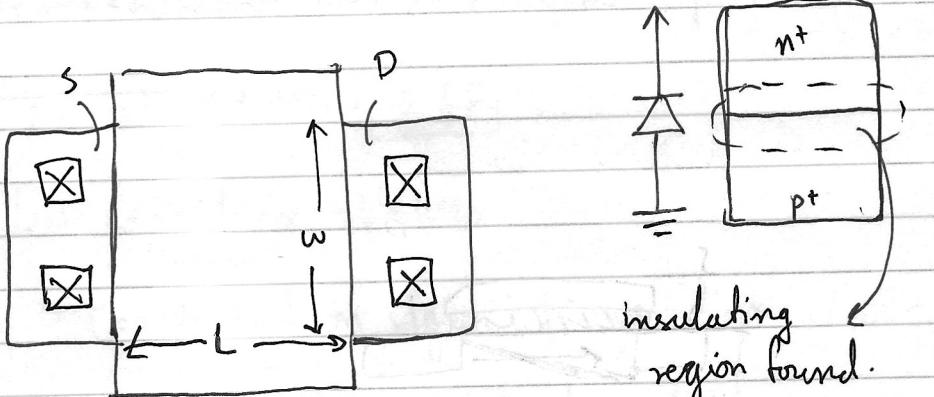
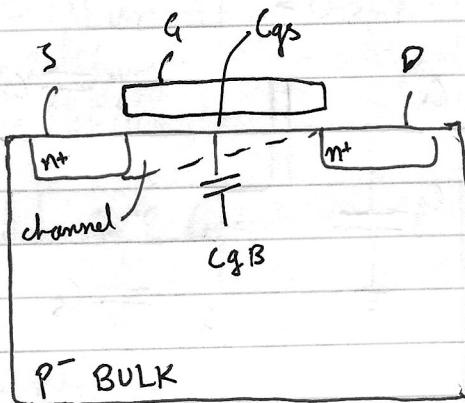
LECTURE 06

Sept 15

MOS Capacitances

- Recall overlap whereby some contaminants ricochet underneath the gate.

- ↳ C_{GS} → gate-source capacitance
- ↳ C_{GD} → gate-drain capacitance
- ↳ C_{AB} & C_{SB}



A cross-sectional diagram of an n-channel MOSFET. The drain region is shaded with diagonal lines. The drain-to-source junction is indicated by a shaded rectangular area. The drain width is labeled 'w' and the drain length is labeled 'L_D'. The source and drain regions are labeled 'n+'.

$$\therefore C_d = C_{jbs} W L_D + C_{jbs w^2 / 2} (W + L_D)$$

- Now consider only drain $\Rightarrow C_{jbs}$ cap per area of pn junction

$$\hookrightarrow C_d = C_{ds} + C_{dp}$$

A cross-sectional diagram of an n-channel MOSFET. The drain region is shaded with diagonal lines. The drain-to-source junction is indicated by a shaded rectangular area. The drain width is labeled 'w' and the drain length is labeled 'L_D'. The source and drain regions are labeled 'n+'.

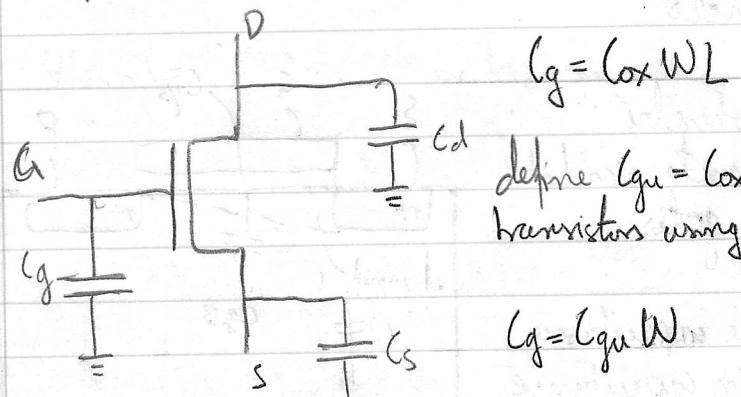
$$C_{ds} = C_{jbs} W L_D$$

$$C_{dp} = 2 C_{jbs} (W + L_D) h_p$$

Define $C_{jbsw} = C_{jbs} h_p [F/m]$

$$C_{dp} = C_{jbsw} \cdot 2 (W + L_D)$$

Simple Model



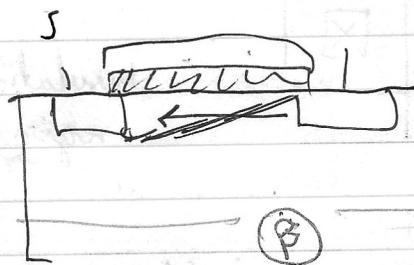
define $C_{gu} = C_{ox} L_{MIN}$ then for transistors using L_{MIN} ,

$$C_g = C_{gu} W$$

typically $C_{gu} \approx 2 \text{ fF}/\mu\text{m}$

- Most technologies t_{ox} scales as L_{MIN} scales

$$C_{ox} = \frac{R_{ox} \epsilon_0}{t_{ox}}$$



ϵ_0

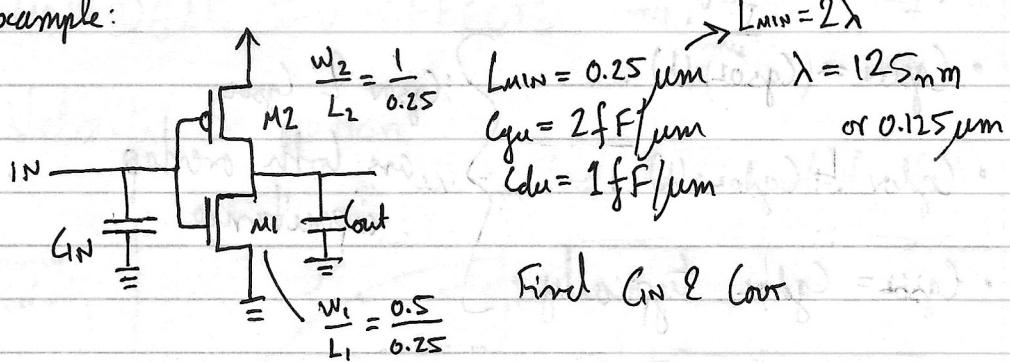
k_{ox}

t_{ox}

LECTURE 07

Sept 17

Example:



$$G_N = C_{gu} W_1 + C_{gu} W_2 = C_{gu} (W_1 + W_2)$$

$$= [2 \text{ fF}/\mu\text{m}] [1.5 \mu\text{m}] = 3 \text{ fF}$$

$$C_{out} = C_{du} W_1 + C_{du} W_2 = 1.5 \text{ fF}$$

Detailed Capacitance Values

★ Table 2.1
Pg 71/70

- G_N & C_{out} depend on region of operation

	cutoff	triode	active
C_{gb}	C_0	0	0
C_{gs}	0	$C_0/2$	$2C_0/3$
C_{gd}	0	$C_0/2$	0

★ $C_0 = C_{ox} WL = C_{gu} W$

- Mainly caused by channel inversion

C_{gso} and C_{gdr}

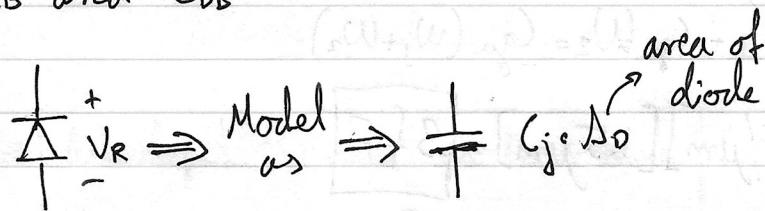
$$\bullet C_{gso} = \frac{(C_{gsoL})}{\text{FF, } \mu\text{m}} \cdot W$$

$$\bullet C_{gdr} = (C_{gdrL} \cdot W)$$

$$\bullet C_{gsoL} = C_{gdrL} \text{ typically}$$

C_{gso} & C_{gdr}
are both overlap
capacitance

C_{SB} and C_{DB}



$$\bullet \text{Could also be given } C_j = \frac{C_{j0}}{(1 + V_R/\phi_0)^{Mj}}$$

↳ C_{j0} is cap per unit area at $V_R=0$

↳ V_R is reverse biased voltage

↳ ϕ_0 is built-in potential $\boxed{\phi_0 = 0.9 \text{ V}}$

↳ Mj varies from 0.3-0.5

Non-Ideal Effects

① Velocity Saturation

⑤ Junction leakage

② Channel length modulation

⑥ Tunneling

③ Body Effect

⑦ Temperature

④ Subthreshold conduction

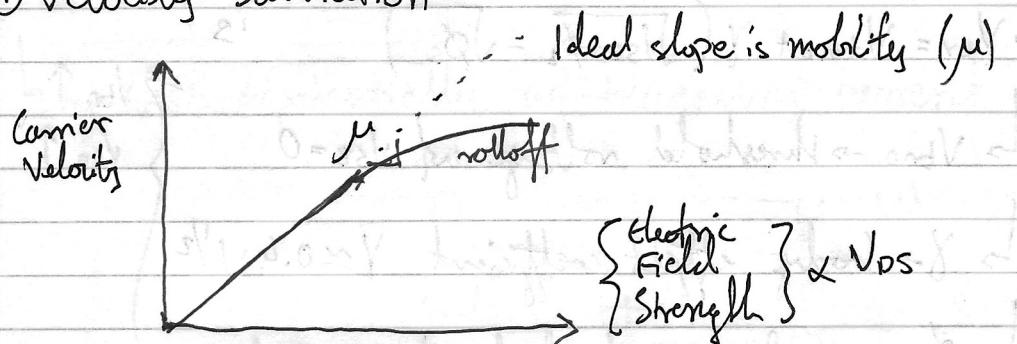
⑧ Geometry variation

LECTURE 08

Sept 18

NON-IDEAL I-V EFFECTS

① Velocity Saturation

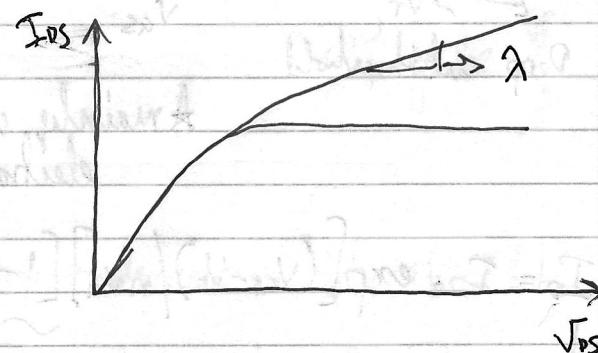


Ideal: $I_D = \frac{\mu n C_o x}{2} \left(\frac{W}{L}\right) (V_{DS})^2$ square law in active region
 $I_D \propto V_{DS}^2$

Velocity saturated: $I_D \propto (V_{DS} - V_t)$ or $I_D \propto V_{DS}$

② Channel Length Modulation

In active region $I_{DS} = \frac{\mu n C_o x}{2} \left(\frac{W}{L}\right) (V_{DS} - V_t)^2 (1 + \lambda V_{DS})$

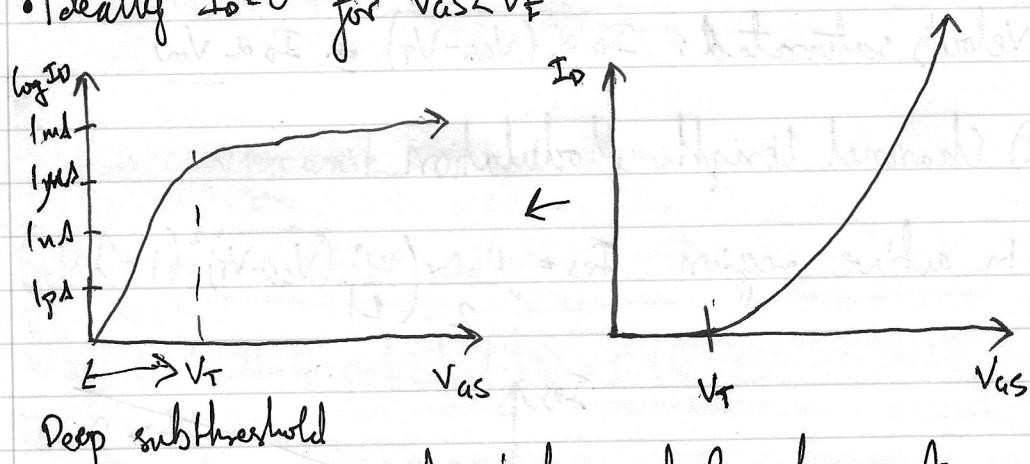


③ The Body effect

- V_{tn} is a function of V_{SB}
- $V_{tn} = V_{tn0} + \gamma (\sqrt{V_{SB} + \phi_s} - \sqrt{\phi_s})$
- ↳ V_{tn0} → threshold voltage w/ $V_{SB}=0$
- ↳ γ → body effect coefficient $\gamma \approx 0.4 \text{ V}^{1/2}$
- ↳ ϕ_s → surface potential $\phi_s \approx 0.9 \text{ V}$

④ Subthreshold Conduction

- Ideally $I_D = 0$ for $V_{AS} < V_T$



Deep subthreshold

* mainly used for slower *
electronics devices

$$I_D = I_{D0} \exp \left[\frac{(V_{AS} - V_T)}{nV_T} \right] \left[1 - e^{-\frac{V_{OS}}{V_T}} \right]$$

↳ $I_D = I_{D0}$ when $V_{OS} = V_T$

↳ $V_T = kT/q \approx 25 \text{ mV}$ at room temp

↳ ensures $I_D \approx 0$ as

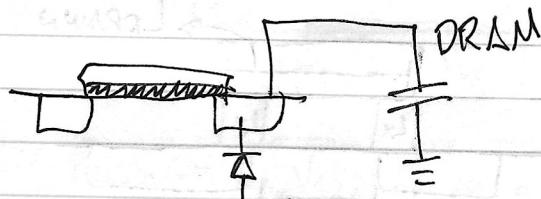
$V_{OS} \rightarrow 0$ but

$\beta \approx 1$ when
 $V_{OS} \gg V_{TN}$

⑤ Junction Leakage

- Reverse-biased diodes leak some current but small $\Rightarrow I_s \approx 0.1 \rightarrow 0.001 \text{ fA}/\mu\text{m}^2$

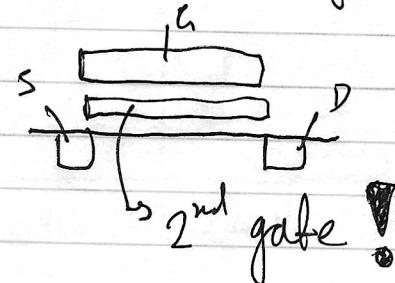
↳ increases exponentially w/ increasing current



⑥ Tunneling

- Gate oxides so thin, electrons "tunnel" through the gate oxide \rightarrow quantum mechanically

↳ used in EEPROM, EEPROM, & flash memory



⑦ Temperature

- When we increase temp \rightarrow decrease mobility of e^- & h^+ * old chips
* airplanes

\rightarrow Digital circuits are slower.

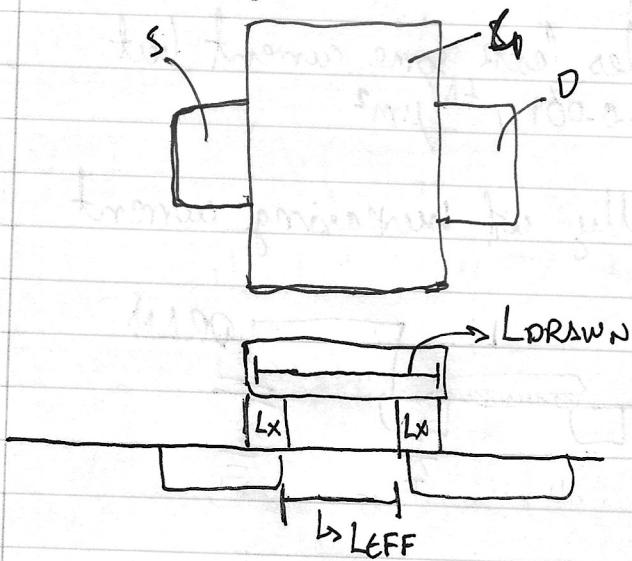
\rightarrow \uparrow tunneling & leakage

\rightarrow ~~\rightarrow~~ $\uparrow V_t$ & subthreshold current

MTBF

\rightarrow reduces device lifetime

⑧ Geometry Dependence

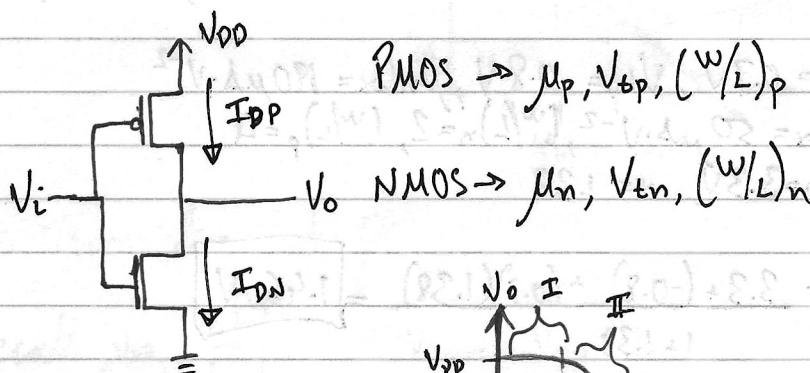
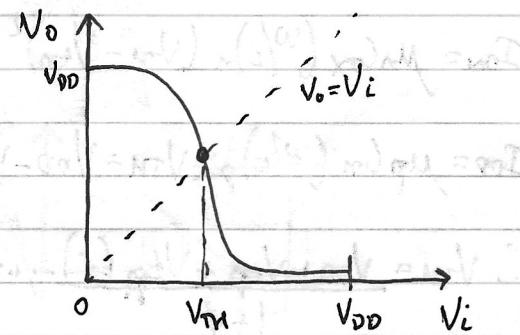


LECTURE 09

Sept 22

DC TRANSFER CHARACTERISTICS

- V_m defined as point where $V_o = V_i$



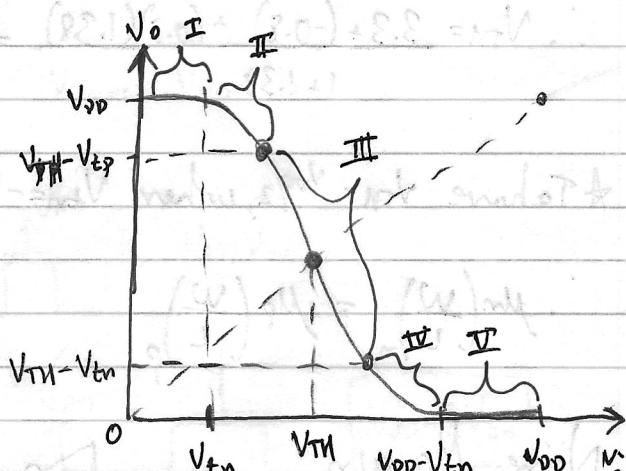
• In region II

↳ NMOS active

$$V_{GS} > V_{TN}, V_{DS} > V_{DA}$$

↳ PMOS triode

$$V_{GS} \leq V_{TP}, V_{DD} > V_{DS} - V_{TP}$$



- In all regions, $I_{DN} = I_{DP}$ as we're only looking at DC characteristics

↳ KCL applies at output node

Region	NMOS	PMOS
I	cutoff	triode
II	active	triode
III	active	active
IV	triode	active
V	triode	cutoff

- V_{TH} occurs when both NMOS & PMOS are in active regions of operation as $V_{DS} = V_{GS}$ for both

$$I_{DN} = \mu_n C_{ox} \left(\frac{W}{L}\right)_n (V_{TH} - V_{TN})^2 \quad \dots \quad \left. \begin{array}{l} \text{Solve for} \\ V_{TH}, \text{ equate} \\ I_{DP} = I_{DN} \end{array} \right\}$$

$$I_{DP} = \mu_p C_{ox} \left(\frac{W}{L}\right)_p (V_{TH} - V_{DD} - V_{TP})^2 \quad \dots \quad \left. \begin{array}{l} \text{Solve for} \\ V_{TH}, \text{ equate} \\ I_{DP} = I_{DN} \end{array} \right\}$$

$$\therefore V_{TH} = \frac{V_{DD} + V_{TP} + V_{TN}(r)}{1+r} \text{, where } r = \sqrt{\frac{\mu_n (W/L)_n}{\mu_p (W/L)_p}}$$

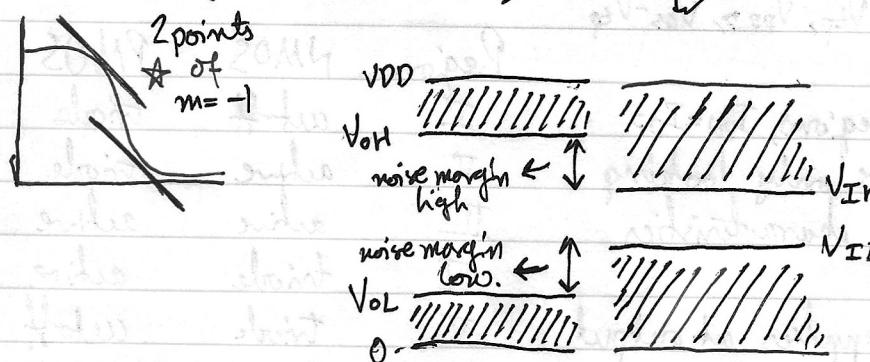
E.g. $V_{TN} = 0.7 \text{ V}$, $V_{TP} = -0.8 \text{ V}$, $\mu_n C_{ox} = 190 \mu\text{A} \cdot \text{V}^{-2}$
 $\mu_p C_{ox} = 50 \mu\text{A} \cdot \text{V}^{-2}$, $(W/L)_n = 2$, $(W/L)_p = 4$
 $V_{DD} = 3.3 \text{ V}$, $r = 1.38$

$$\therefore V_{TH} = \frac{3.3 + (-0.8) + (0.7)(1.38)}{1+1.38} = \boxed{1.46 \text{ V}}$$

* To have $V_{TH} = V_{DD}/2$, when $V_{TN} = -V_{TP}$, let $r = 1$

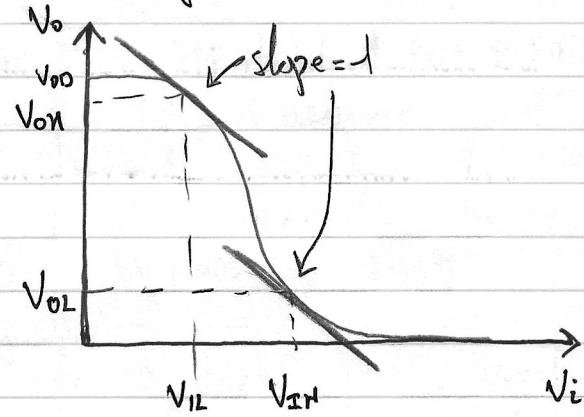
$$\therefore \mu_n \left(\frac{W}{L}\right)_n = \mu_p \left(\frac{W}{L}\right)_p$$

Noise Margin



$$NM_H = V_{OH} - V_{IN}, \quad NM_L = V_{IL} - V_{OL}$$

- V_{IH} → max high input voltage
- V_{IL} → max low \rightarrow $0V$ voltage
- V_{on} → max high output voltage
- V_{oL} → max low \rightarrow $0V$ voltage



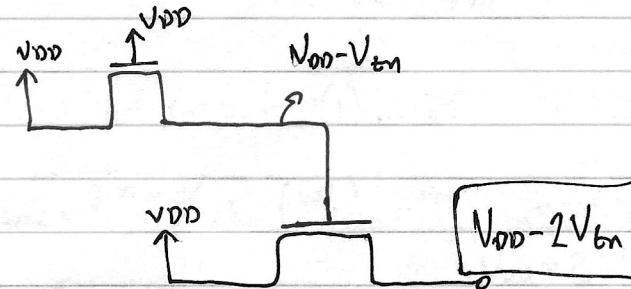
Pass Transistors at DC

- Recall V_{DD}



$$V_{max} = V_{DD} - V_{tn}$$

- Now consider



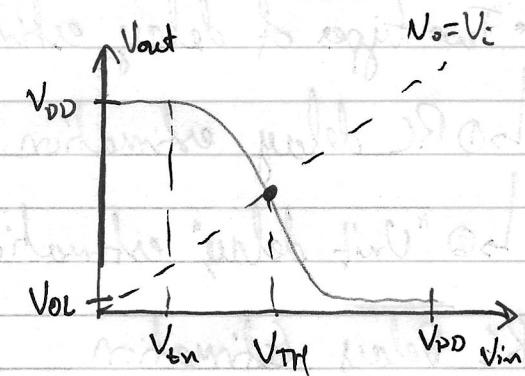
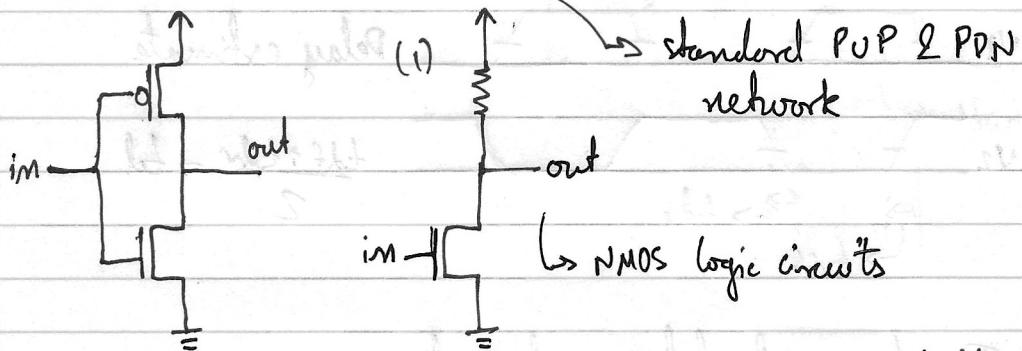
$$V_{DD} - 2V_{tn}$$

LECTURE 10

Sept 24

DC TRANSFER CHARACTERISTICS

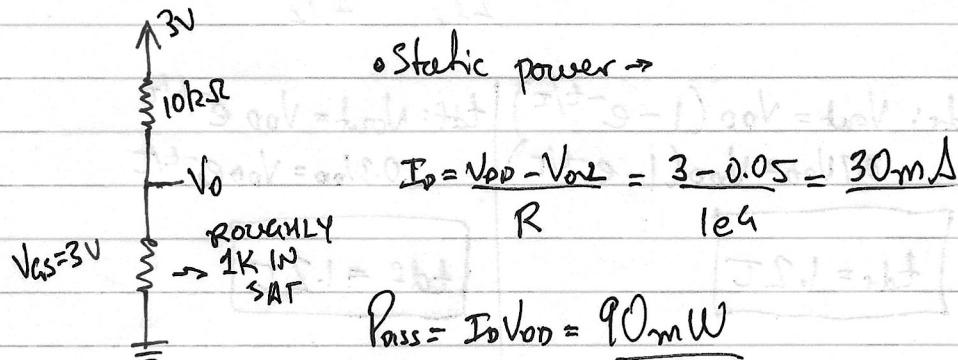
- Applies to conventional and unconventional circuit fam's



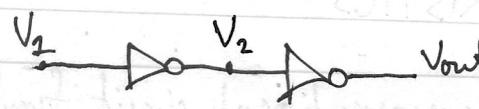
E.g. Find V_{OL} & V_{TM} for circuit (i)

$$\begin{aligned} \rightarrow R &= 10k\Omega, \mu_n C_{ox} = 200 \text{ mV/V}^2, (w/l)_n = 10, V_{tn} = 0.5 \text{ V} \\ V_{DD} &= 3 \text{ V} \end{aligned}$$

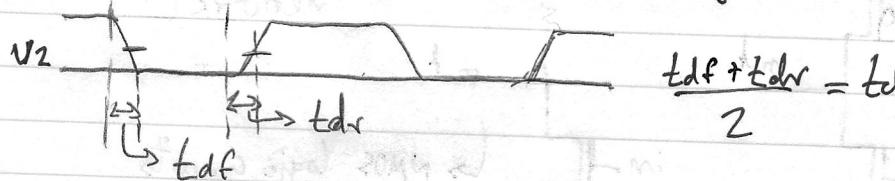
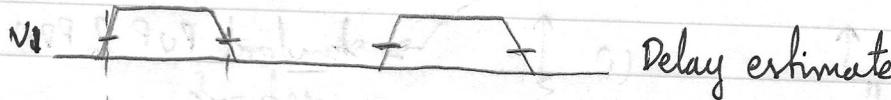
- Model NMOS as resistor



Logic Delay Estimation



Realistic load conditions including gate & wire loading.



