Lower-power VLSI ~~CMOS~~ Circuit Design

- why low-power design

① For battery-powered applications where battery life is an important concern.

e.g., PDA, cellular phone, all portable devices.

② High speed circuits.

power dissipation is a problem, e.g., first generation of α chip generates several tens watts of heat.

⇒ increase the cost of packaging (for power dissipation)

⇒ Reduce the life time of the chip.

③ Circuit reliability

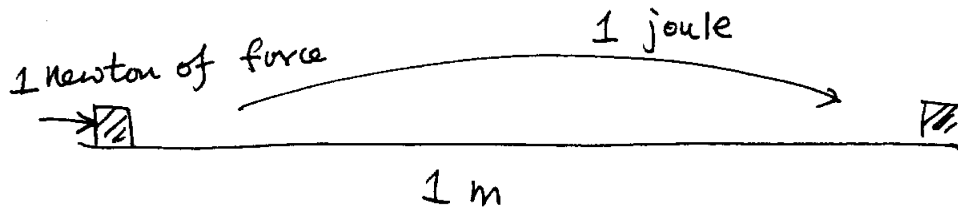
Circuit activity is a measure of the "stress" that can cause failures (run time) in digital circuits,

power ↑ circuit activity ↑ electromigration, hot carrier ↑ → reliability ↓

∴ should try to reduce power consumption for circuit reliability.

Power Modeling and Sources

- Energy: Represented by joule.



- power: The rate at which work is done.

  Represented by watts.

  $$1 \text{ watt} = \frac{1 \text{ joule}}{1 \text{ sec}}$$ ← consumed or produced

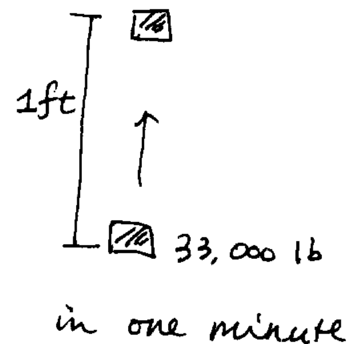
- Relationship between power & energy

$$P(\text{watts}) = \frac{W(\text{joules})}{t(\text{seconds})}$$

$$P(\text{Watts}) = I(\text{Amp}) \cdot V(\text{volt})$$

$$\mathcal{E} = QV$$

horse power:



in one minute

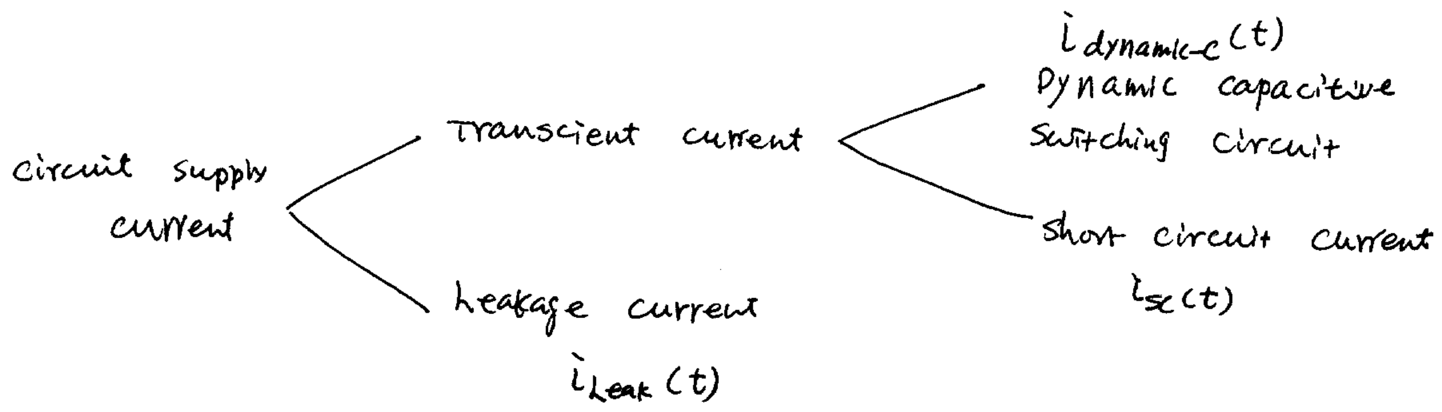
1 hp = 746 w.