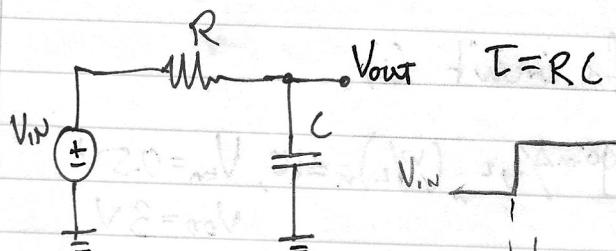
$$\frac{tdr + tdf}{2} = tdl$$

- Two types of delay estimates

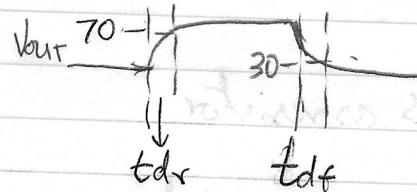
① RC delay estimation

② "Unit delay" estimation

RC Delay Estimation



RISE
2 FALL



$$tdr: V_{out} = V_{dd} (1 - e^{-t/\tau})$$

$$0.7V_{dd} = V_{dd} (1 - e^{-t/\tau})$$

$$tdr = 1.2\tau$$

$$tdf: V_{out} = V_{dd} e^{-t/\tau}$$

$$0.3V_{dd} = V_{dd} e^{-t/\tau}$$

$$tdf = 1.2\tau$$

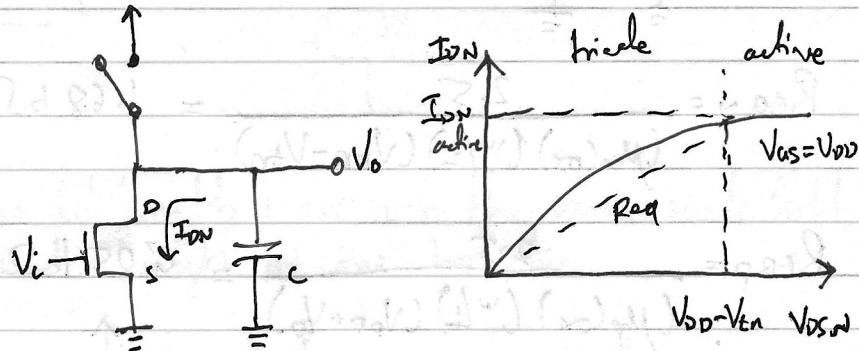
LECTURE 11

Sept 25

Last Lecture

$$\begin{aligned} \cdot t_{dhr} &= 1.2 T \\ \left. \begin{aligned} t_d &= t_{dhr} + t_{df} \\ t_{df} &= 1.2 T \end{aligned} \right\} \quad \tau = RC \quad t_d = \frac{t_{dhr} + t_{df}}{2} \end{aligned}$$

• Consider



- \$I_{DN}\$ starts off in active and when \$V_o\$ falls below \$V_{DD} - V_{TN}\$, the transistor enters triode region

$$R_{eqN} = \frac{V_{DD} - V_{TN}}{I_{DN}} = \frac{(V_{DD} - V_{TN})}{(\mu_n C_{ox}/2)(w/L)(V_{DD} - V_{TN})^2}$$

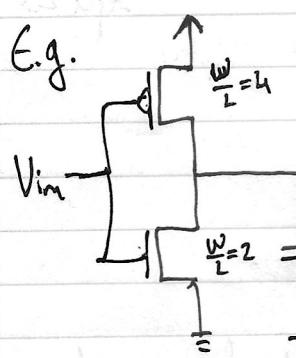
$$= \frac{2}{(\mu_n C_{ox})(w/L)(V_{DD} - V_{TN})}$$

- However, the formula above does not take into account velocity saturation which slows down circuits by 25%.

$$\bullet R_{edn} = \frac{2.5}{\mu_n C_{ox}(w/L)(V_{DD} - V_{TN})}$$

$$\bullet PMOS \rightarrow R_{eqP} = \frac{2.5}{\mu_p C_{ox}(w/L)(V_{DD} + V_{bp})}$$

↳ (-)



$$\mu_n C_{ox} = 360 \mu\text{A}/\text{V}^2 \quad V_{tn} = 0.43\text{V}$$

$$\mu_p C_{ox} = 82 \mu\text{A}/\text{V}^2 \quad V_{tp} = -0.62\text{V}$$

$$R_{Q_N} = \frac{2.5}{(\mu_n C_{ox})(w/L)(V_{DD} - V_{tn})} = 1.68 \text{ k}\Omega$$

$$R_{Q_P} = \frac{2.5}{(\mu_p C_{ox})(w/L)(V_{DD} + V_p)} = 4.05 \text{ k}\Omega$$

- Implication is t_{dr} (PMOS) is longer

$$t_d = \frac{t_{dr} + t_{df}}{2} = 688 \text{ ps}$$

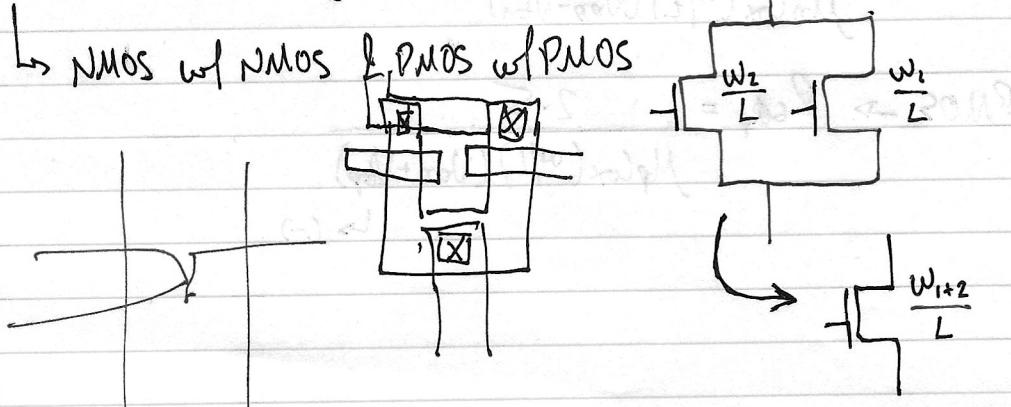
$$\hookrightarrow t_{dr} = 1.2 R_{Q_P} (L = 973 \text{ ps})$$

$$\hookrightarrow t_{df} = 1.2 R_{Q_N} (L = 403 \text{ ps})$$

Transistor Equivalency

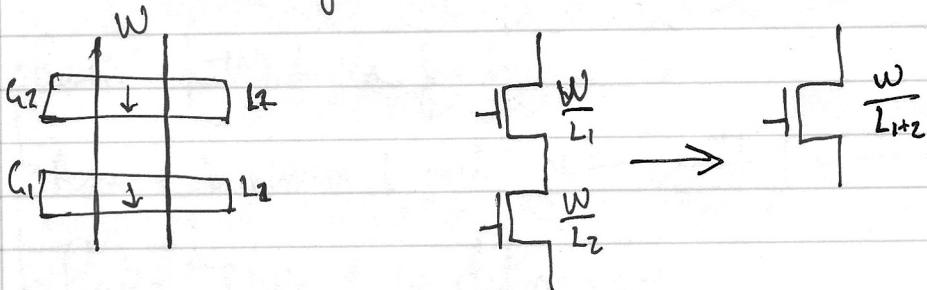
① Combine parallel transistors by adding their widths and assuming same lengths.

\hookrightarrow NMOS w/ NMOS & PMOS w/ PMOS

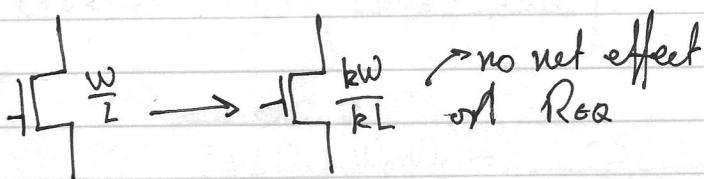


A layout diagram
for this?

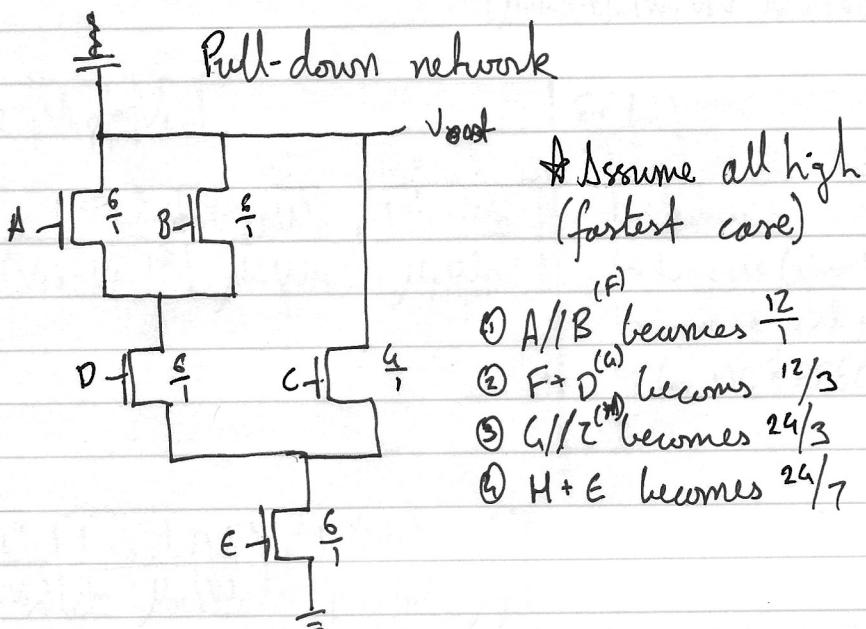
② Combine series transistors by adding lengths and assuming widths are the same



③ If widths (or lengths) are not the same, scale both width and length of one transistor



E.g. Pull-down network



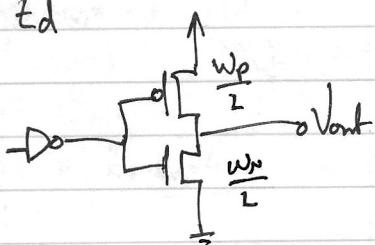
* Assume all high
(fastest case)

* Why fastest
why not average?

- ① A/B becomes $\frac{12}{7}$
- ② F = D becomes $\frac{12}{3}$
- ③ G/Z becomes $\frac{24}{3}$
- ④ H+E becomes $\frac{24}{7}$

• Ratio of W_p to W_n for lowest I_d

↳ Assume each L is minimum
length Ratio of $\frac{W_p}{W_n}$ to be determined.



LECTURE 12

Sept 29

Last Lecture

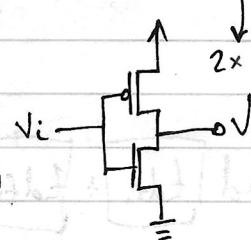
- Ratio of W_p to W_n for lowest t_d

↳ Assume minimum L each $V_i \rightarrow V_o$

↳ Ratio W_p/W_n to be determined.

- Input capacitance ($C_i \approx C_{ox}L(W_p + W_n)$)

$$t_d = \frac{t_{dN} + t_{dP}}{2} = 1.2 C_L \left(\frac{R_{eqN} + R_{eqP}}{2} \right)$$



$$= 1.2 C_{ox} L (W_p + W_n) \left(\frac{1}{2} \right) \left[\frac{2.5}{\mu_n C_{ox} \left(\frac{W}{2} \right)_n (V_{DD} - V_{TN})} + \frac{2.5}{\mu_p C_{ox} \left(\frac{W}{2} \right)_p (V_{DD} + V_{TP})} \right]$$

assume $|V_{TP}| = |V_{TN}|$

$$t_d = \frac{1.5 L^2}{(V_{DD} - V_{TN})^2} \left[\frac{W_p + W_n}{\mu_n W_n} + \frac{W_p + W_n}{\mu_p W_p} \right]$$

For $t_d \downarrow$

↳ Reduce L

↳ increase $(V_{DD} - V_{TN})$

↳ increase μ_n & μ_p

↳ ~~Cox NO EFFECT~~

$$t_d = \frac{1.5 L^2}{(V_{DD} - V_{TN})} \left[\frac{1}{\mu_n} + \frac{1}{\mu_n} \left(\frac{W_p}{W_n} \right) + \frac{1}{\mu_p} \left(\frac{W_p}{W_n} \right)^2 + \frac{1}{\mu_p} \right]$$

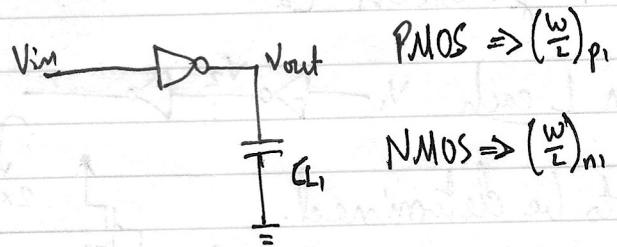
$$\frac{\partial t_d}{\partial \left(\frac{W_p}{W_n} \right)} = \frac{1.5 L^2}{(V_{DD} - V_{TN})} \left[\frac{1}{\mu_n} - \frac{1}{\mu_p} \left(\frac{W_p}{W_n} \right)^2 \right] \rightarrow \text{optimum at } \frac{\partial t_d}{\partial \left(\frac{W_p}{W_n} \right)} = 0$$

$$\frac{1}{\mu_n} - \frac{1}{\mu_p} \left(\frac{W_p}{W_n} \right)^2 = 0 \Rightarrow \boxed{\frac{W_p}{W_n}_{opt} = \sqrt{\frac{\mu_n}{\mu_p}}} \quad \text{e.g. for } \underline{\mu_n = 4 \mu_p}$$

$W_p = 2 W_n$ for opt t_d

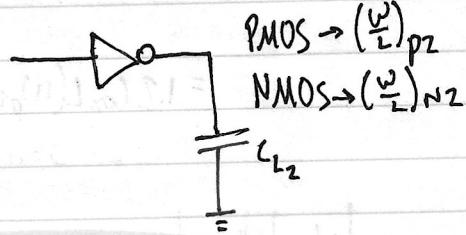
UNIT DELAY ESTIMATION

- Basic Idea \rightarrow Simulate w/ SPICE inverter delay



• Find t_{dr1} + t_{df1}

- To estimate delay for another inverter driving a different capacitive load.



$$\frac{t_{dr2}}{t_{df1}} = \frac{C_{L2}}{C_{L1}} \times \frac{(\frac{w}{l})_{N1}}{(\frac{w}{l})_{N2}} \quad \& \quad t_{dr2} = \frac{C_{L2}}{C_{L1}} \times \frac{(\frac{w}{l})_{p1}}{(\frac{w}{l})_{p2}}$$

\hookrightarrow implicit assumption of same voltage & same technology

E.g. Given $\frac{1}{I} = 100\text{fF}$, $(\frac{w}{l})_p = (\frac{3}{1})$, $(\frac{w}{l})_n = (\frac{3}{1})$

$$t_{df} = 0.48\text{ns} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{measured.}$$

$$t_{dr} = 1.2\text{ns} \quad \left. \begin{array}{l} \\ \end{array} \right\}$$

Find t_{dr} & t_{df} for $\rightarrow t_{df} = (0.48) \left(\frac{200}{100} \right) \left(\frac{3/1}{2/1} \right) = 1.44\text{ns}$

Given $\frac{1}{I} = 200\text{fF}$, $(\frac{w}{l})_p = (\frac{1}{2})$, $t_{df} = (1.2) \left(\frac{200}{100} \right) \left(\frac{3/1}{2/1} \right) = 14.4\text{ns}$

LECTURE 13

Oct 01

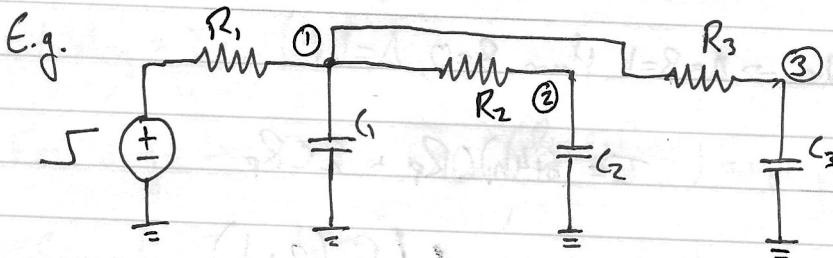
Today \rightarrow Elmore Delay for RC Tree Network

Something wrong w/ Vor. Vor discussed in tutorial

- Elmore was designer at Bell Labs, working on RC circuit theory

↳ Could we come up with an approx for T ?

- $T_i = \sum_k C_k R_{ik}$ \rightarrow i is node of interest
 - C_k is capacitance at node k
 - R_{ik} is sum of all resistances in common from source to node i & source to node k.



For node ① (delay from node input to node ①)

$$\hookrightarrow T_1 = R_1 C_1 + R_1 C_2 + R_3 C_3$$

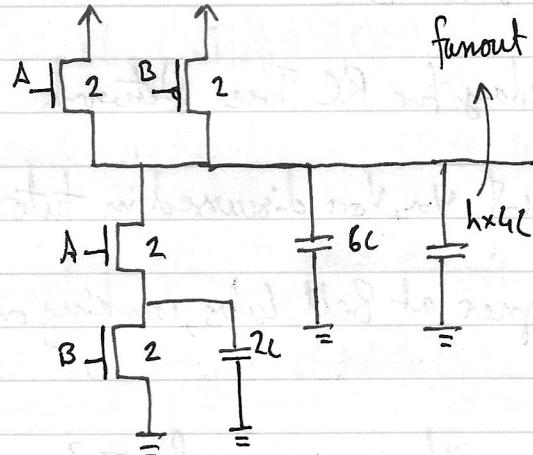
$$\text{For node } ② \rightarrow T_2 = R_1 C_1 + (R_1 + R_2) C_2 + R_3 C_3$$

$$\text{For node } ③ \rightarrow T_3 = R_1 C_1 + R_1 C_2 + (R_1 + R_3) C_3$$

- Elmore delay is realistically accurate

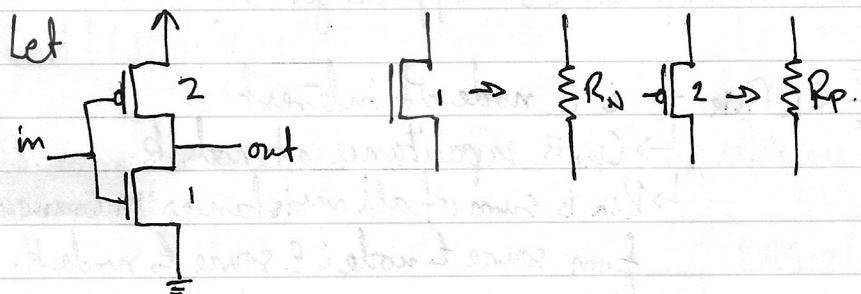
- Even when off, it is almost always useful for optimization in that reducing Elmore delay will almost always reduce true delay.

E.g.

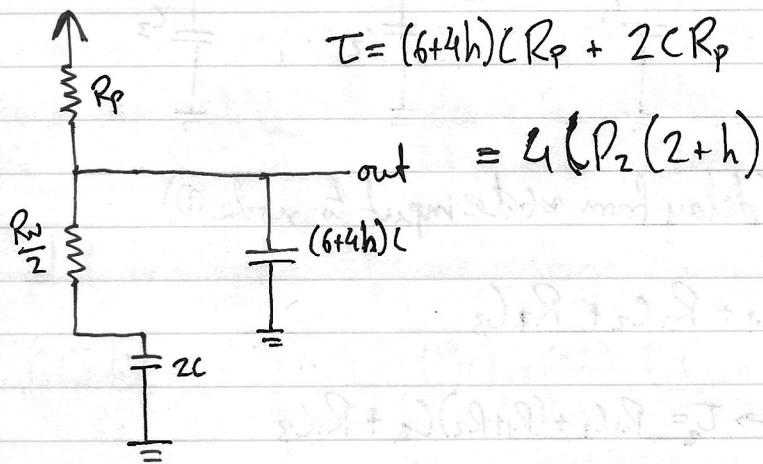


Simple
2-input
NAND
Gate

Let

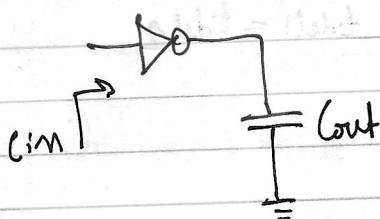


Rising tdr $\rightarrow A=B=1$ then $B=0$, $\Delta=1$



LECTURE 14

- Optimizing gate swing by sizing to minimize delay when driving large capacitive loads
- First define delay through a gate \rightarrow "intrinsic delay"



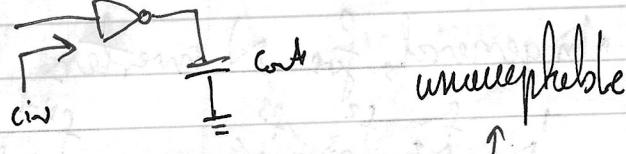
$$t_{\text{delay}} = t_{\text{d,f}} + t_{\text{d,hc}} = 1.2 \text{inv} \left(\frac{R_{\text{cap,in}}}{2} \right)$$

- t_{inv} is the delay of a gate driving same size gate.
- ↳ independent of gate size, actually

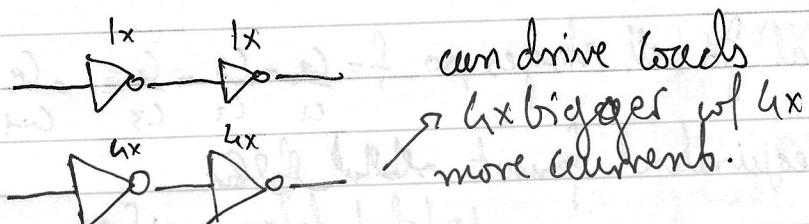
$$t_{\text{delay}} = t_{\text{inv}} \left(\frac{C_{\text{out}}}{C_{\text{in}}} \right)$$

For e.g. $\rightarrow t_{\text{inv}} = 1.2 \text{inv} \left(\frac{R_{\text{cap}} + R_{\text{out}}}{2} \right) = 40 \text{ ps}$

Single inverter



$$t_{\text{delay}} = t_{\text{inv}} \left(\frac{C_{\text{out}}}{C_{\text{in}}} \right) = (40 \text{ ps}) \left(\frac{10 \text{ pF}}{41 \text{ fF}} \right) = \underline{\underline{90.7 \text{ ns}}}$$



For 2 inverters, $N=2$

Total delay

$$= t_{\text{del1}} + t_{\text{del2}}$$

$$= T_{\text{INV}} \left(\frac{C_2}{C_1} \right) + T_{\text{INV}} \left(\frac{C_{\text{out}}}{C_2} \right)$$

To minimize total delay, can show $t_{\text{del1}} = t_{\text{del2}}$

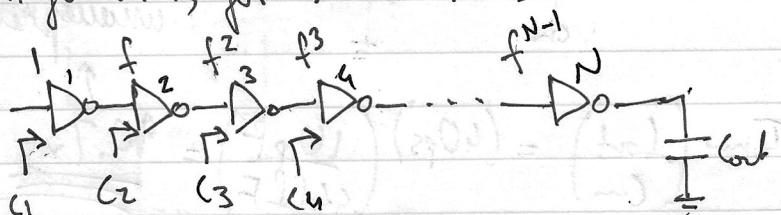
$$\text{so } \frac{C_{\text{out}}}{C_2} = \frac{C_2}{C_1} \rightarrow C_2 = \sqrt{C_{\text{out}} C_1}$$

$$\therefore \text{total delay} = T_{\text{INV}} \left(\frac{f_{\text{out}} + f_1}{C_1} \right) + T_{\text{INV}} \left(\frac{\sqrt{C_{\text{out}} C_1} + f_2}{\sqrt{C_{\text{out}} C_1}} \right)$$

$$= 2 T_{\text{INV}} \left(\sqrt{\frac{C_{\text{out}}}{C_1}} \right)$$

$$= (2)(40 \text{ ps}) \left(\sqrt{\frac{10 \text{ pF}}{4.4 \text{ fF}}} \right) = \boxed{3.8 \text{ ns}}$$

In general, for N inverters



$$t_{\text{del}} \text{ equal for all stages, so } f = \frac{C_2}{C_1} = \frac{C_3}{C_2} = \frac{C_4}{C_3} = \dots = \frac{C_N}{C_{N-1}} = \frac{C_{\text{out}}}{C_N}$$

f is the equivalent fanout ~~or load factor~~

$$\hookrightarrow \text{total delay} = N f T_{\text{INV}}$$

$$\text{if } f^N C_{\text{in}} = C_{\text{out}}$$

$$f = \left[\frac{C_{\text{out}}}{C_{\text{in}}} \right]^{1/N} = \exp \left[\frac{1}{N} \cdot \ln \left(\frac{C_{\text{out}}}{C_{\text{in}}} \right) \right]$$

E.g. $N=8$ inverters

$$f = \exp\left[\frac{\ln\left(\frac{10\text{ pF}}{4.61\text{ pF}}\right)}{8}\right] = 2.63$$

$$\text{total delay} = \frac{(8)}{N} \frac{(2.63)}{f} \frac{(40\text{ ps})}{T_{INV}}$$

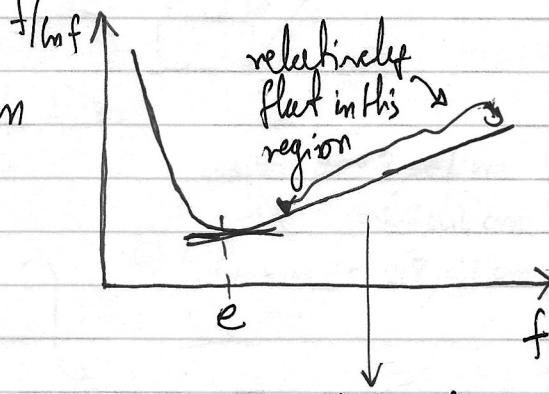
Optimum value of $f \rightarrow \text{total del} = f N T_{INV} - ①$

$$N = \frac{\ln\left(\frac{C_{out}}{C_{in}}\right)}{\ln(f)} - ②$$

$$② \rightarrow ① \quad \text{total del} = \frac{\ln\left(\frac{C_{out}}{C_{in}}\right)}{\ln(f)} f T_{INV}$$

$$\frac{\partial(\text{total del})}{\partial f} = 0 \rightarrow f = e = 2.718$$

- Four actually better than e because we neglected internal & parasitic cap.



For minimal loss of tot del we can have fewer stages w/ greater fanout.

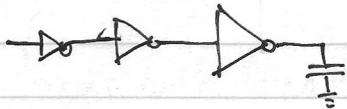
LECTURE 15

Oct 06

Last Lecture

- Driving large capacitive loads \rightarrow bus wires, clk, off-chip

end-to-end delay $\rightarrow f_{opt} = c = 2.718$

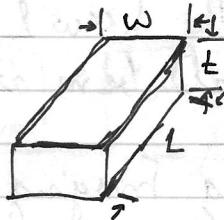


E.g. Let $N_f = 4$

$$\hookrightarrow N = \lceil \frac{\ln(10\text{pF}/2.41\text{fF})}{\ln(4)} \rceil = 5.57 \rightarrow \lfloor 5.57 \rfloor \text{ or } \lceil 5.57 \rceil$$

Interconnect.

- Resistance $R = \frac{\rho \cdot l}{t \cdot w}$



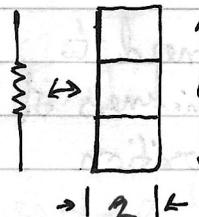
\hookrightarrow Resistivity ρ has units of $\Omega \cdot m$ or powers thereof

- Define $R_D = \rho/t$ units Ω/\square

$$\begin{aligned} \text{copper } \rho &= 1.7 \mu\Omega \cdot \text{cm} \\ \text{gold } \rho &= 2.2 \mu\Omega \cdot \text{cm} \\ \text{aluminum } \rho &= 2.8 \mu\Omega \cdot \text{cm} \end{aligned}$$

- If $R_D = 10\Omega/\square$ then

$$\hookrightarrow R = \frac{10\Omega}{\square} \times 3\Omega = \boxed{30\Omega}$$

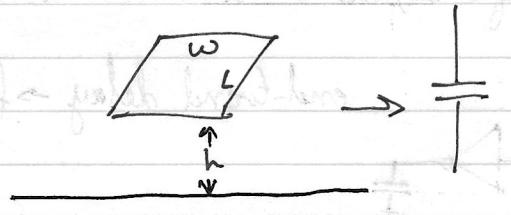


20/60

Capacitance

- Classical parallel plate (ignore fringing)

↳ $C = \epsilon_0 \frac{wL}{h}$



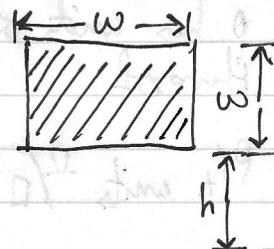
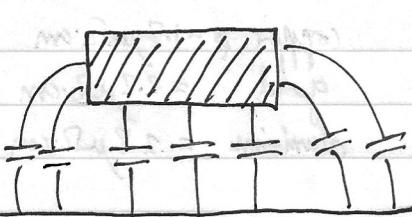
- $\epsilon_{ox} \approx 3.9 \epsilon_0 \rightarrow$ usually SiO₂ oxide

↳ $\epsilon_0 = 8.85 \times 10^{-12} \text{ F/m}$

- This model is good for longer w & L compared to h

↳ Usually $h \ll w \& h \ll L$ ends up being true, as SiO₂ layer would be in atoms range.

Including Fringing Effects



- Now we need to consider thickness of metal deposition

• $C = \epsilon_0 L \left[\frac{w}{h} + 0.77 + 1.06 \left(\frac{w}{h} \right)^{1/2} + 1.06 \left(\frac{t}{h} \right)^{1/2} \right]$

↳ within 6% for $w/h > 0.3$ and $t/h \leq 10$

* This was
a PhD's work
for Heris, ref'd
in textbook

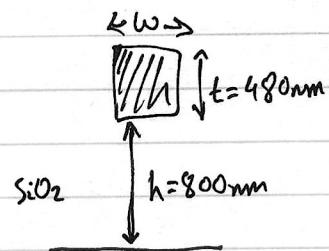
E.g. \rightarrow Metal 1 (M1)

- $W = 250\text{nm}$, $h = 800\text{nm}$, $t = 480\text{nm}$

$$\begin{aligned} C_{M1} &= \epsilon_0 \times [0.3125 + 0.77 + 0.7925 + 0.8211] \\ &= \epsilon_0 \times [0.3125 + 2.384] \end{aligned}$$

\hookrightarrow per unit length

$$= 0.094 \text{ fF}/\mu\text{m}$$

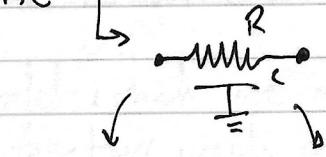


\hookrightarrow fringe cap is 8x parallel plate

- A wire is a distributed RC

- We want to take a wire and break it into N segments

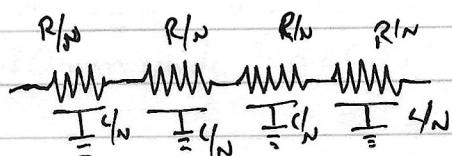
$\hookrightarrow N$ doesn't have to be very high in order to get a good estimate of τ



$$R = R_w l \quad C = C_w l$$

$$R_w = \frac{R}{l} \text{ }\mu\text{m}$$

$$C_w = \frac{C}{l} \text{ }\mu\text{m}$$



Edmore Delay \rightarrow RC Ladder Networks

$$\tau = \left(\frac{C}{N}\right)\left(\frac{R}{N}\right) + \left(\frac{C}{N}\right)\left(\frac{2R}{N}\right) + \left(\frac{C}{N}\right)\left(\frac{3R}{N}\right) + \dots + \left(\frac{C}{N}\right)\left(\frac{NR}{N}\right)$$

$$= \frac{CR}{N^2} (1 + 2 + 3 + \dots + N) \quad \text{as } N \rightarrow \infty, \boxed{\tau = \frac{RC}{2}} \rightarrow \tau = \frac{R_w C_w l^2}{2} \frac{l^2}{\frac{N^2}{2}}$$

$$= \frac{CR}{N^2} \left(\frac{N(N+1)}{2} \right) = \boxed{\frac{CR(N+1)}{2N}}$$

big implications

LECTURE 16

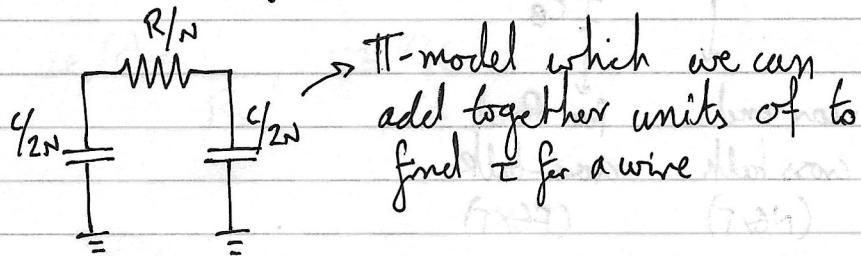
Oct 08

Elmore Delay (of a Wire)

- $T = \frac{RC}{2}$ as $N \rightarrow \infty$ (# of segments)

- $T = \frac{(R_w l)(C_w l)}{2} = \frac{R_w C_w l^2}{2} \rightarrow$ keep wires short!

- When simulating in SPICE we don't use $N \rightarrow \infty$, we can actually get good results w/ small N



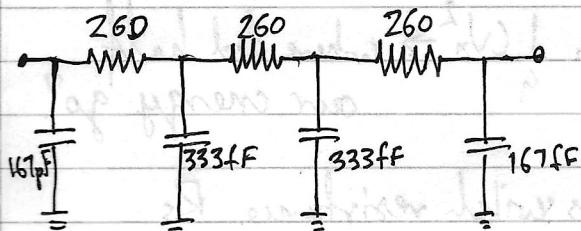
- Usually $N=3$ is good enough, captures all the major behaviours.

E.g. 5mm long wire, 320 mm wide, $R_D = 0.05 \Omega/\square$
 $C_w = 0.2 \text{ fF}/\mu\text{m}$

↳ Construct a 3-segment T-model for this wire.

$$R = R_D \left(\frac{l}{w} \right) = (0.05) \left(\frac{500}{0.32} \right) = 781 \Omega$$

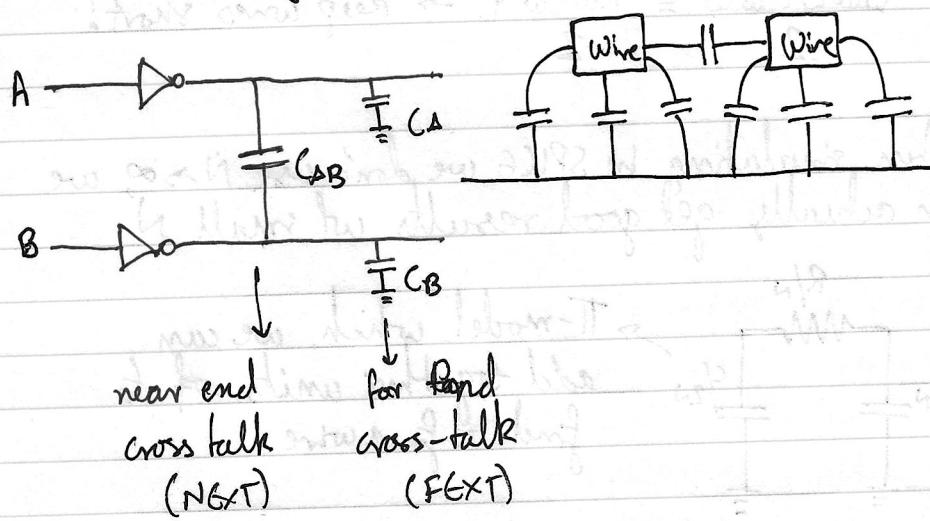
$$C = C_w l = (0.2) (5000) = 1000 \text{ fF}$$



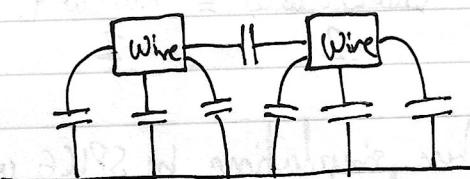
Now compare delay of 3-segment model to
 $RC/2 = 390 \text{ ps} = T_1$

$$T_2 = (333)(260) + (333)(520) + (167)(781) = 390 \text{ ps} !$$

Gross-Talk Delay



Cross-section

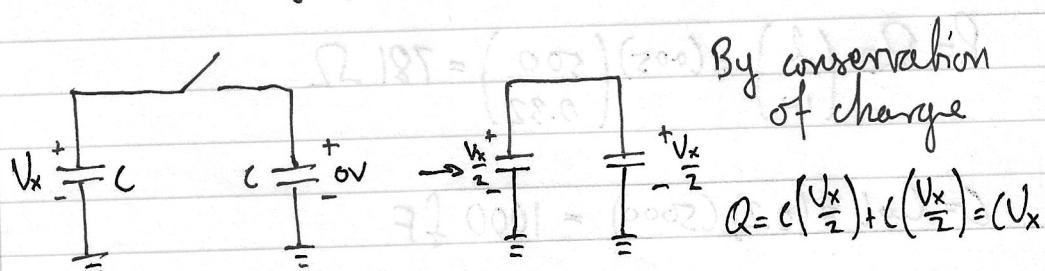


Power Dissipation

- Capacitors \rightarrow do not dissipate power but.....

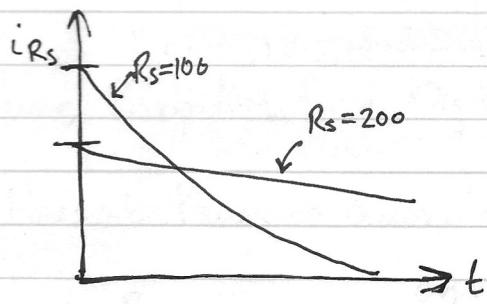
$$\text{Energy} \rightarrow E = \frac{1}{2} CV_x^2$$

$$\text{Charge} \rightarrow Q = CV_x$$



But $E_{\text{total}} = \frac{1}{2}(2C)\left(\frac{V_x}{2}\right)^2 = \frac{1}{4}(V_x)^2 \rightarrow$ where did half our energy go?

↳ Energy dissipated across switch resistance R_s .



If R_S is small then the initial current is large but for a short time

On the other hand, if R_S is large then the initial current is small but takes longer to transfer charge over.

$$i_{RS}(0) = \frac{V_x}{R_S}, \quad T = \frac{R_S C}{2}$$

LECTURE 17

Oct 09

Power Dissipation for a Digital Chip

① Dynamic Power \rightarrow occurs only on transitions

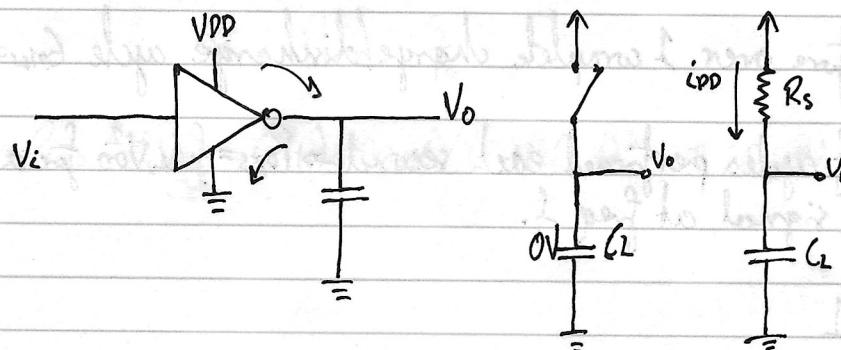
② Static Power \rightarrow occurs ~~even~~ ^{only} if signal clock held high or low

- Dynamic power consists of both

- \hookrightarrow Capacitive power [80% of DP] \rightarrow thermal design power

- \hookrightarrow Direct path power [20% of DP]

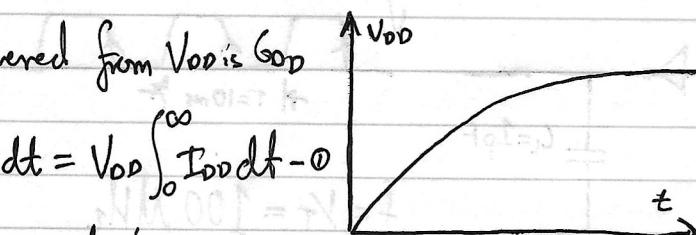
Dynamic Capacitive Power Low to High



- Energy delivered from V_{DD} is G_{DD}

$$E_{DD} = \int_0^{\infty} V_{DD} I_{DD} dt = V_{DD} \int_0^{\infty} I_{DD} dt - 0$$

Recall $q = CV$ $\& i = dq/dt$



$$Q = \int i dt + C$$

$$\therefore \text{using } ① \quad G_{DD} = V_{DD} Q = V_{DD} C_L V_{DD} = C_L V_{DD}^2$$

$$\text{BUT energy stored on } C_L, E_{C_L} \rightarrow E_{C_L} = \frac{1}{2} C_L V_{DD}^2$$

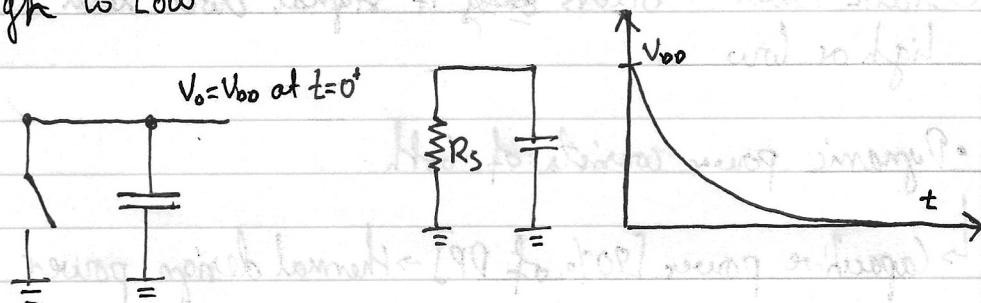
on 60

1397

$$\text{Energy dissipated in resistor} \rightarrow E_{Diss} = \frac{1}{2} C_L V_{DD}^2 = E_{Diss}$$

- We see that E_{Diss} is independent of R_s whether it is linear (MOS transistor) or not

High to Low



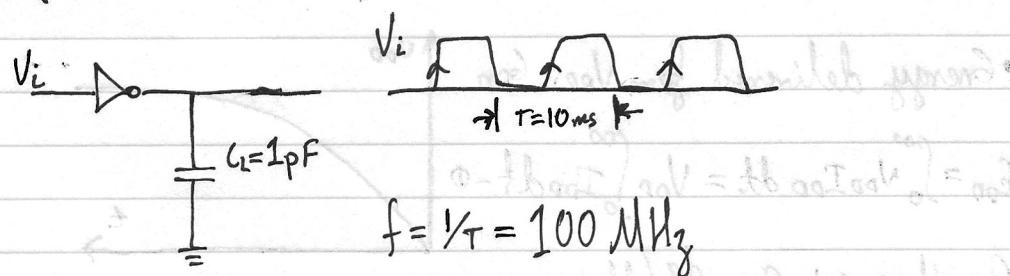
$$\text{Energy before in cap} \rightarrow E_{CL} = \frac{1}{2} C_L V_{DD}^2$$

$$\text{after} \rightarrow E_{CL} = 0; E_{Diss} = \frac{1}{2} C_L V_{DD}^2$$

$$\therefore \text{Therefore over 1 complete charge/discharge cycle } E_{Diss} = C_L V_{DD}^2$$

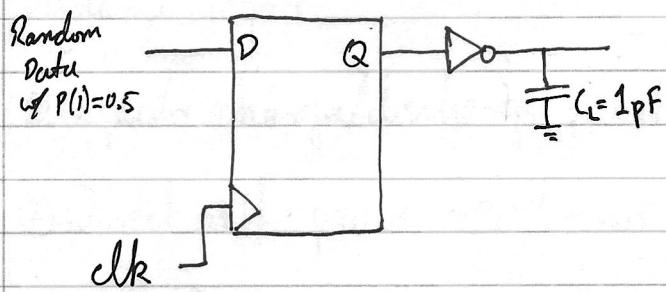
- For f cycles performed each second $\rightarrow P_{Diss} = f C_L V_{DD}^2$ for a clock signal at freq f .