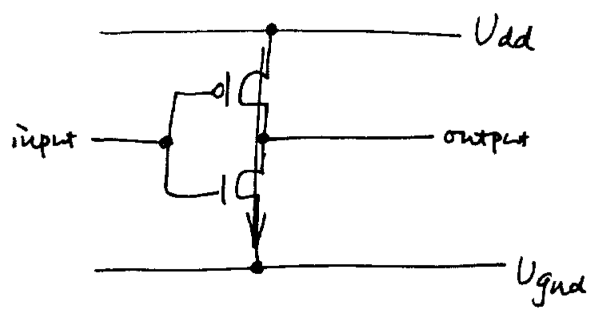
- Components of power in CMOS.

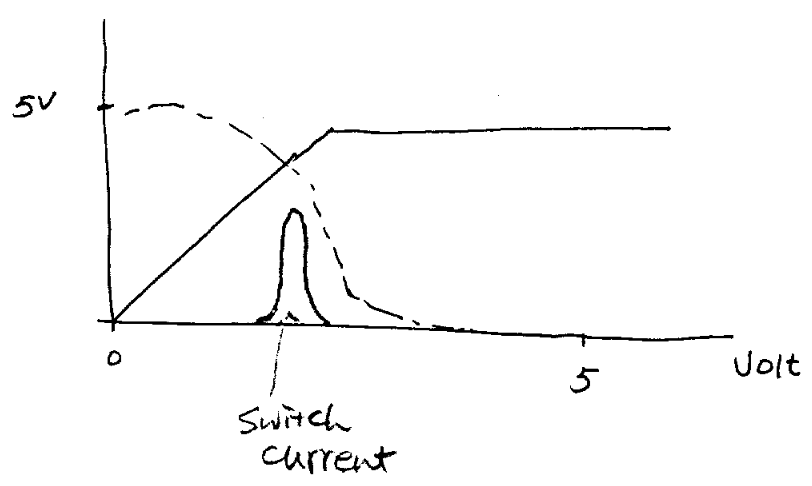
CMOS: low power device

∵ no current flow when inputs & clock are not changing

* Each signal switch ⟶ a certain amount of energy is consumed from the power supply

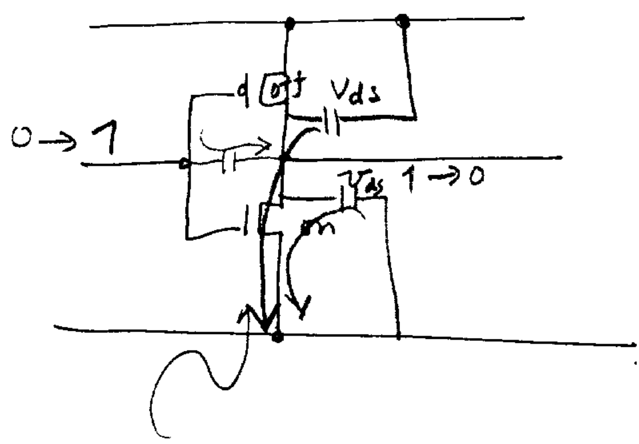






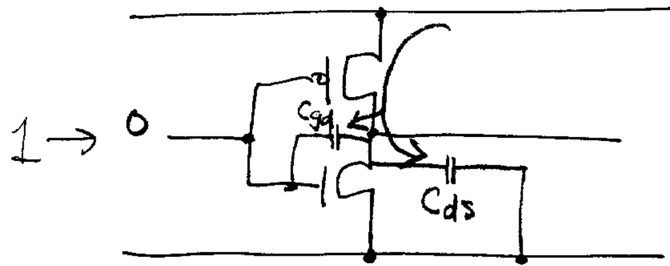
- Current induced by the simultaneous conduction of pMOS & nMOS.

This is called short-circuit current.



* Current required to charge or discharge the capacitors of the ckt.