E.g. 1



$$P_{Diss} = f C_L V_{DD}^2 = (100 \times 10^6)(1 \times 10^{-12})(3.3^2) = 1.1\text{ mW}$$

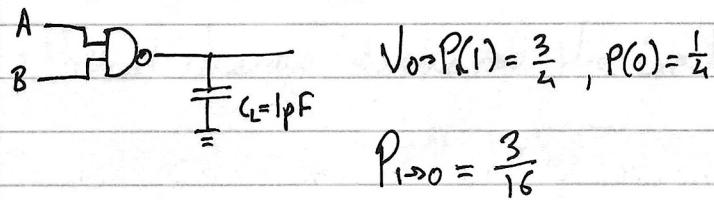
E.g. 2



$$P_{1 \rightarrow 0} = P(1)P(0) = (0.5)(0.5) = \frac{1}{4}$$

$$P_{DISS} = P_{1 \rightarrow 0} f C_L V_{DD}^2 = \left(\frac{1}{4}\right) (100e6) (1e-12) (3.3)^2 = 0.27 \text{ mW}$$

E.g. 3



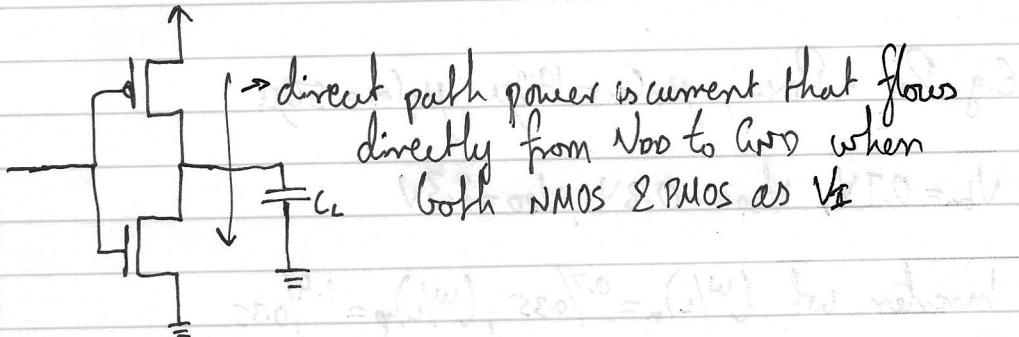
$$P_{1 \rightarrow 0} = \frac{3}{16}$$

$$P_{DISS} = P_{1 \rightarrow 0} f C_L V_{DD}^2 = \left(\frac{3}{16}\right) (100e6) (1e-12) (3.3)^2 = 0.20 \text{ mW}$$

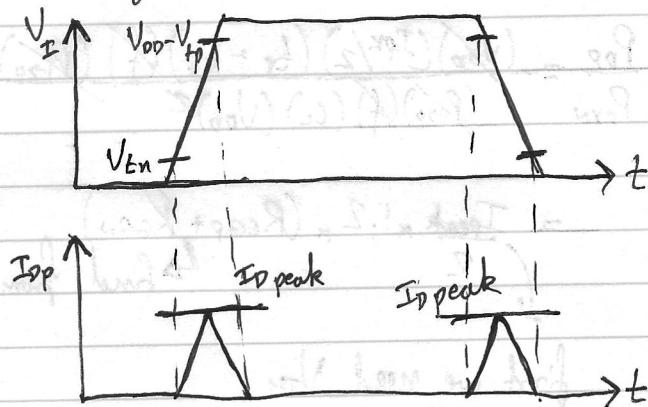
Dynamic Power

① Capacitor charge/discharge power $\sim 80\%$

② Direct path power $\sim 20\% \rightarrow$ not deterministic ratio



- Doesn't charge/discharge the load capacitance



- V_{TN} is the inverter threshold

$$\hookrightarrow I_{D\text{ peak}} = \frac{\mu_n C_{ox}}{2} (\omega/L) n (V_{dd} - V_{tn})^2$$

- Direct path energy $G_{DP} = V_{DD} (Q_r + Q_f)$

$\hookrightarrow Q_r \& Q_f$ are charge delivered for rising & falling edges

$$\bullet Q_r = I_{avg} \times \Delta t = I_{peak}/2 (t_r + t_f)$$

$$\bullet \text{Similarly } Q_f = \frac{I_{\text{peak}}}{2} (t_f)$$

$\therefore E_{DD} = \frac{V_{DD} \times I_{pk}}{2} (t_r + t_f) \rightarrow \text{now we can find corresponding } P_{DD}$

$$\hookrightarrow P_{DD} = \frac{V_{DD} I_{pk} (t_r + t_f) (f) (P_{i \rightarrow o})}{2}$$

E.g. $P_{DD} + P_{DYN}$, $\mu_n C_{ox} = 190 \mu$, $\mu_p C_{ox} = 50 \mu$

$$V_{bn} = 0.7V, V_{bp} = -0.8V, V_{DD} = 3.3V$$

$$\text{Inverter w/l } (\omega/L)_n = 0.7/0.35, (\omega/L)_p = 1.4/0.35$$

Find ratio of P_{DD}/P_{DYN}

$$\frac{P_{DD}}{P_{DYN}} = \frac{\left(\frac{V_{DD}}{2} \right) \left(I_{pk}/2 \right) (t_r + t_f) f_f (P_{i \rightarrow o})}{(P_{i \rightarrow o}) f_f (L) (V_{DD})^2} = \frac{I_{pk} (1.2 R_{eq})}{2 L V_{DD}}$$

$$= \frac{I_{\text{peak}} \times 1.2 \times (R_{eqn} + R_{eqp})}{2} \quad \hookrightarrow \text{find from w/l ratios}$$

* first we need V_{TH}

$$V_{TH} = \frac{V_{DD} + V_{bp} + V_{bn}}{1+r}, r = \sqrt{\frac{\mu_p (\omega/L)_p}{\mu_n (\omega/L)_n}} = 1.46 V$$

$$I_{\text{peak}} = \left(\frac{\mu_n C_{ox}}{2} \right) \left(\frac{W}{L} \right)_n (V_{TH} - V_{bn})^2 = 109.7 \mu A$$

$$R_{eqn} = \frac{2.5}{\mu_n C_{ox} (\omega/L)_n (V_{DD} - V_{bn})} = 2.53 k\Omega$$

$$\hookrightarrow \text{Similarly } R_{eqp} = 5 k\Omega$$

Gives us $\frac{P_{DP}}{P_{DYN}} = \frac{(109.7\mu)(1.2)(2.53k + 5k)}{(2)(3.3)} \approx 0.15 = 15\%$

LECTURE 19

Oct 16

CMOS Circuit Families

① Static CMOS

* Non-idealities *
on midterm

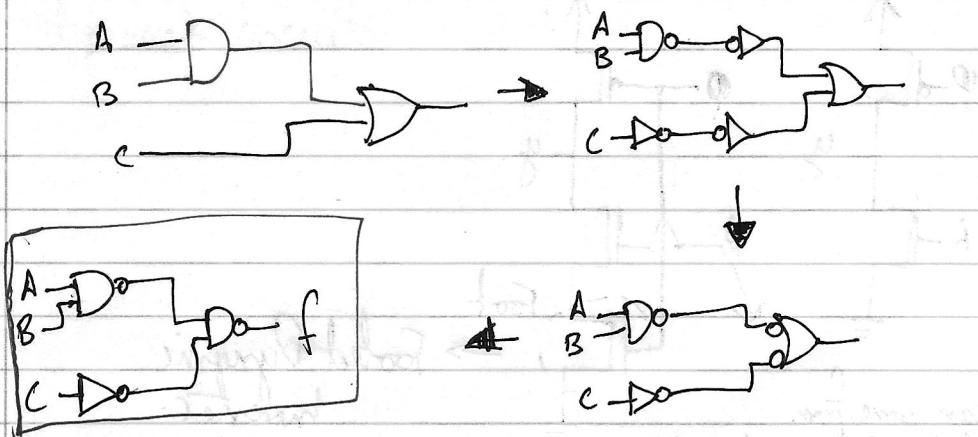
② Dynamic Circuits

③ Pass Transistor Circuits

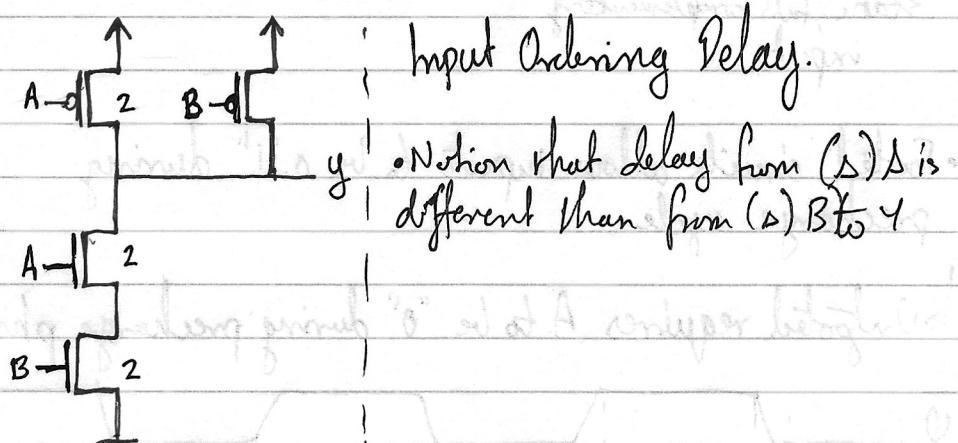
Brown 241 textbook
examples

- Don't forget De Morgan's Law!

- "Pushing bubbles"

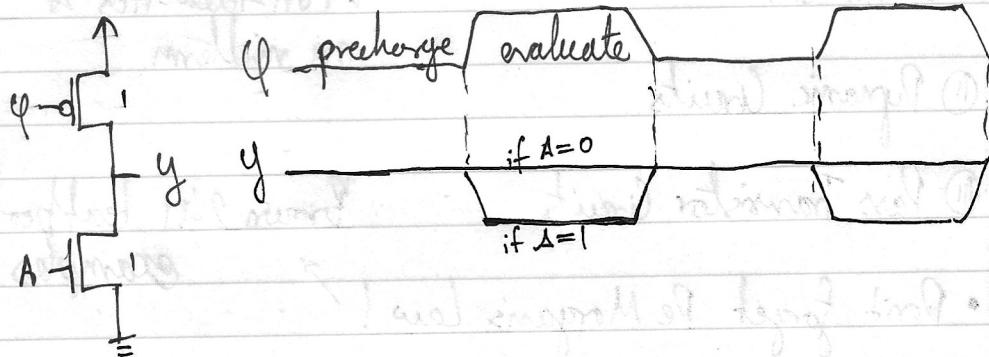


Input Ordering Delay.

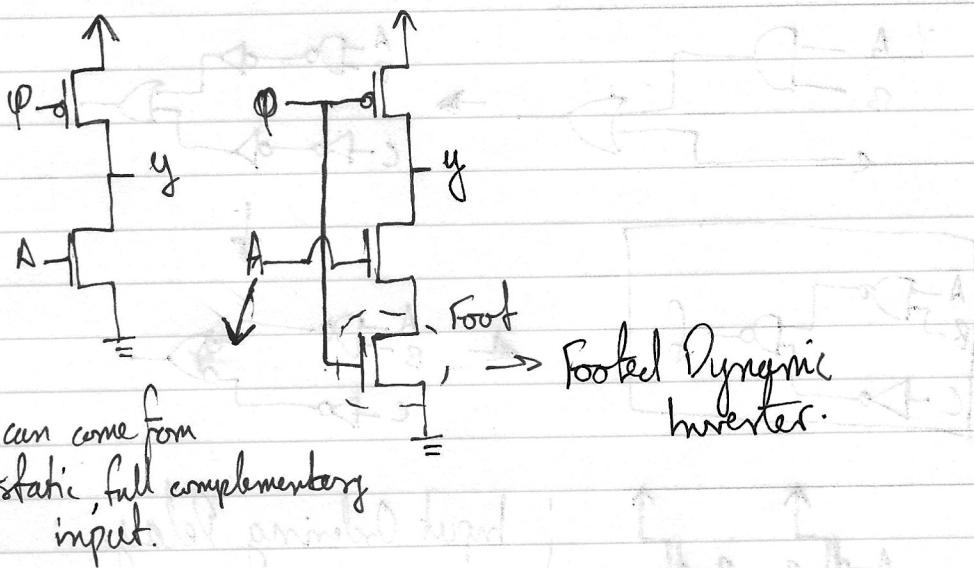


Dynamic Circuits

- Requires clock φ to operate

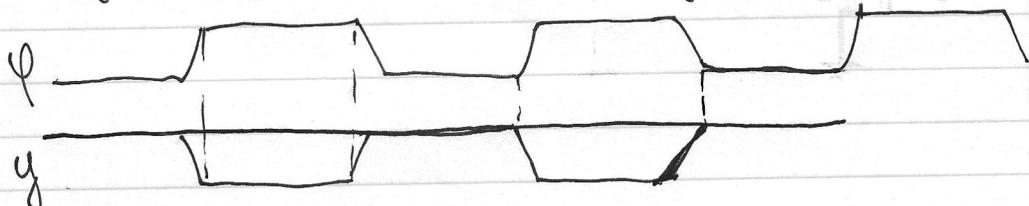


- Two different topologies:



- Footed circuits allow inputs to be a "1" during precharge cycle

↳ Unfooted requires A to be "0" during precharge phase



- However there is a strict requirement of inputs that they must be monotonically rising
- Inputs must: low and stay low \rightarrow allowed if low and rise high \rightarrow footed, not allowed high and stay high if unfooted.
- High and going low is not allowed!
 - ↳ Since input is high & precharged and cannot go high again if discharged.
- Footed circuits allow outputs to drive other dynamic circuits.

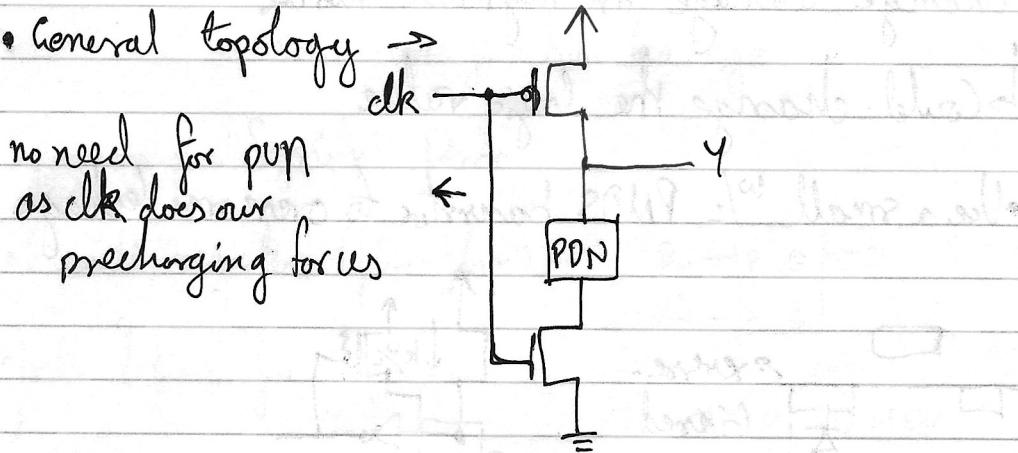
LECTURE 20

Oct 20

Dynamic Circuits

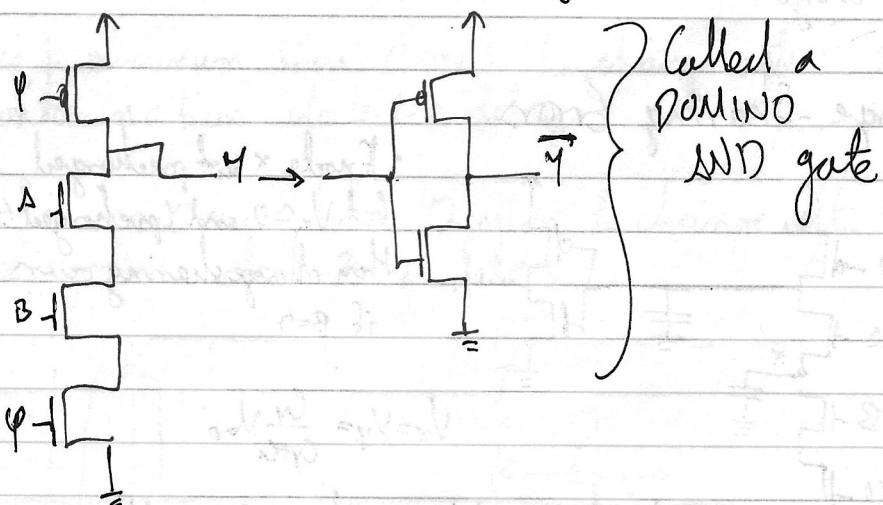
- Requires a clock to operate \rightarrow has precharge & eval phases
- Output precharged to V_{DD} during precharge phase
- Output only valid during EVAL phase

- General topology \rightarrow



- Connect Y directly to another gate as it will generate monotonically falling outputs

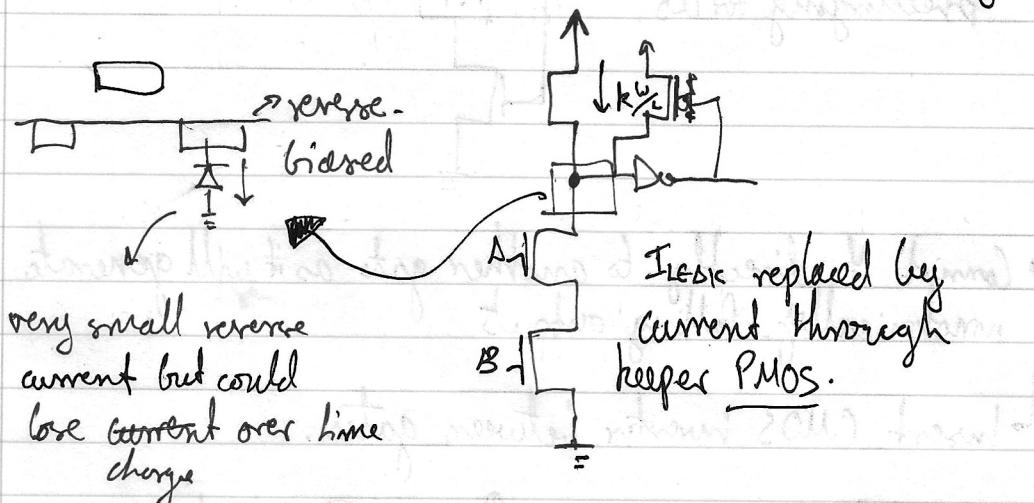
↳ Insert CMOS inverter between gates.



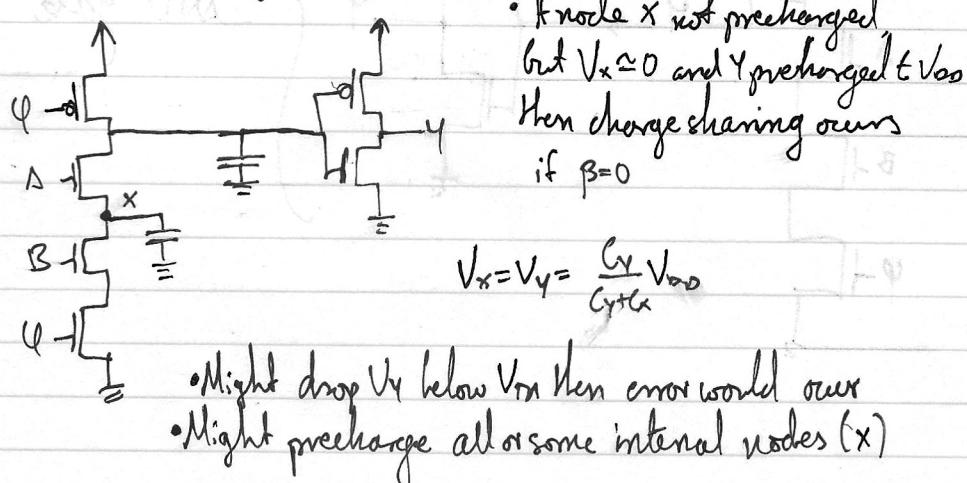
- Domino gates are inherently non-inverting
- If the last domino gate in a chain we can use an inverter

Keepers (adding sink off bypassing bus):

- If φ is stopped (clk stops ticking) in EVAL phase, change leakage on dynamic mode
 - ↳ Could change the logic value
- Use a small w/l PMOS transistor to overcome leakage



Charge-Sharing Errors



- If node x not precharged, but $V_x \approx 0$ and y precharged to V_{DD} . Then charge sharing occurs if $\beta = 0$

LECTURE 21

Oct 22

Static Power Dissipation

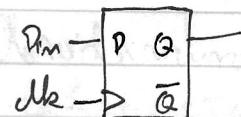
MOSAID

- Power dissipation is when logic is not switching

- In CMOS \rightarrow
 - ① Subthreshold current ($V_{GS} \approx V_{TN/P}$) \rightarrow FinFET gate all around
 - ② Gate leakage
 - ③ Reverse-bias junction leakage

Synchronous Systems Design

- Recall registers (flip-flops)



- Let's define for register

t_{su} , t_h , t_{cq} , and t_{pcq}

- t_{su} is time D is held before \nearrow

$CLK 0 \rightarrow Q = 0$

$CLK 1 \rightarrow Q = 1$

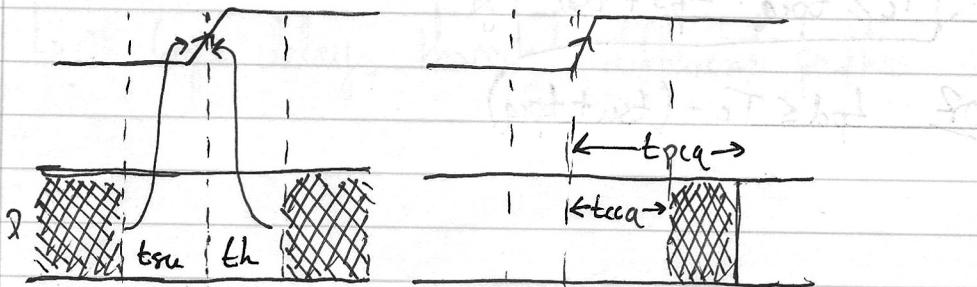
$\nearrow \rightarrow Q = 1$

$\searrow \rightarrow Q = 0$

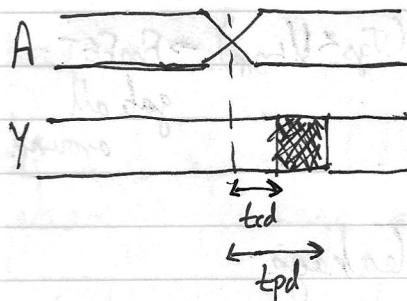
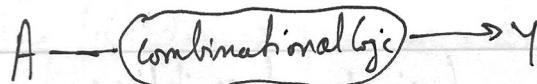
- t_h is time D is held after \nearrow

- t_{cq} is minimum time Q needs to change after \nearrow contamination clock-2-Q delay.

- t_{pcq} is maximum time Q needs to change after \nearrow propagation clock-2-Q delay.

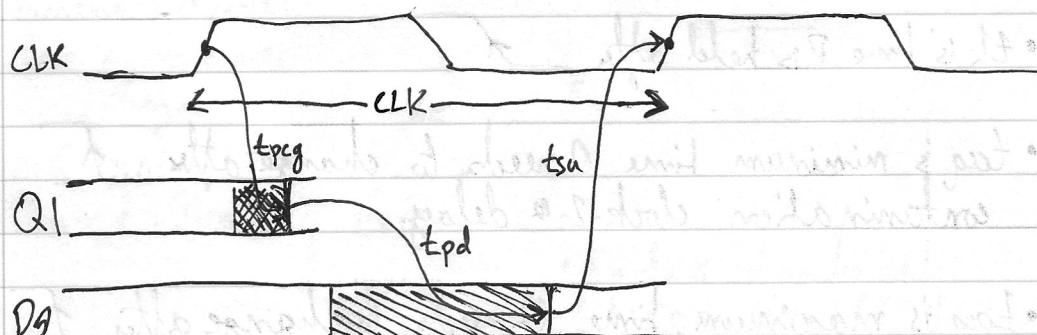
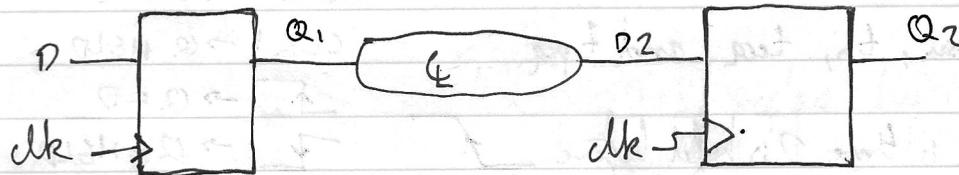


- Similar for combinational logic



- t_{cd} is combinational logic contamination delay
- t_{pd} is combinational logic propagation delay

Max Delay Constraint



$$T_c \geq t_{pcq} + t_{pd} + t_{su} \quad \star$$

$$\text{or } t_{pd} \leq T_c - (t_{su} + t_{pcq})$$

LECTURE 22

Oct 23

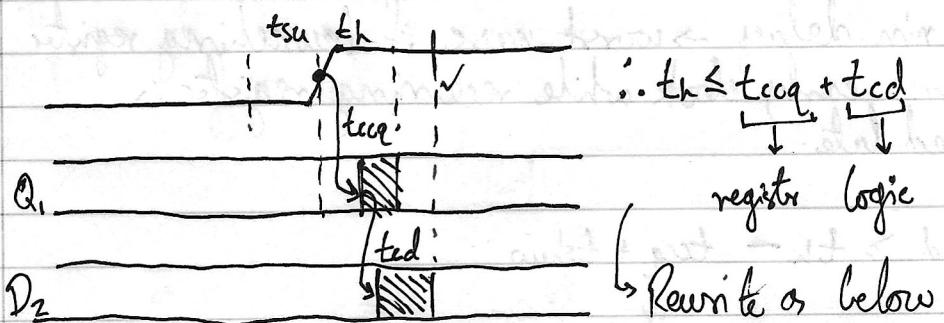
Max Delay Constraint

- $t_c \geq t_{pcq} + t_{pd} + t_{su}$

- If not satisfied, known as "setup time failure" or "max delay failure"

Minimum Delay Failure

- If "Q" not held enough often \rightarrow Then data may propagate through 2 registers on one clock edge.



- $t_{cd} \geq t_h - t_{pcq} \rightarrow$ If $t_{pcq} \geq t_h$ then you can use back-to-back registers.

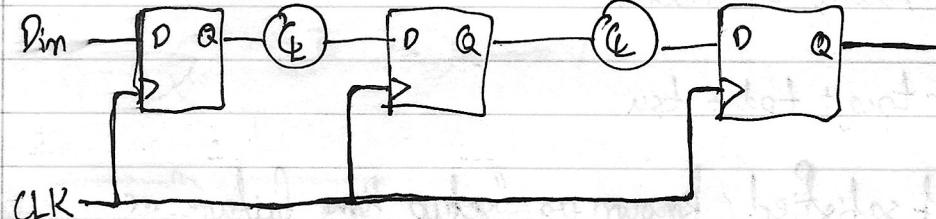
- If not satisfied, logic must be inserted b/w registers.

- $t_{cd} \neq t_{pd}$ can be due to PVT variations

\hookrightarrow PVT \rightarrow process, voltage, temperature

\hookrightarrow OR logic delays through different paths

- Consider: CLOCK SKEW t_{skew}



- For max delay \rightarrow worst case is if launching register receives late clock while receiving register is clocked early.

$$\hookrightarrow t_{pd} \leq T_c - (t_{pcq} + t_{su} + t_{skew})$$

- For min delay \rightarrow worst case if launching register receives early clock while receiving register is clocked late.

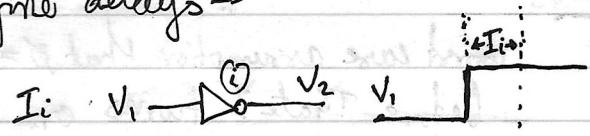
$$\hookrightarrow t_{cd} \geq t_h - t_{cq} + t_{skew}$$

LECTURE 23

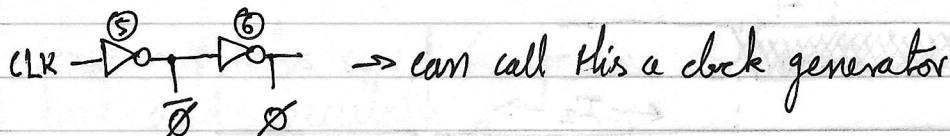
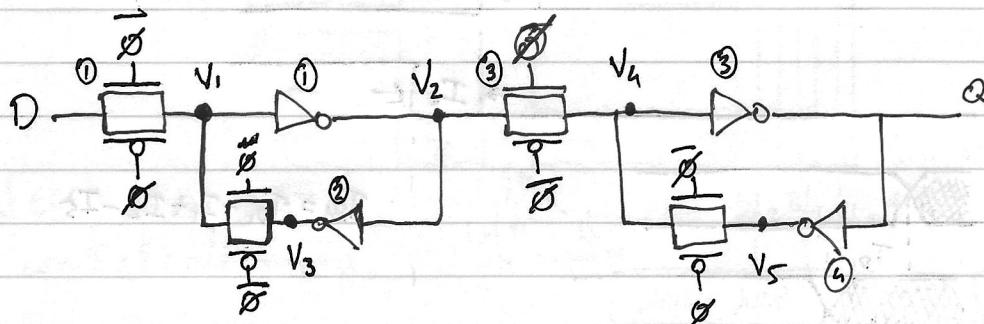
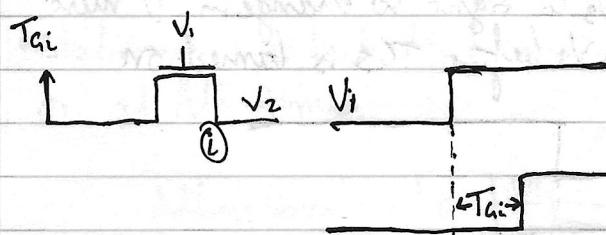
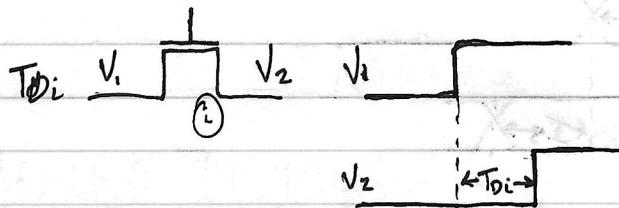
Nov 03

Estimating Delays t_{pq} , t_{su} , t_L

- Define delays →



* Not an accurate method for design, but gives us a good intuition



→ can call this a clock generator

80 ns

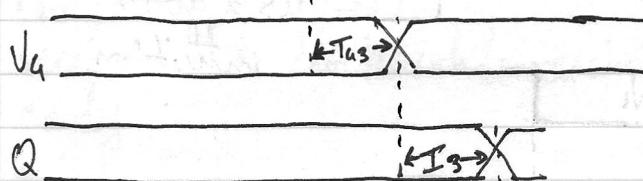
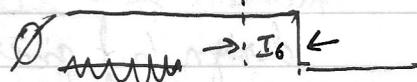
88 ns (Total)

t_{pq} \rightarrow

$$\text{CLK} \quad \therefore t_{pq} = I_5 + I_6 + T_{G3} + I_3$$



worst case assumption that $\phi=1$
before TG3 turns on



$t_{su} \rightarrow$ Do not want V_1, V_2, V_3 to "fight" so change in "D" must settle through to V_3 before TG3 is turned on

$$\text{CLK} \quad \text{---} \quad \text{---} \quad \text{---}$$

$$\phi \quad \rightarrow I_5 \quad \leftarrow$$

$$\phi \quad \rightarrow I_6 \quad \leftarrow$$

$$D \quad \text{---}$$

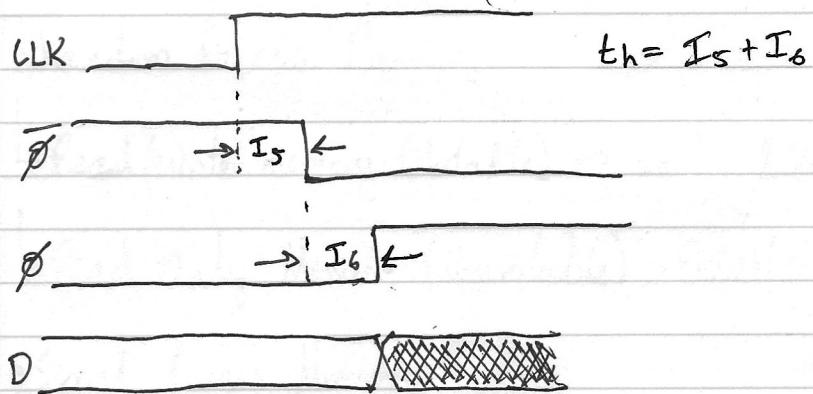
$$t_{su} = T_{D1} + I_1 + I_2 - I_5$$

$$V_1 \quad \text{---} \quad T_{D1} \quad \text{---}$$

$$V_2 \quad \text{---} \quad I_1 \quad \text{---}$$

$$V_3 \quad \text{---} \quad I_2 \quad \text{---}$$

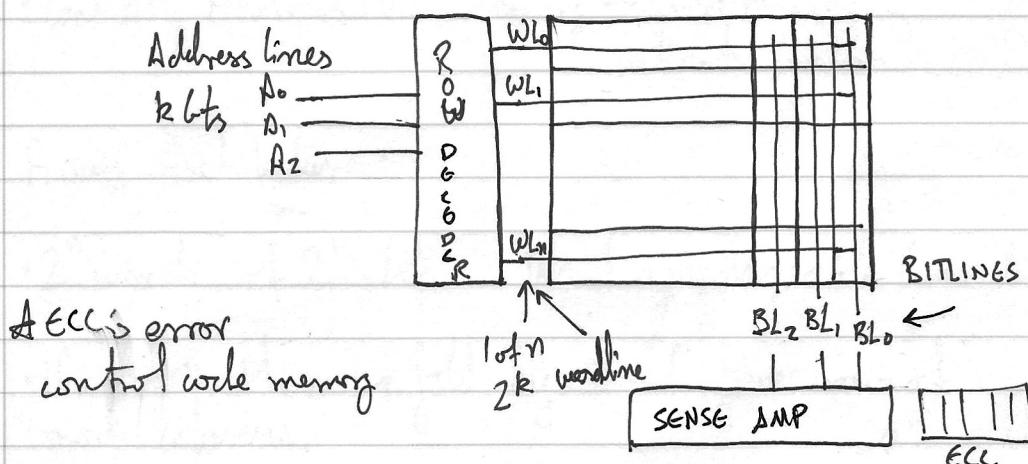
$t_h \rightarrow$ Do not want D to change until after TG0 is fully turned off



MEMORY SYSTEMS

- Static refers to fact that memory must be powered on at all times

CORE ARRAY



- Sense amp helps speed by looking at signal development time and speculates a change.

↳ Say BL₁ goes from 0 to 100mV then sense amp will amplify signal and commit to sending VDD voltage

↳ If we predicted wrong just read again, chances are error was just caused by VDD/GND noise

LECTURE 24

Nov 05

MEMORY ARRAYS

- Random Access Memory

- ↳ Read/Write Memory (Volatile) → Static & Dynamic RAM

- ↳ Read Only Memory (Nonvolatile) → PROM/EPROM/EEPROM

- Serial Access Memory

- ↳ Shift Registers → SIPO/PISO

- ↳ Queues → FIFO/LIFO

- Content Addressable Memory → used mainly in networks/routers.

Array Architecture

- 2^n words of 2^m bits Want a manageable t_{more delay}

- If $n \gg m$ then we fold by 2^k into fewer rows of more columns

- Good regularity and easy to design

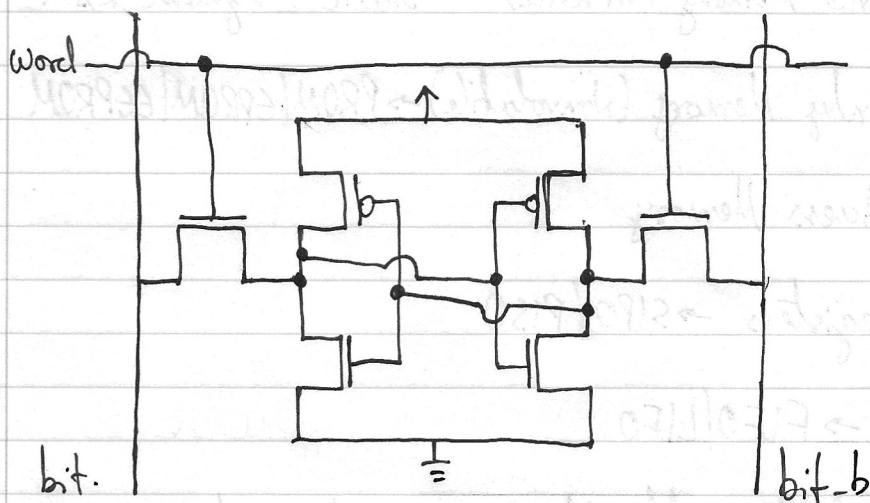
- Very high density if good cells are used.

20.4

As required

6T SRAM Cell

- Cell size accounts for most of array size
- 6T SRAM used in most commercial devices

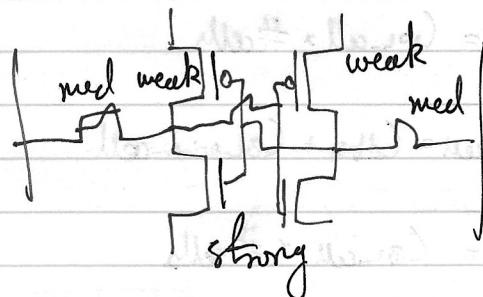


SRAM Read

- Precharge both bitlines high THEN turn off on wordline
- One of two bitlines will be pulled down by the cell.
 - E.g. $A=0, A_b=1$
 - ↳ bit discharges, bit-b stays high
 - ↳ A jumps up high slightly
 - Read stability, A must not flip $\therefore N_1 \gg N_2$

SRAM Write

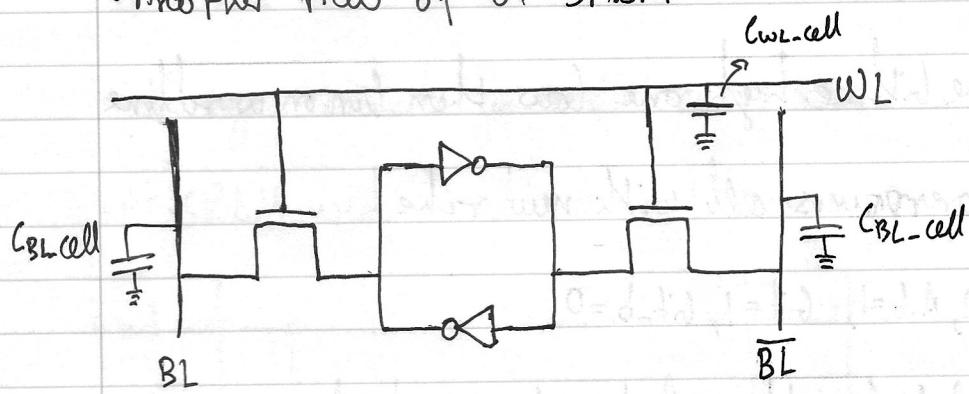
- Drive one bitline high, one low then turn on wordline
- Bitline overpowers cell with new value
 - Eg. $A=0, A-b=1, G_{it}=1, G_{it-b}=0$
 - ↳ Force $A-b$ low, then A has to rise high
- Writeability \rightarrow need overpower feedback inverter
 - ↳ $N \gg P$



Thin CMOS

- In nanometer CMOS
 - ↳ Avoid bends in polysilicon and diffusion
 - ↳ Orient all transistors either horizontally or vertically
- Lithographically friendly or thin cell layout fixes this
 - ↳ Also reduces length and capacitance of G_b lines.

• Another view of 6T SRAM



SRAM Capacitance

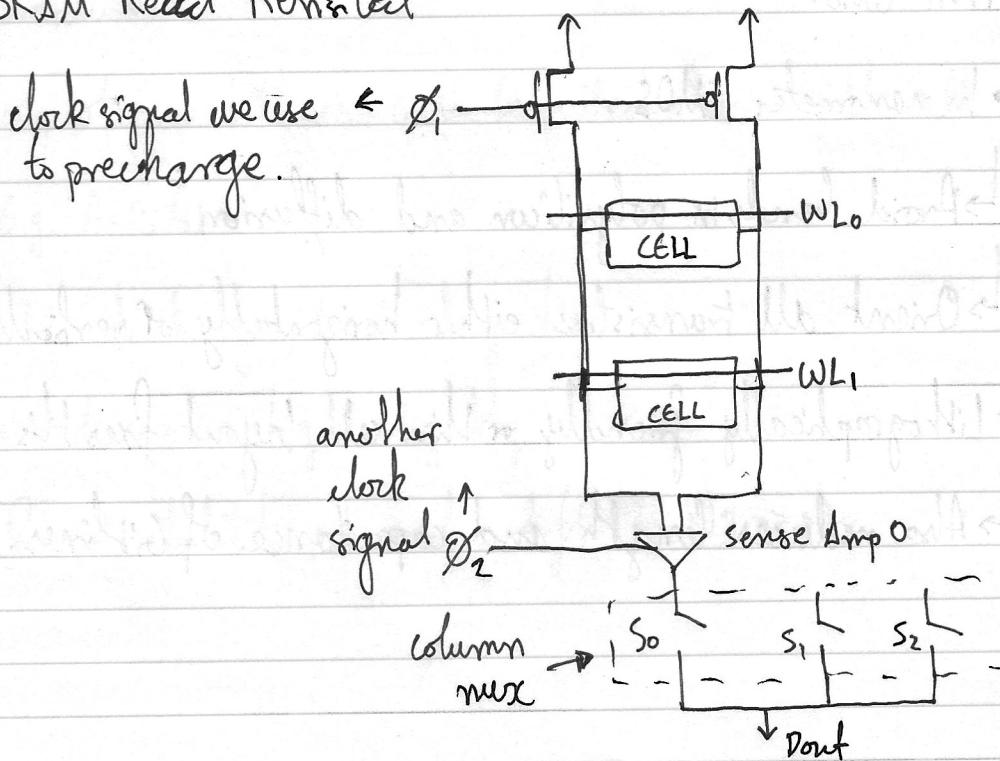
- $C_{WL\text{-cell}} = C_{g_3} + C_{g_4} + C_{WL\text{-wire-cell}}$

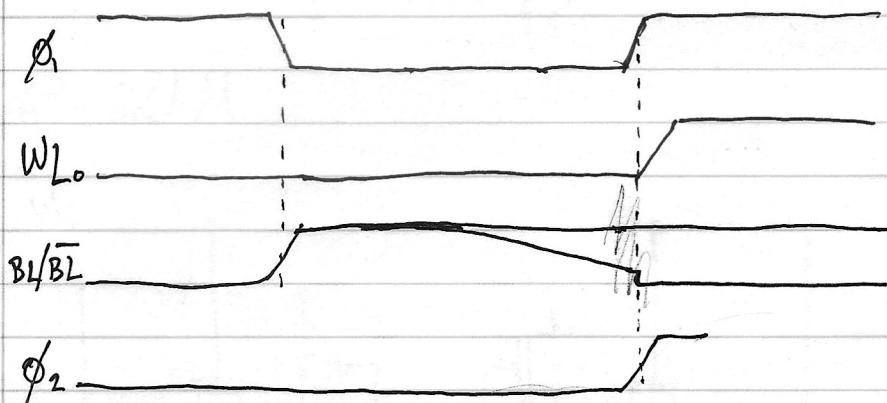
- $C_{WL} = C_{WL\text{-cell}} \times \# \text{cells}$

- $C_{BL\text{-cell}} = C_{db4} + C_{BL\text{-wire-cell}}$

- $C_{BL} = C_{BL\text{-cell}} \times \# \text{cells}$

SRAM Read Revisited

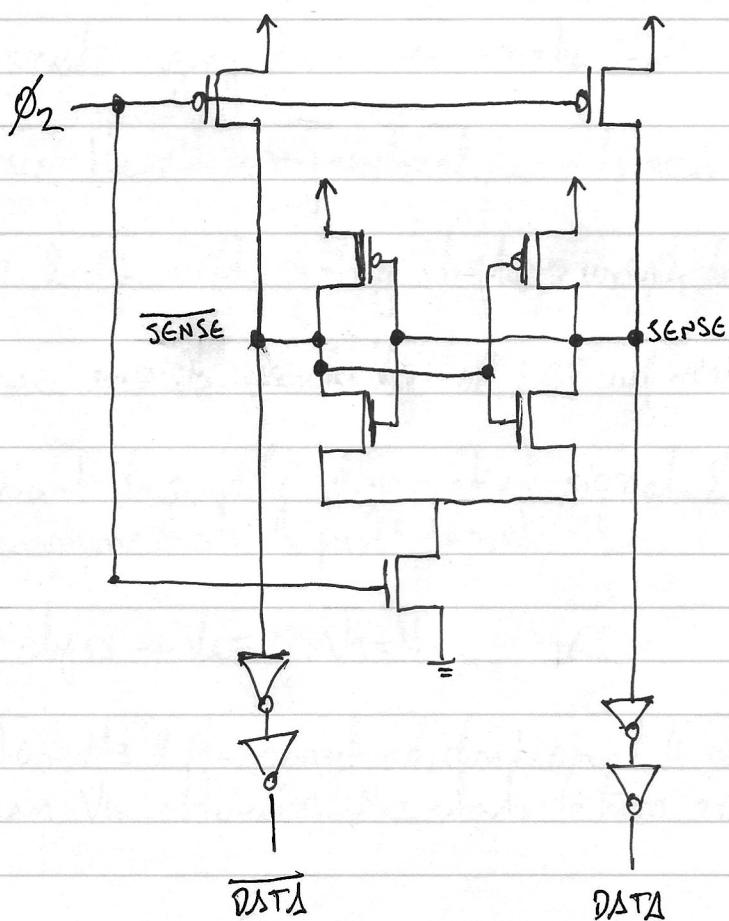




• ΔV might be only 200mV

• Reason for sense-amp is that C_{BL} large compared to size of transistors in cell so it would take much longer to be pulled down.

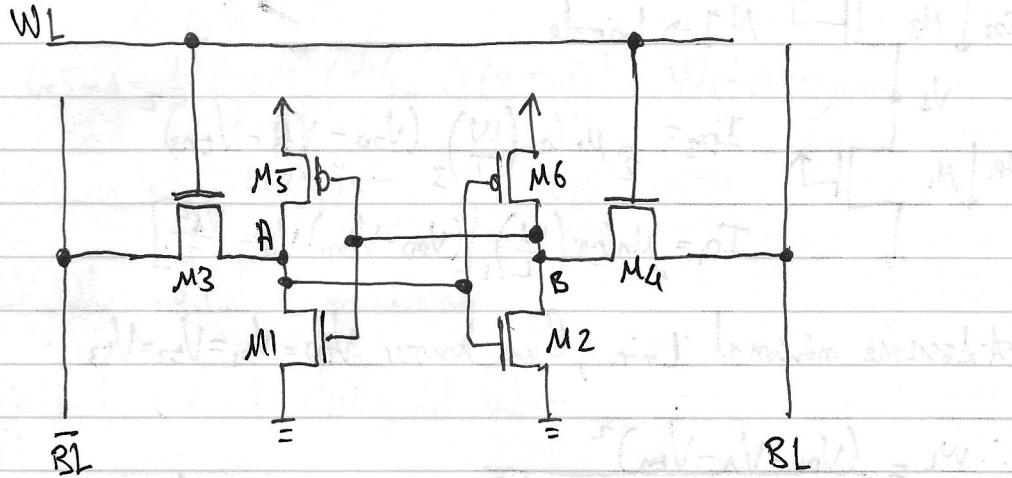
Typical Sense Amp



LECTURE 25

Nov 06

GT-SRAM



* Bitlines are differential

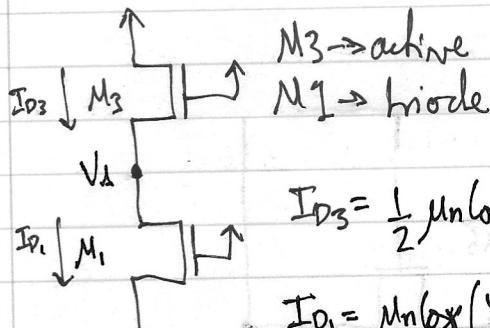
SRAM Read

- Transistor sizing (for no error) \rightarrow aka READ STABILITY
- Since the cell is symmetrical, we only need to M1, M3, M5
- BL & \bar{BL} both going high \rightarrow use ϕ_1 to precharge.
- Sense amp is essentially OR AMP w/ really high CMRR
 ↳ Want to amplify differential signal & cancel out common-mode "signal" (noise)

Analysis $\rightarrow V_A = 0, V_B = V_{DD}$

- When M3 & M4 turn on, don't want to pull V_A higher than V_{DD} otherwise M2 starts to turn on.

- EHS becomes:



$M_3 \rightarrow$ active
 $M_1 \rightarrow$ triode

$$I_{D3} = \frac{1}{2} \mu_n C_s \left(\frac{W}{L} \right)_3 (V_{DD} - V_A - V_{TN3})^2$$

$$I_{D1} = \mu_n C_s \left(\frac{W}{L} \right)_1 [(V_{DD} - V_{TN1}) V_A - \frac{V_A^2}{2}]$$

* Assume minimal L_{min} , we know $V_{TN} = V_{T1} = V_{T2} = V_{T3} \dots$

$$\therefore \frac{W_1}{W_3} = \frac{(V_{DD} - V_A - V_{TN})^2}{2[(V_{DD} - V_{TN}) V_A - \frac{V_A^2}{2}]}$$

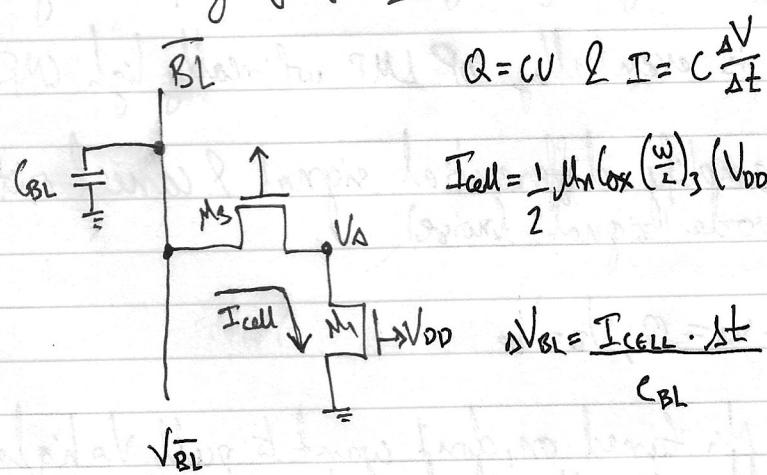
* As a designer
I should want to
start w/ smallest
value i.e. W_3

e.g. $V_{DD} = 2.5$, $V_{TN} = 0.5V$, $V_A = 0.5V$

$$\hookrightarrow \text{Then } \frac{W_1}{W_3} = 1.3 \rightarrow W_1 = 1.3 W_3$$

SRAM Read

- Transistor sizing for SPEED



$$Q = CV \quad \& \quad I = C \frac{\Delta V}{\Delta t}$$

$$I_{cell} = \frac{1}{2} \mu_n C_s \left(\frac{W}{L} \right)_3 (V_{DD} - 2V_{TN})^2$$

$$I_{cell} \downarrow M_2 \rightarrow V_{DD} \quad \Delta V_{BL} = I_{CELL} \cdot \frac{t}{C_{BL}}$$

E.g. $\mu_{nLc} = 200 \mu A/V^2$, $V_{thn} = 0.5V$, $V_A = 0.5V$, $V_{DD} = 2.5V$, $C_{BL} = 1 pF$

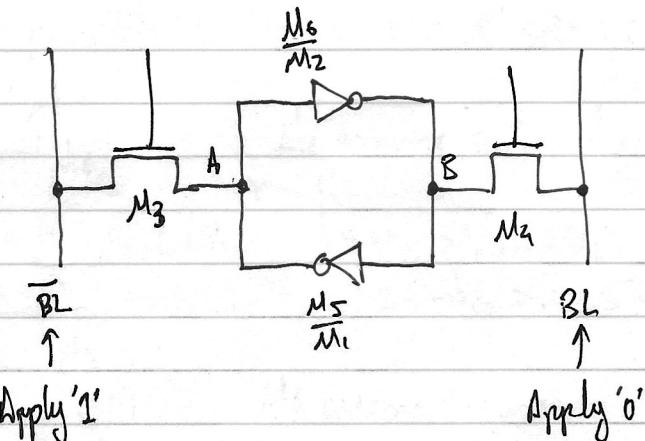
- Find size of $M_1 \& M_3$ such that $\Delta V_{BL} = 0.5V$ in $1\mu s$

$$0.5 \text{ mV} = \frac{1}{2}(200 \times 10^{-6})(W_3/L_{min}) (2.5 - 1)^2 \rightarrow W_3 = 0.55 \mu m$$

$$\text{Thus } W_1 = (1.3)(0.55 \mu m) = 0.72 \mu m$$

SRAM Write Operations

- Want to find W_{M6} and W_{M5}
- ↳ These are not critical for read
- Say we want to write a '1' into SRAM cell.



Recall for read stability $M_3 \& M_1$ sized so to never exceeds threshold of $\frac{M_6}{M_2}$ when $\overline{BL} = '1'$

- Need to go through whole analysis but

$$\underline{\underline{W_{M4} = 1.15(W_{M5})}}$$

SRAM Supplementary Notes

SRAM Write Operation

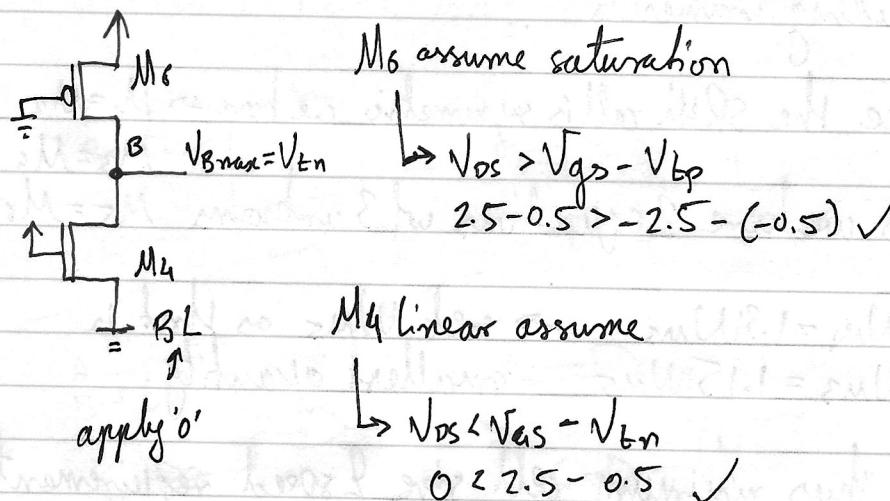
- Want W for M_6 & M_5
not critical to our
read but impacts
write operation

- E.g. if we wish to write
'1' into SRAM cell

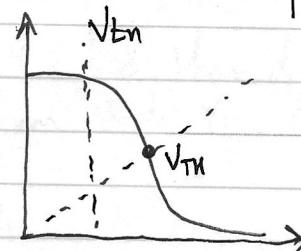
$\hookrightarrow \bar{BL} \rightarrow$ apply '1', $BL \rightarrow$ apply 0

- Recall for read stability M_3 & M_1 sized so that node A
can't exceed M_2^6 threshold V_T when $\bar{BL}=1$

\hookrightarrow Not possible to write a '1' into SRAM input (via \bar{BL})



- Make sure node B is below low threshold of M_5/M_1
inverter input



- Again as before we can set $V_B = V_{BN}$

$$M_6: I_{DM6} = \frac{1}{2} \mu_p C_{ox} \left(\frac{W}{L}\right) M_6 (V_{AS} - V_{BP})^2$$

$$= \frac{1}{2} \mu_p C_{ox} \left(\frac{W}{L}\right) M_6 (-V_{DD} - V_{BP})^2$$

$$M_4: I_{DM4} = \mu_n C_{ox} \left(\frac{W}{L}\right) M_4 [(V_{DD} - V_{BN}) V_A - \frac{V_A^2}{2}]$$

Set $I_{DM6} = I_{DM4}$ (KCL)

$$\frac{1}{2} \mu_n W M_6 (-2.5 - (-0.5))^2 = \mu_n W M_4 [(2.5 - 0.5) 0.5 - \frac{0.5^2}{2}]$$

$$\mu_p = \frac{1}{2} \mu_n \quad W_{M6} = W_{M4} \cdot \frac{7}{8}$$

or $W_{M4} = 1.15 W_{M6}$

can make this
longer to be
safe.

Concluding Comments

- Since the SRAM cell is symmetric we have $\rightarrow M_1 = M_2$
 $M_3 = M_4$

- Thus we have 2 equations w/ 3 unknowns $M_5 = M_6$

\rightarrow ① $W_{M1} = 1.3 W_{M3} \rightarrow$ set W_{M5} as that is
② $W_{M3} = 1.15 W_{M5}$ smallest quantity.

- And thus minimum cell size & speed requirements
and stability satisfied.

CLASS NOTES 25 - CIRCUIT FAMILIES

MOSFET Circuit Families

- ① Static CMOS
- ② Ratioed Circuit
- ③ Dynamic Circuit
- ④ Cascade Voltage Switch Logic (CVSL)
- ⑤ Pass transistor circuits

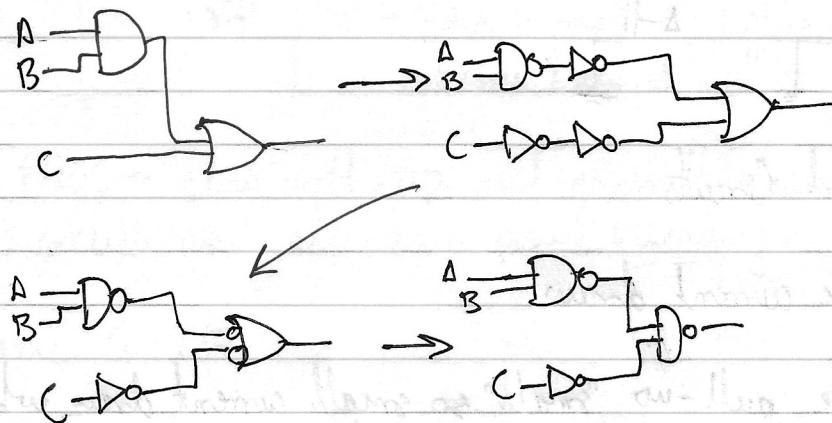
Static CMOS

- Bubble pushing (de Morgan's Law)

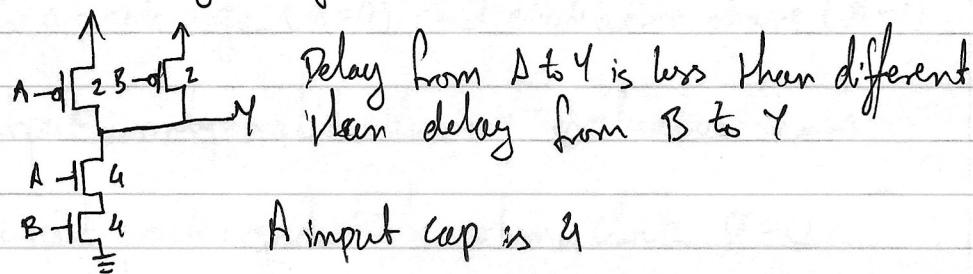
$$\overline{A} \overline{B} \rightarrow \overline{D}$$

$$\overline{D} \rightarrow \overline{D}$$

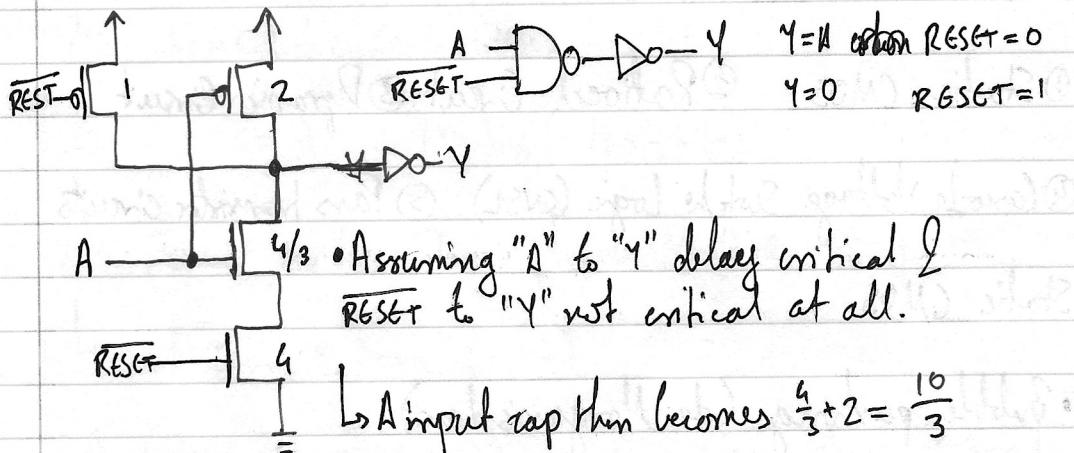
E.g. compute $F = AB + C$ using NANDs & NORs



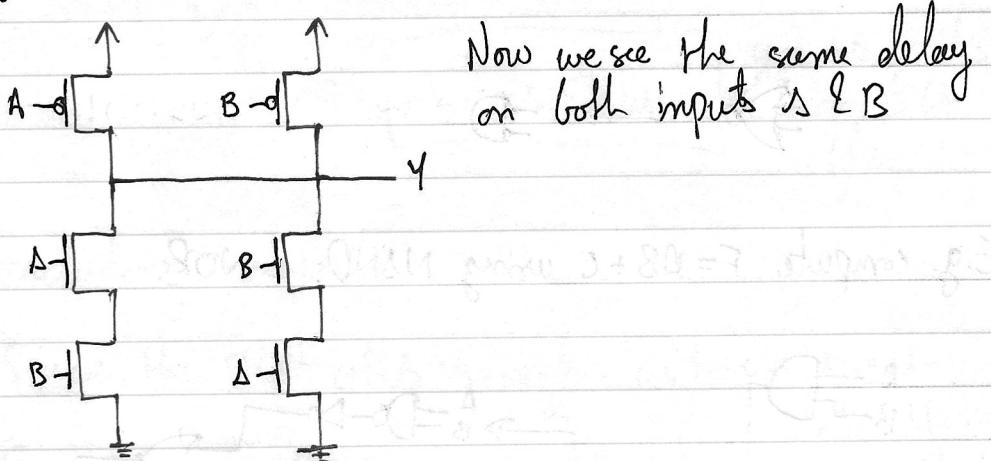
- Input ordering delay



Asymmetric Gates



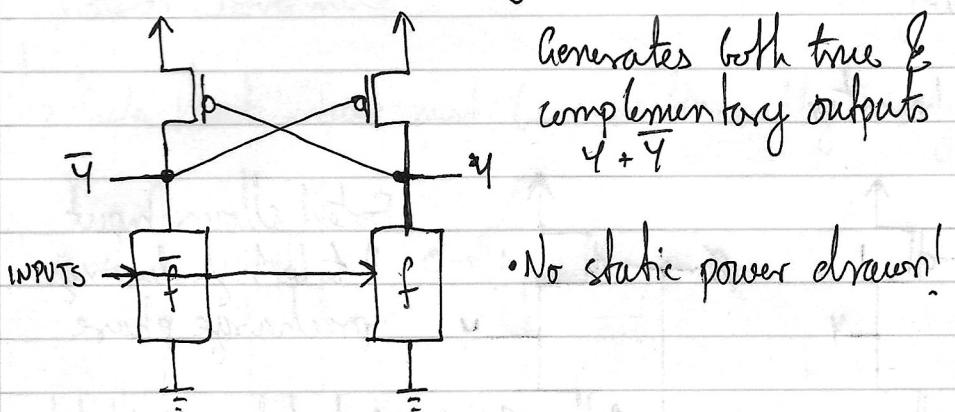
Symmetric Circuits \rightarrow NAND Example



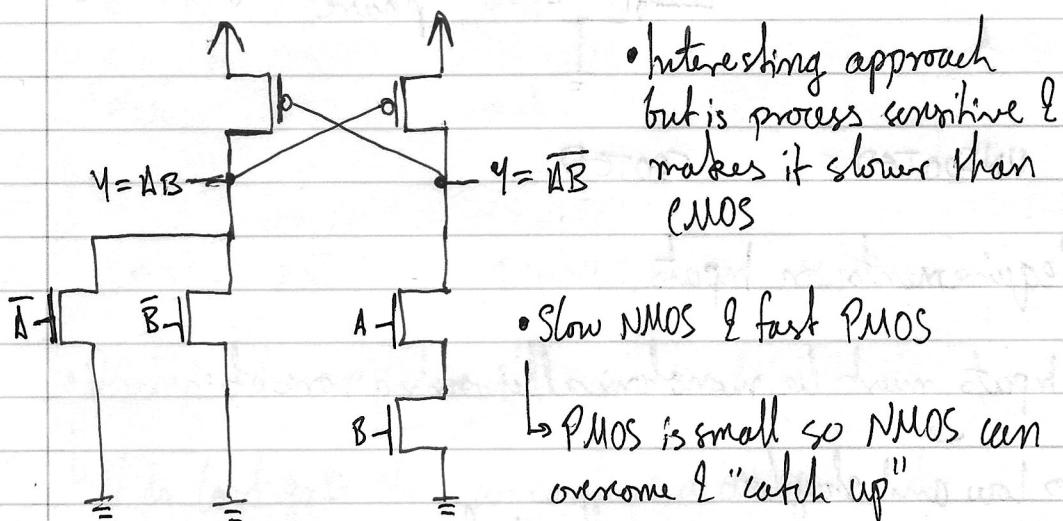
Ratioed Circuits

- Static current draw
- Make pull-up "small" so small current draw when the output is low.

Cascade Voltage Switch Logic



- NAND example



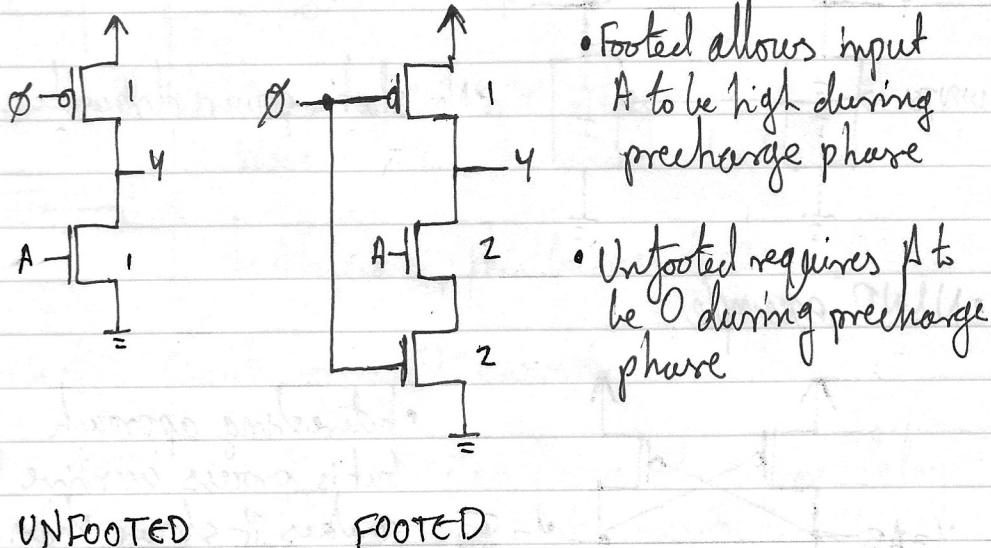
- Slow PMOS → slow pull-up and generally slower than NMOS circuits and also more power hungry.

Dynamic Circuits

- Requires a clock signal ϕ to operate
- Has a precharge ($\phi=0$) and evaluation phase ($\phi=1$)
- Output is only valid during evaluation phase
 ↳ Output precharged high when clock $\phi=0$

- Fast, no static power consumption and high dynamic power

E.g. Inverter

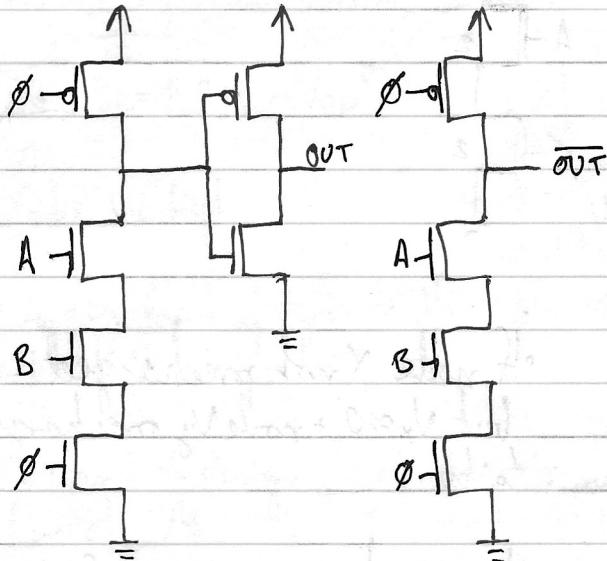


Requirements on Inputs

- Inputs must be monotonically rising, one of:
 - ↳ low and stay low
 - ↳ low and rise high allowed if
footed only.
 - ↳ High and stay high
- High and falling low is never allowed.
 - ↳ Input is precharged high and cannot be discharged as it would not be able to go high again.
- Footed allows CMOS gates to drive dynamic gates

- Cannot connect \bar{Y} directly to another gate as it violates above rule

↳ Rule that input must be monotonically rising.



DOMINO AND

DYNAMIC FOOTED NAND

- Domino gates are inherently non-inverting

↳ If its last gate then we can use a CMOS inverter

↳ If within domino gates, use dual-rail domino logic

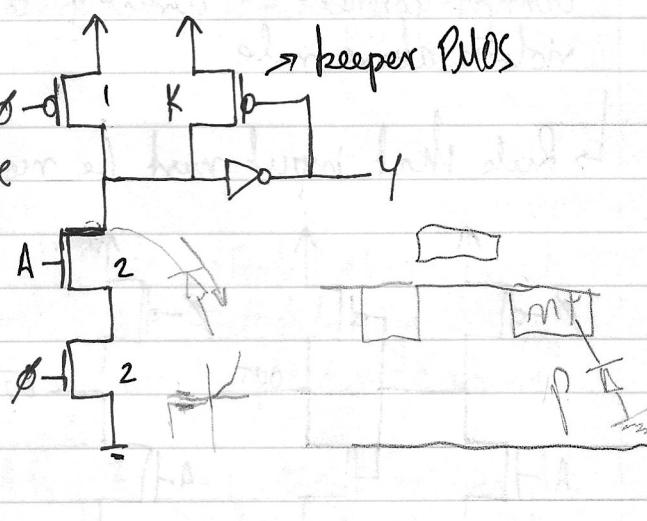
- All input/output signals encoded as a pair of signals

- ϕ needs to always run for logic to operate

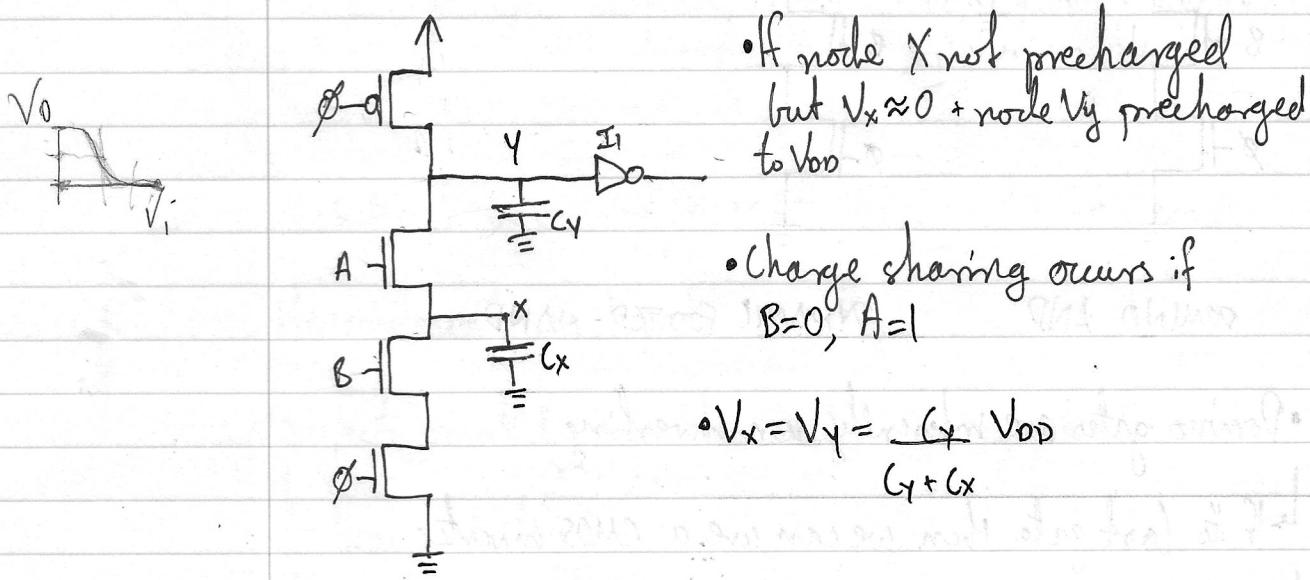
- Dynamic power high due to dual-rail, activity factor high due to precharge phase

Keepers

- If ϕ stopped on eval $\phi-d$ phase, charge leakage on dynamic node can change logic value



- Use small $\frac{V}{2}$ PMOS transistors to overcome leakage



- Might drop V_y below V_{TH} of I_1 and error occurs

↳ Might precharge all or some internal nodes to reduce this effect.

$$\underline{Q} = CV$$

LECTURE 26

Nov 10

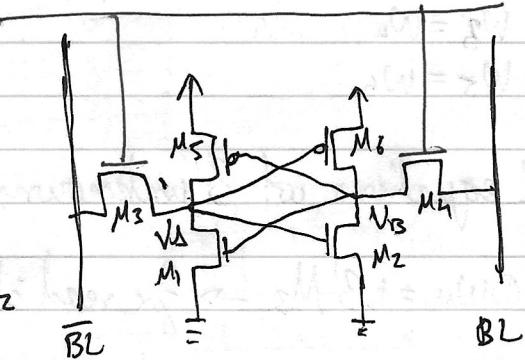
GT Write Operation

- Assume a cell has stored $V_A = V_{DD}$, $V_B = 0$ and we want to write opposite

1) Force $\bar{BL} = 0 \& BL = V_{DD}$

2) Take WL high to V_{DD}

3) V_A goes low and turns off M_2
so that V_B goes high



- We want to design (size) M_3 & M_5 such that $V_A \approx V_{Bn}$ to turn off M_2

At edge of active/triode

$$I_{DM5} = \frac{1}{2} \mu_p C_{ox} \left(\frac{W_5}{L} \right) (-V_{DD} - V_{Bp})^2 \quad \text{--- (1)}$$

$$I_{DM3} = \frac{\mu_n C_{ox}}{2} \left(\frac{W_3}{L} \right) \left[(V_{DD} - V_{Bn}) V_A - \frac{V_A^2}{2} \right] \quad \text{--- (2)}$$

$$\frac{W_3}{W_5} = \frac{\mu_p (+V_{DD} + V_{Bp})^2}{2 \mu_n [(V_{DD} - V_{Bn}) V_A - \frac{V_A^2}{2}]}$$

E.g. $\mu_n = 6 \mu_p$, $V_{DD} = 2.5V$, $|V_{Bn}| = |V_{Bp}| = 0.5V$ want $V_A = V_{Bn} = 0.5$

$$\frac{W_3}{W_5} \geq 0.57 \rightarrow \text{see if } W_3 = 0.55 \mu\text{m then } W_5 \leq 0.96 \mu\text{m}$$

- Generally we want to restrict ourselves to quantized values regardless of scaling values

Concluding Comments

- We only have 3 unknowns (not 6 due to symmetry)

$$\hookrightarrow w_1 = w_2$$

$$w_3 = w_4$$

$$w_5 = w_6$$

- 2 equations w/ 3 unknowns

$$\hookrightarrow \textcircled{1} w_{M_1} = 1.3 M_3 \rightarrow \text{for real stability} \quad \left. \begin{array}{l} M_3 \text{ has to be} \\ \text{smallest} \end{array} \right\}$$

$$\hookrightarrow \textcircled{2} w_{M_3} = 1.15 M_5 \rightarrow \text{for unstabiliy} \quad \left. \begin{array}{l} \text{pick that first.} \end{array} \right.$$

\therefore we achieve min cell size, & speed and stability are satisfied

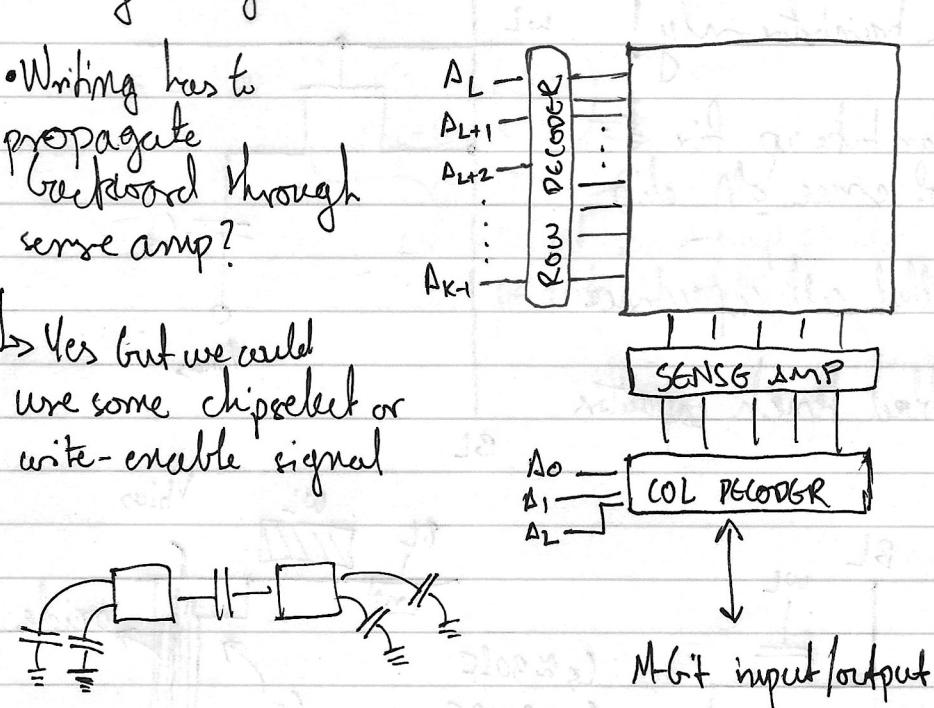
LECTURE 27

Nov 12

Memory Array

- Writing has to propagate backword through sense amp?

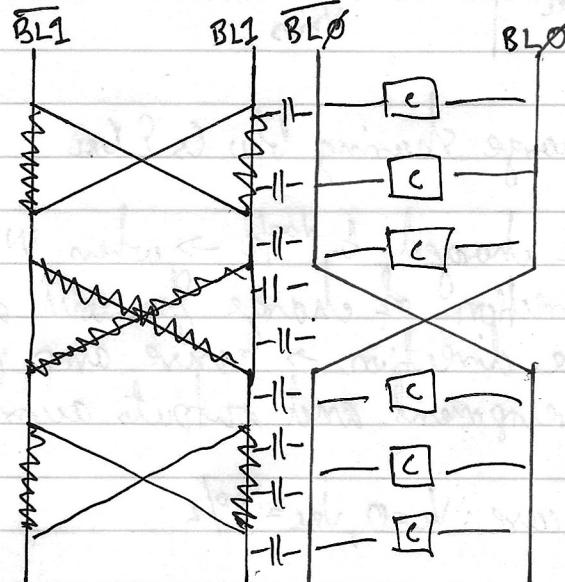
↳ Yes but we could use some chipselect or write-enable signal



Bitline Twists

- Noise on BL1 gets transferred equally to coupled BL0 and BL0'

↳ Sense amp CMRR won't pick up on crosstalk from BL1



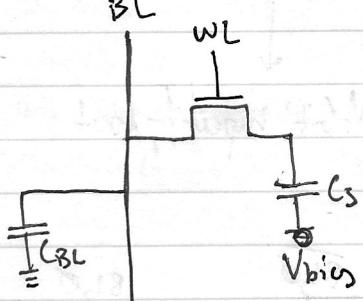
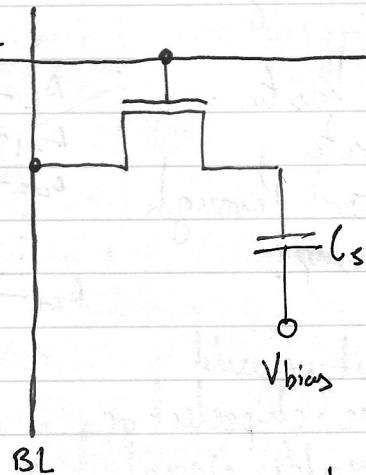
Dynamic RDN

- One transistor only!

• C_s can take up fair bit of space off chip

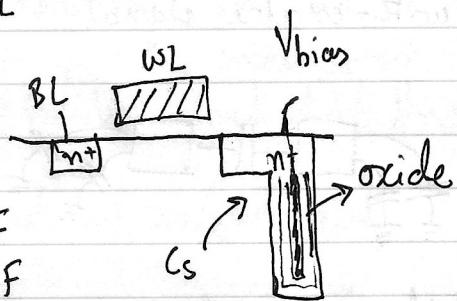
↳ called cell capacitance

↳ fabbed trench capacitors



$$C_s \approx 30\text{fF}$$

$$C_{BL} \approx 300\text{fF}$$



- Charge sharing b/w C_s & C_{BL}

• BL charged to $V_{DD}/2$ → when WL goes high, direction of charge flow will ~~not~~ bump BL in one direction → sense amp picks up signed development and outputs accordingly.

• "0" case: $V_C = 0, V_{BL} = V_{DD}/2$

$$\hookrightarrow WL=0 \rightarrow Q_i = \frac{1}{2}V_{DD}(C_{BL} + (0 - V_{bias})C_s) \quad \left. \begin{array}{l} \\ \end{array} \right\} Q_i = Q_f$$

$$\hookrightarrow WL=1 \rightarrow Q_f = V_0 C_{BL} + (V_0 - V_{bias})C_s \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{no dependence on } V_{bias}$$

$$\hookrightarrow \frac{V_{DD}}{2} C_{BL} - V_{bias} C_s = V_0 C_{BL} + V_0 C_s + V_{bias} C_s$$

$$\therefore V_0 = \frac{1}{2} V_{DD} \left(\frac{C_{BL}}{C_{BL} + C_S} \right)$$

$$= \frac{1}{2} V_{DD} - \left(\frac{C_S}{C_{BL} + C_S} \right) \left(\frac{V_{DD}}{2} \right)$$

$\Delta V \rightarrow$ what sense amp should
be able to detect.

LECTURE 28

Nov 13

1T-DRAM

Read

- "0" Case $\Rightarrow V_0 = \frac{1}{2}V_{DD} - \left(\frac{C_s}{C_s + C_{BL}}\right)\left(\frac{V_{DD}}{2}\right) = \boxed{\frac{1}{2}V_{DD} - \Delta V}$

- "1" Case $\Rightarrow V_1 = \frac{1}{2}V_{DD} - \left(\frac{C_s}{C_s + C_{BL}}\right)\left(\frac{V_{DD}}{2}\right) = \boxed{\frac{1}{2}V_{DD} + \Delta V}$

$\hookrightarrow V_{CS} = V_{DD}, V_{BL} = \frac{1}{2}V_{DD}$

$\hookrightarrow WL=0, Q_i = \frac{1}{2}V_{DD}(C_{BL} + (V_{DD} - V_{bias})C_s)$

$\hookrightarrow WL=1, Q_f = V_1 C_{BL} + (V_1 - V_{bias})C_s$

$$\Delta V = V_{DD} \left(\frac{C_s}{2(C_s + 2C_{BL})} \right)$$

• What to choose for V_{bias} ?

\hookrightarrow No effect on capacitance but affects dielectric and electric field in cap

\hookrightarrow Recall from ECE221 $\vec{E} = V/d \Rightarrow \vec{E} \propto 1/d$ $| C = \epsilon_r \epsilon_0 \frac{A}{d} \uparrow$

\hookrightarrow Want trench capacitor oxide as thin as possible to max capacitance

\hookrightarrow Thin oxide means we want minimum voltage across capacitor. Choose $V_{bias} = \frac{1}{2}V_{DD}$ so max voltage over cap is $\frac{1}{2}V_{DD}$

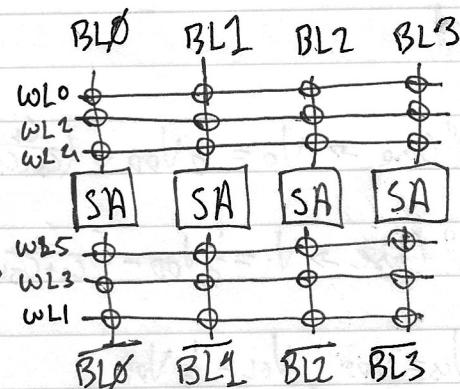


• Researchers at IBM asked about read speed and found best V_{bias} for reading is $\frac{2}{3}V_{DD}$

- DRAM Bitlines are not all differential \rightarrow issue is now we don't have CMRR from sense amp.

- Only 1 wordline goes high at any time

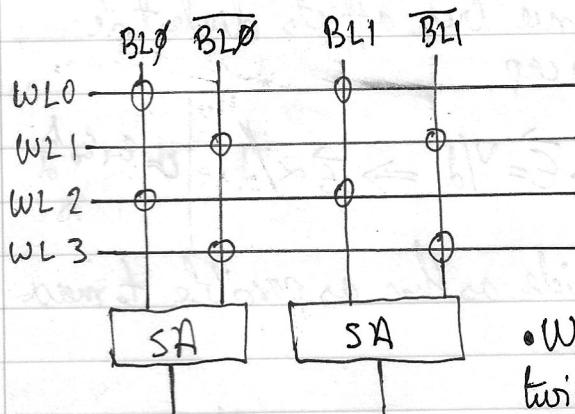
- Say WL0 then $\overline{BL_0}, \overline{BL_1}$, $\overline{BL_2}$ and $\overline{BL_3}$ are reference voltages for sense amp, however, local noise is not same on BL_2 and $\overline{BL_2}$



*Cell of each circle drawn

\hookrightarrow Fix using a folded architecture

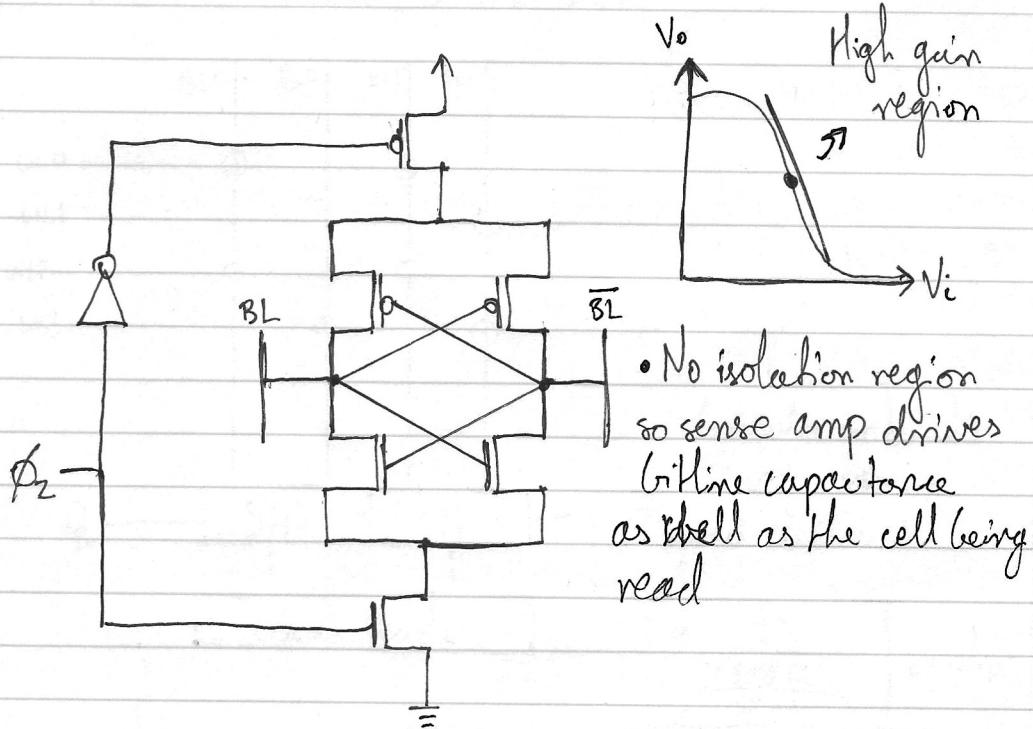
Folded Bitline Architecture



• Can't have all cells on BL_0 as capacitance would be mismatched and CMRR would be harder to compute.

• We can also have twisted wires to mitigate crosstalk imbalance.

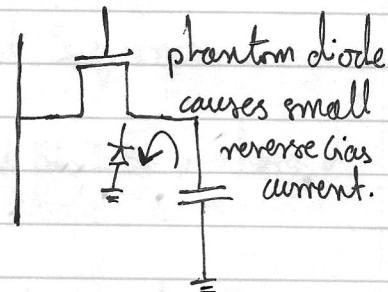
Sense Amp



- No isolation region so sense amp drives bitline capacitance as well as the cell being read

• Therefore the cell also refreshes cells on the chosen word line

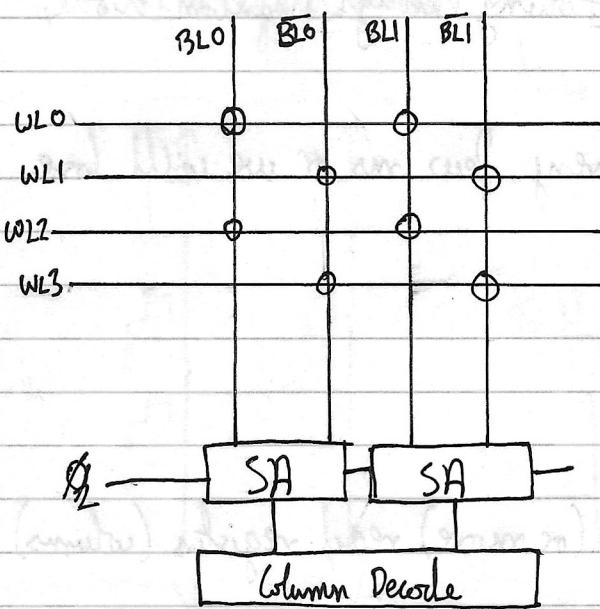
↳ Periodic refreshes must be done as cell is always slowly losing charge.



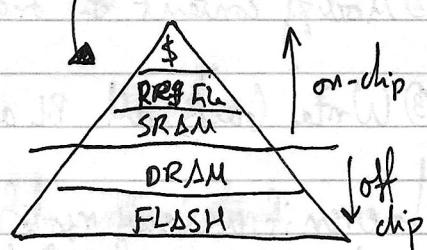
LECTURE 29

Nov 17

1T-DRAM \rightarrow Folded Bitline Arch

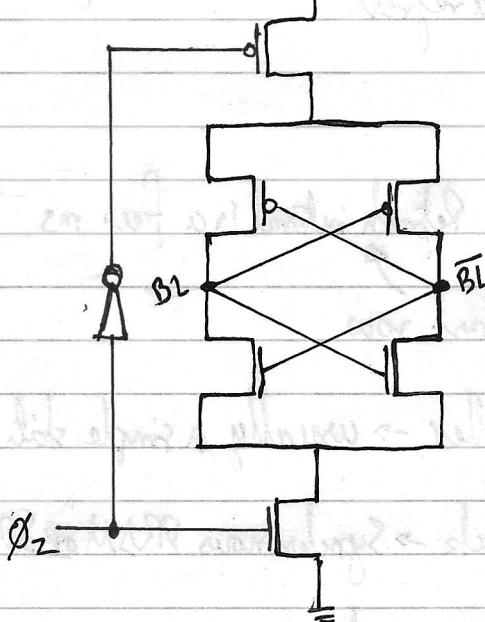


Computer architect's view on memory and electronics



Sense Amp

Read Sequence



① All BL & \bar{BL} precharged to $V_{DD}/2$

② One WL is selected, all others are low

③ V_{BL} or $V_{\bar{BL}}$ will go to $\frac{V_{DD}}{2} + \Delta V$ or $\frac{V_{DD}}{2} - \Delta V$

④ Other of V_{BL} and $V_{\bar{BL}}$ (depending on architecture) remains to $\frac{V_{DD}}{2}$

⑤ ϕ_2 goes high, driving BL and \bar{BL} to either "0" or " V_{DD} "

- Recall that reading in DRAM will refresh value stored

↳ Due to current always flowing through phantom diode, even if W_L not high

↳ We need to refresh every few ms or we will lose our data.

Write Sequence

① Read entire row

② Modify content of one (or more) read registers (columns)

③ Write back all BL and \overline{BL}

↳ even if we say modify only 32 bits, we can refresh all 1024 lines. Performs a refresh

Refresh Sequence

① Read one row at a time Refresh interval is a few ms

② Repeat when you get to last row

• Every DRAM needs a controller \rightarrow usually a simple state machine

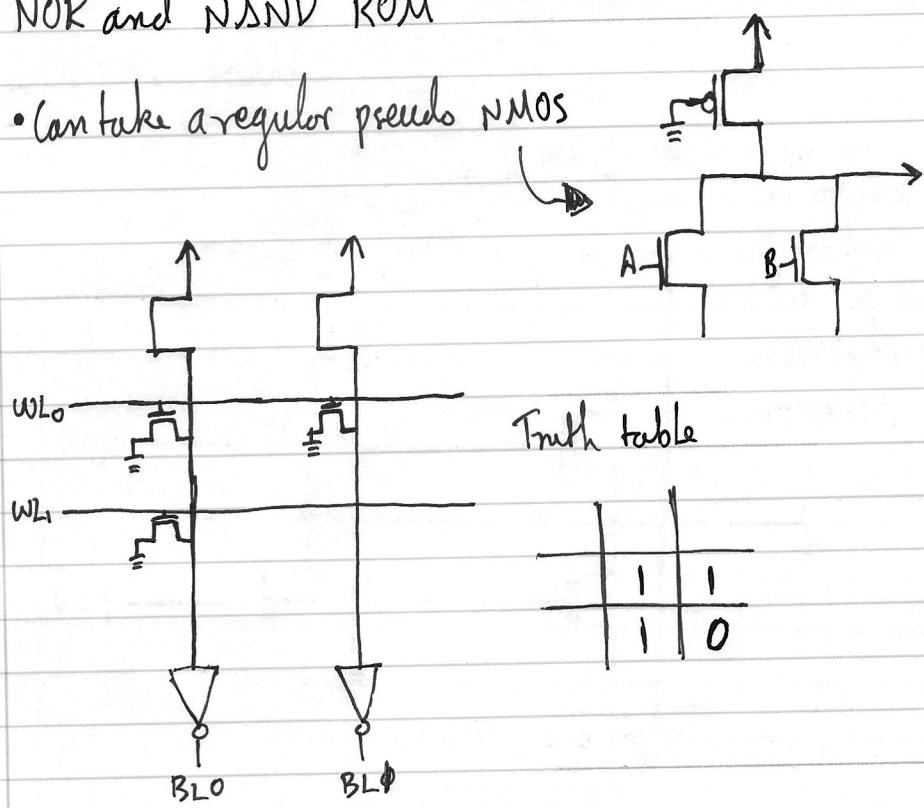
↳ If running off system clock \rightarrow Synchronous DRAM or SDRAM

• DRAM in regular ASIC process?



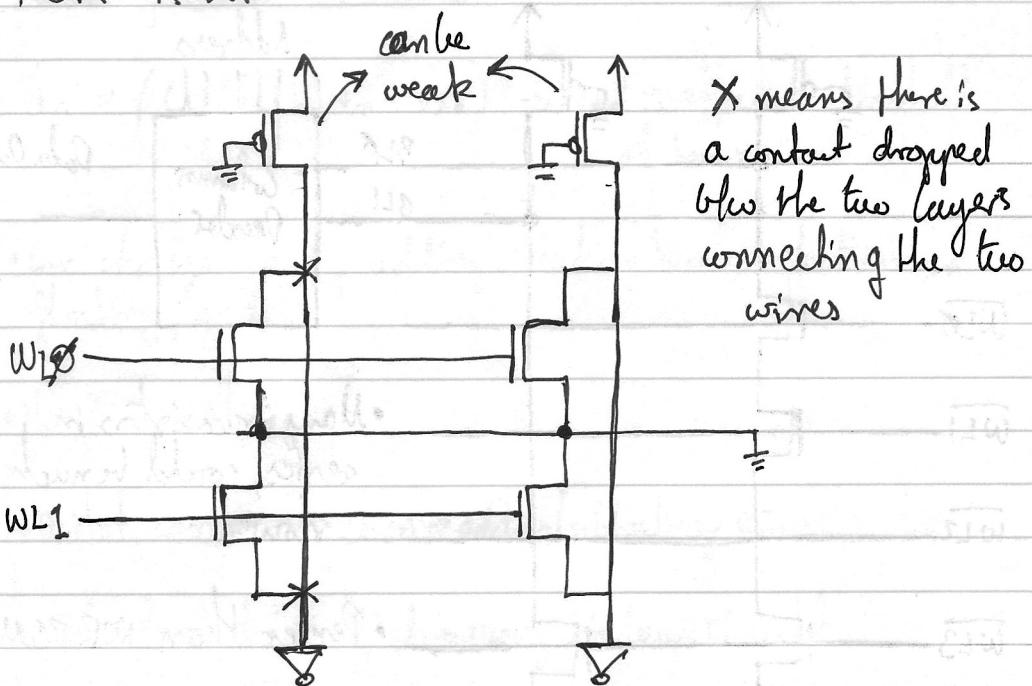
NOR and NSND ROM

- Can take a regular pseudo NMOS



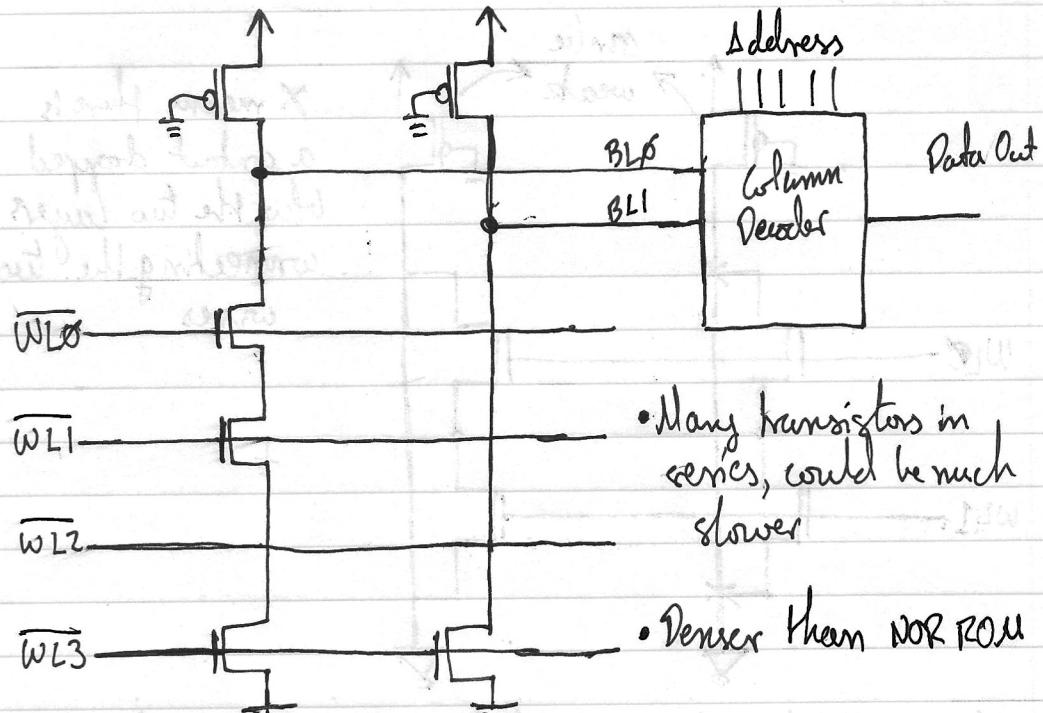
- Take this one step further in the next lecture.

NOR ROM

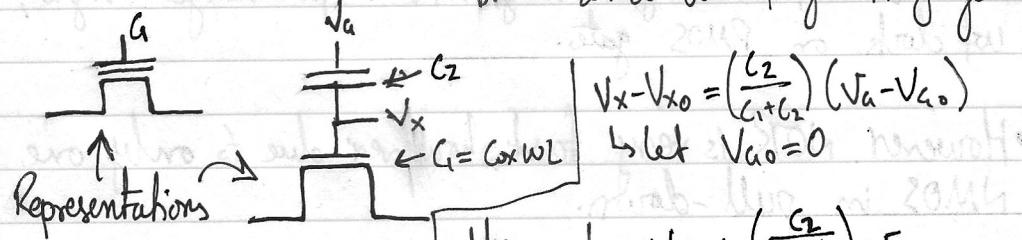
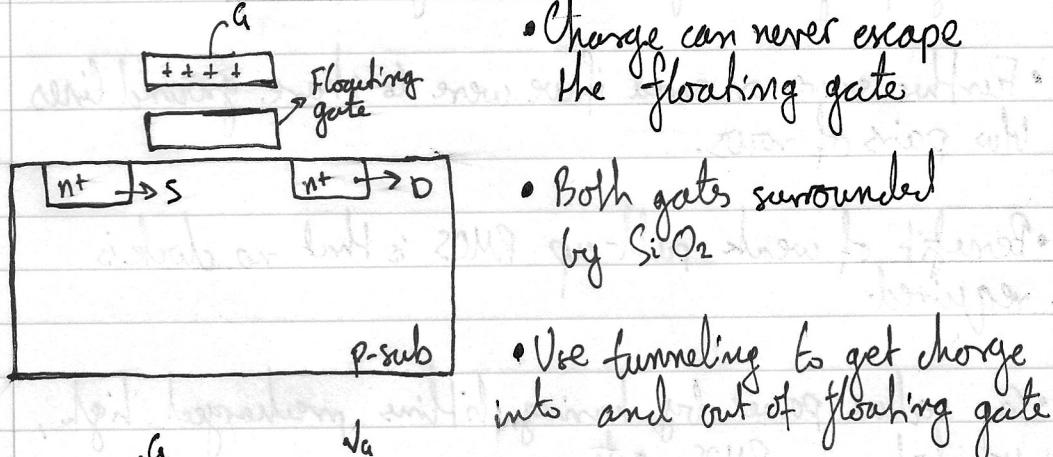


- Easy way to write ROM data → say BIOS load
- We can have half-fabricated chips, and drop metal contacts during programming → not very area efficient.
- Further save on area if we were to share ground lines b/w pairs of rows.
- Benefit of weak pull-up PMOS is that no clock is required.
- Can reduce power by having liflines precharged high, w/ clock on PMOS gate.
- However NOR is very fast, however due to only one NMOS in pull-down.

NAND ROM



Flash Memory



$$\Delta V_x = \frac{C_2}{C_1 + C_2} \Delta V_a$$

$$\hookrightarrow V_{x_0} > \frac{C_1 + C_2}{C_2} (V_{tn} - V_{x_0}) = V_{tn}' \quad \text{we set } V_{tn}'$$

LECTURE 31

Nov 20

Floating Gate Transistor

$$V_{tn'} = \frac{(C_1 + C_2)}{C_2} (V_{tn} - V_{xo}) \rightarrow V_{tn'} \text{ is effective threshold of transition}$$

- Can change $V_{tn'}$ by adding or taking away electrons on floating gate.

- How? 3 methods

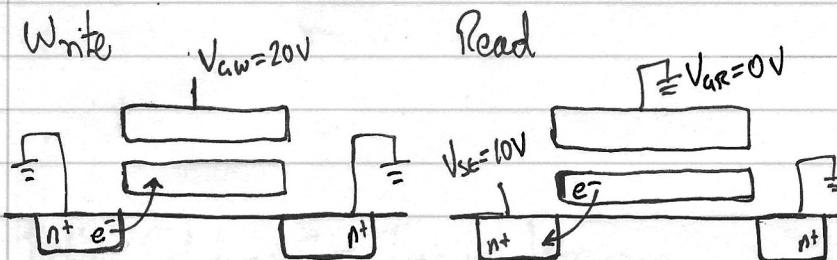
- ↳ ① Hot carrier injection \rightarrow adding/subtracting e^-

- ↳ ② Photoelectric effect \rightarrow quartz window/UV light

- ↳ ③ Fowler-Nordheim tunneling \rightarrow most common today

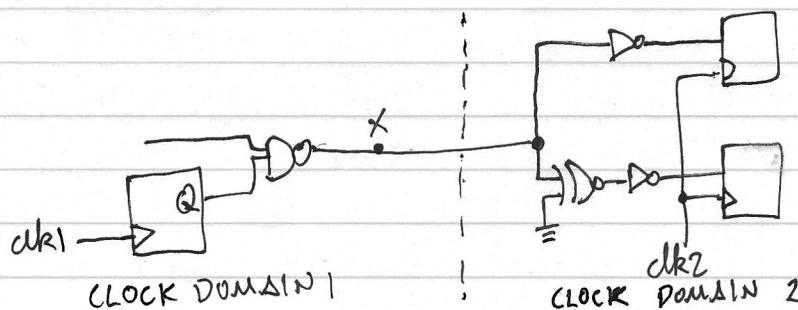
- ↳ Add/remove e^- electrons from floating gate, needs higher V_{DD}

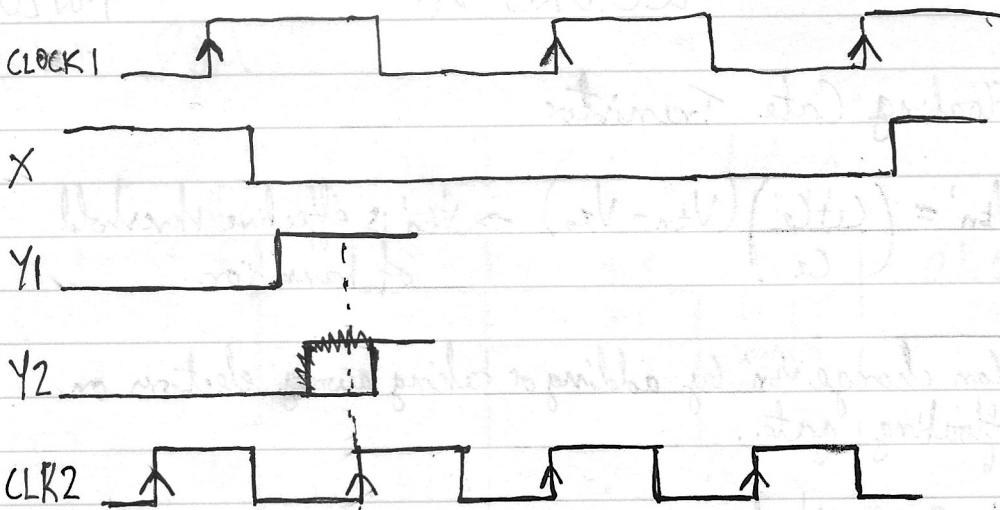
* Wear-leveling algorithms



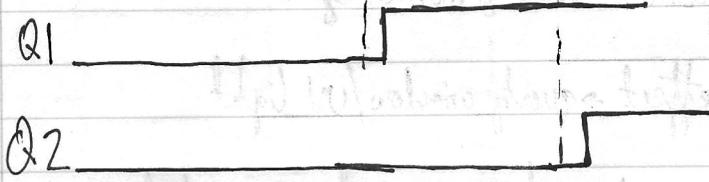
SYNCHRONIZERS AND METASTABILITY

- Consider asynchronous signal on 2 different clock domains

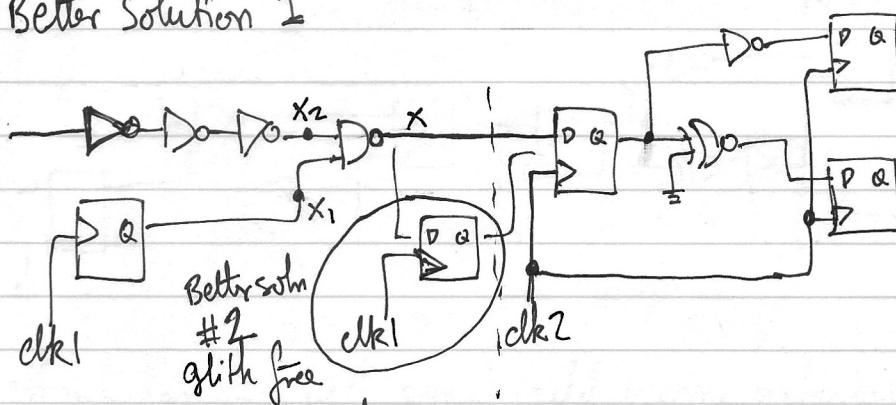




① Y1 captured but Y2 has wrong value captured until the NEXT clock edge.



Better Solution 1

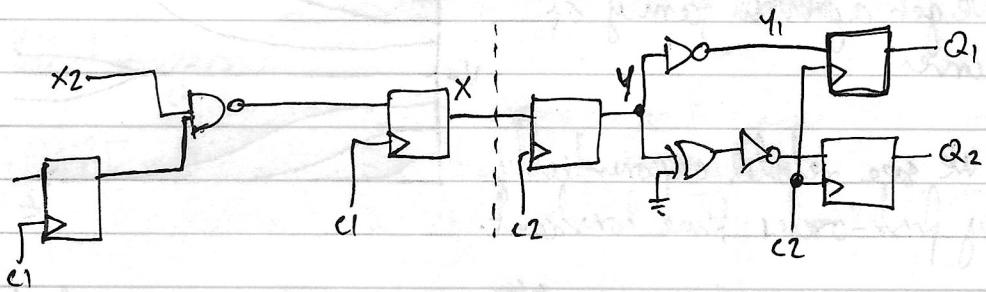


- Store x into a register BEFORE it lands in output register, having passed thru logic ~~at~~ in domain 2.

Better Solution 2 - Timing Diagram

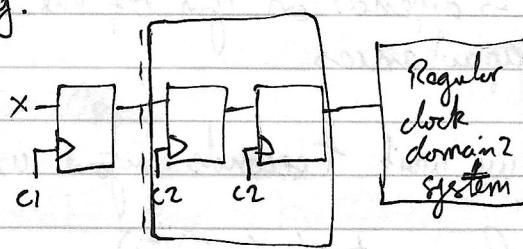
LECTURE 32

Nov 24



- If not handled well, signal could become metastable or worse, start oscillating.

- Box is a synchronizer circuit, used to reduce metastability errors

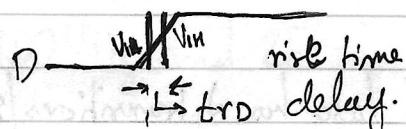


↳ More latches is better.

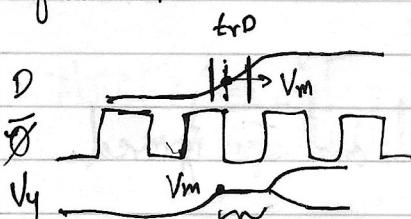
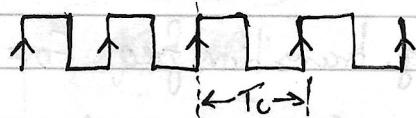
- Usually 2 is enough to reduce error to perhaps 1 in 1K years
- Sample a single data transition into an inverter w/o positive feedback

- Say $t_{RD} = 100\text{ps}$, $T_C = 10\text{ns}$

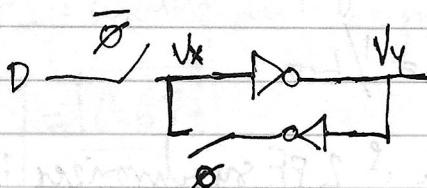
↳ $P_{error} = \frac{100\text{ps}}{10\text{ns}} = 0.01 = 1\%$.



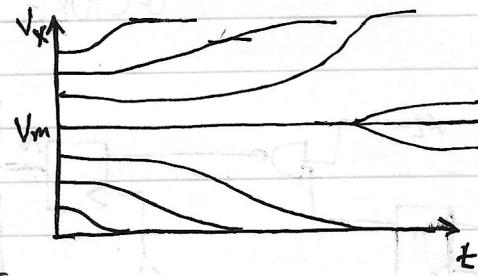
- Now we will sample a single transition on an inverter w/ positive feedback



↳ Regeneration region



- We get a family of curves



- ↳ We see related exponentially by first-order time constant.

$$V_x(t) = V_m + (V_x(0) - V_m) e^{-t/T_s} \rightarrow T_s \text{ is latch time constant}$$

- ★ ★ ★ ↳ T_s depends on g_m of the transistors and on internal capacitances

- If we wait T seconds before using signal

$$\hookrightarrow P_{\text{error}} = t_{\text{rd}}/T_c (e^{-T/T_s}) \quad P_{\text{err}} = N \frac{t_{\text{rd}}}{T_c} e^{-T/T_s}$$

- ↳ Sampled signal can be e^{T/T_s} times smaller. If N data transmissions per second (average)

- ↳ We call P_{err} the mean time b/w failures (MTBF)

$$\hookrightarrow \text{MTBF} = 1/P_{\text{err}} = \frac{T_c e^{T/T_s}}{N t_{\text{rd}}} \rightarrow t_{\text{rd}} \text{ sometimes called } T_0$$

- If data has N transitions/sec, we can define

$$\hookrightarrow \text{Avg. transition freq } F_0 = N$$

$$\hookrightarrow \text{clock freq } F_{\text{CLK}} = 1/T_c$$

$$\hookrightarrow k_1 = 1/t_{\text{rd}}$$

$$\text{MTBF} = \frac{k_1 e^{T/T_s}}{F_0 F_{\text{CLK}}}$$

- $T = T_c$, if 2 FF synchronizer is used and T_{su} ignored.

LECTURE 33

Nov 26

Synchronizers

$$\text{• MTBF} = \frac{k_1 e^{-T_c/T_s}}{F_D F_{CLK}}$$

$F_D = N$ average data frequency
 $F_{CLK} = 1/T_c$ clock frequency
 $k_1 = 1/t_{RD}$ rise time delay. \rightarrow textbook calls this T_c

- $T = T_c$ if we have 2 flip-flop synchronizers used
 ↳ and t_{SU} ignored.

Example $\rightarrow t_{RD} = 15\text{ps}$, $T_s = 20\text{ps}$

- Assume data average transition freq is $F_D = 50\text{MHz}$

- a) Find max clock rate if 2-flip synchronizers are used and want $\text{MTBF} \geq 1000$ years = $3.15e10$ seconds

$$\text{MTBF} = 3.15e10 = \frac{T_c e^{-T_c/(20e-12)}}{(50e6)(15e-12)} \rightarrow T_c = 760\text{ps}$$

} Trial & error approximation method.

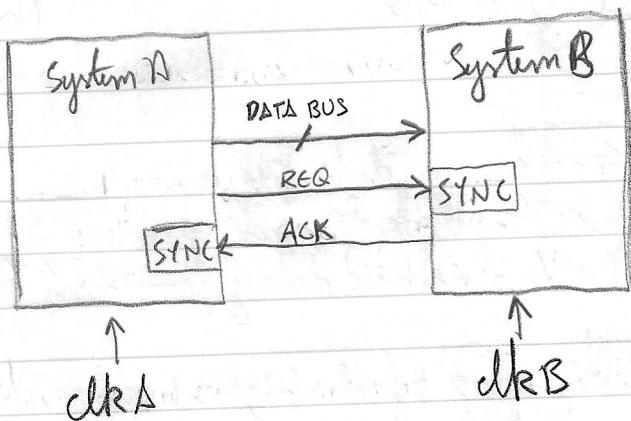
$$\therefore F_{CLK} = 1.32\text{GHz}$$

- b) Find MTBF if $F_D = 1\text{kHz}$, $F_{CLK} = 100\text{MHz}$ and 2FF synchronization used.

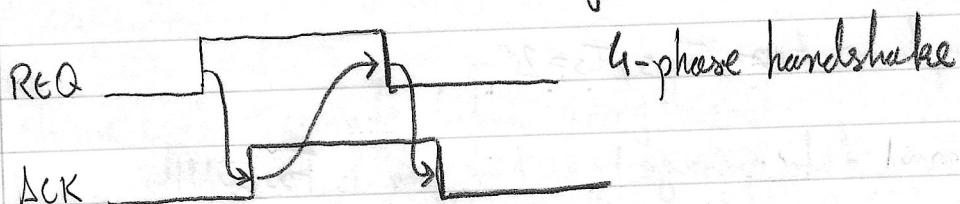
$$\text{MTBF} = \left(\frac{1}{15\text{ps}} \right) \left[\frac{e^{-100e-3/(20e-12)}}{(1e3)(10^8)} \right] = 2.97e209.5$$

- c) Find MTBF if $F_D = 1\text{kHz}$ and $F_{CLK} = 100\text{MHz}$ and no sync

$$\text{MTBF} = \left(\frac{1}{15\text{ps}} \right) \left[\frac{1}{(1e3)(10^8)} \right] = 0.67\text{s}$$



- When N-bit data stable no sync needed on input



- ① REQ high indicates data stable
- ② ACK high indicates data has been read.
- ③ REQ low indicates ACK received.
- ④ ACK low indicates REQ low received.

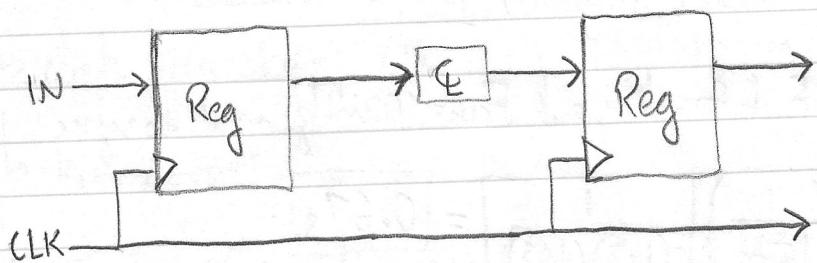
END OF FINAL EXAM COVERAGE

*START OF BONUS TOPICS

Clock Design

- Concept of ideal registers

von
Neumanns



- All registers are clocked at same time, recall

$T_c \geq t_{pq} + t_{su} + t_{pd}$ * Can think of t_{pd} as compute time

$t_h \leq t_{cd} + t_{cq}$ and t_{su} and t_{pq} as "register overhead"

- However clock skew occurs due to \rightarrow not an exhaustive list.

↳ ① Different spatial position

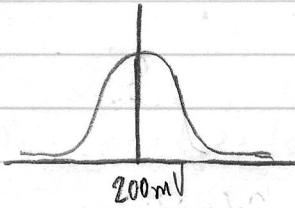
↳ ② Wire delay

↳ ③ Load delay

↳ ④ Power supplies (IR diff)

↳ ⑤ Threshold mismatch →

whole discussion
↗ on its own



$$\Delta V = L \frac{di}{dt}$$

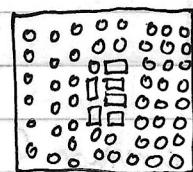
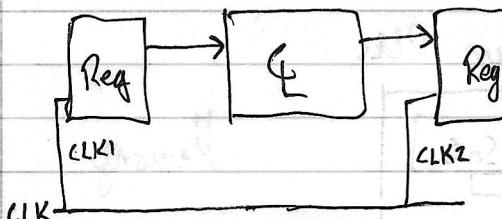
LECTURE 34

Nov 27

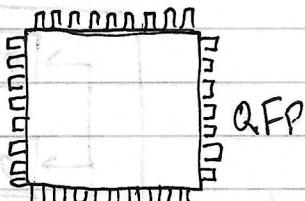
- QFP → Quad Flat Pack

- BGA → Ball Grid Array

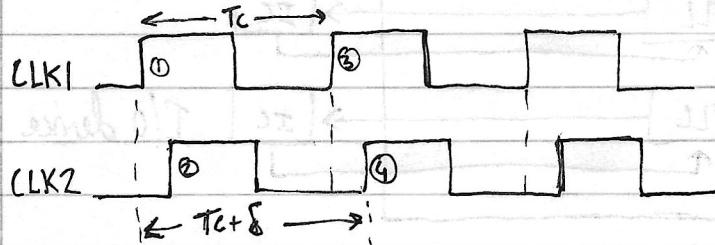
} Types of packaging



BGA



QFP



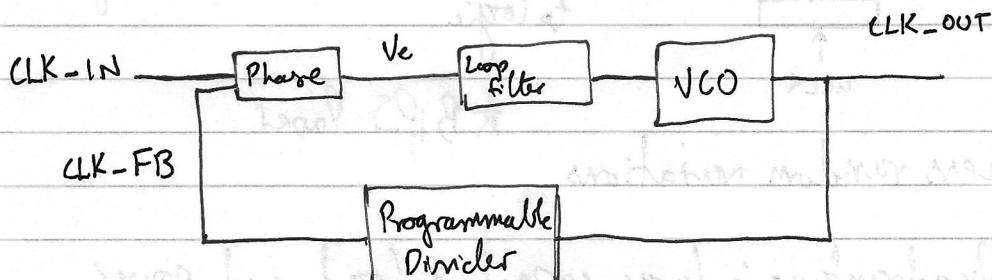
- Now $T_c + \delta \geq t_{pqg} + t_{pd} + t_{su} \rightarrow T_c \geq t_{pqg} + t_{pd} + t_{su} - \delta$

↳ Higher δ reduces value of T_c which is good!

- BUT: $t_n + \delta \leq t_{ccq} + t_d \rightarrow t_{cd} \geq t_n - t_{ccq} + \delta$

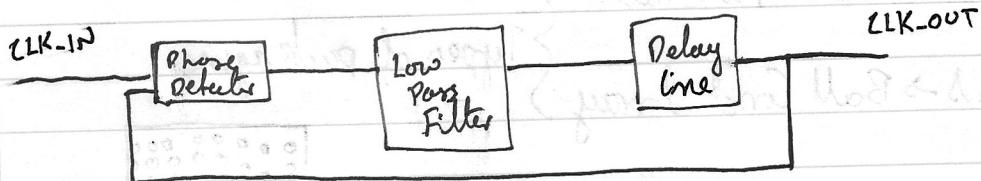
↳ t_{cd} needs to be longer otherwise race condition may occur.

Phase Lock Loop (PLL) *VCO → voltage-controlled oscillator

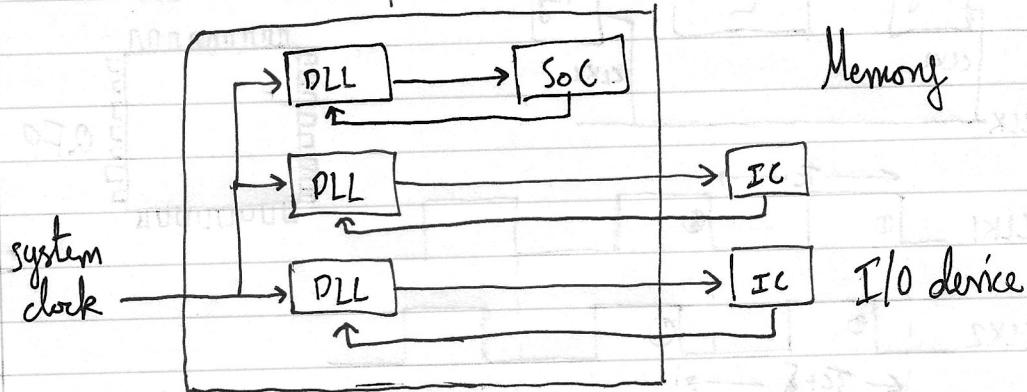


- Need to have analog/mixed-signal teams or buy license to use fab's PLL library

Delay Lock Loop (DLL)



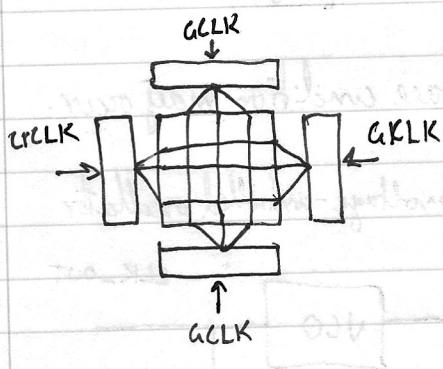
- More common & simpler than PLLs



SoC

Clock Distribution

① Grid clock



• Grid of interconnects to minimize delay variations

• Make last minute changes to logic

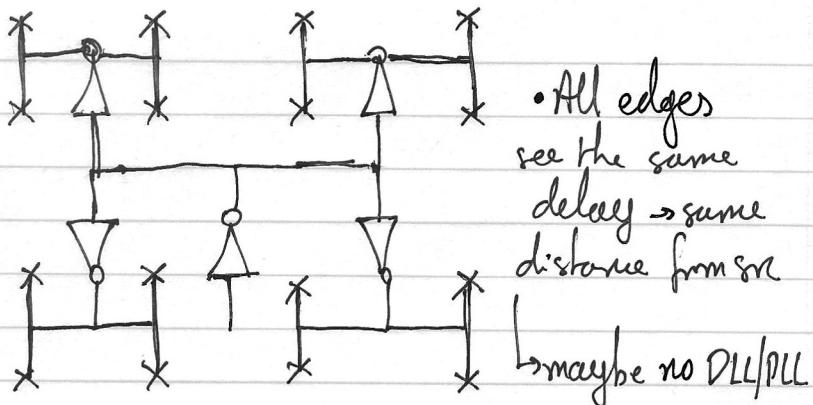
*BIPS Paper

• Less random variations

• Disadvantage is large capacitive load and power

2) H-Tree

Type of
binary tree



- All edges see the same delay \rightarrow same distance from src
- ↳ maybe no DLL/PLL