Dynamic capacitive switching current



$$i_{DD}(t) = i_{dynamic-c}(t) + i_{sc}(t) + i_{Leak}(t)$$

- The instaneous (tan) power $p(t)$:

$$p(t) = V_{DD} \cdot i_{DD}(t)$$

maxi instantaneous power
⇒ used to estimate the ~~expected life of battery~~ circuit reliability
(electromigration, hot carriers)

- The average power P_{avg}

$$P_{avg} = \left(\frac{1}{T}\right) \int_0^T V_{DD} \cdot i_{DD}(t)\, dt$$

$$= \left(\frac{1}{T}\right) \int_0^T V_{DD} \cdot \left(i_{dynamic-c}(t) + i_{sc}(t) + i_{Leak}(t) \right) dt$$

⇒ used to estimate the expected life of battery

$$P_{dynamic-c} = \left(\frac{1}{T}\right) \int_0^T V_{DD} \cdot i_{dynamic-c}(t)\, dt$$

Called <u>dynamic power</u>, the major power consumption source.

$$P_{sc} = \left(\frac{1}{T}\right) \int_0^T V_{DD}\, i_{sc}(t)\, dt$$

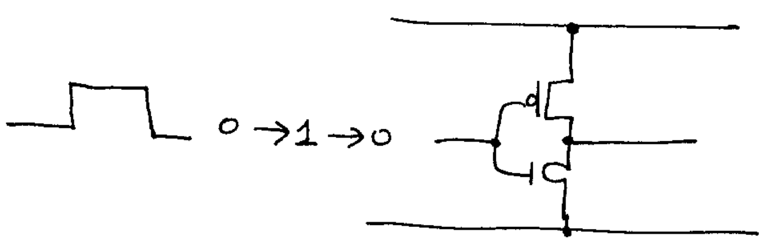
Called <u>short circuit power</u>

$$P_{Leak} = \left(\frac{1}{T}\right) \int_0^T V_{DD} \cdot i_{Leak}(t)\, dt$$

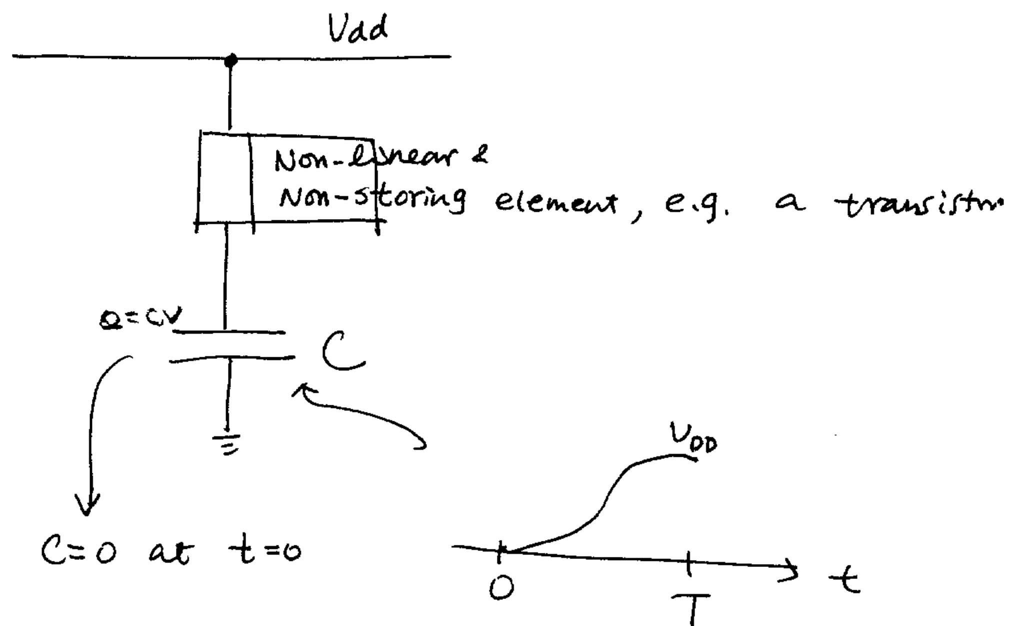
For circuit with long period inactivity ⟶ this becomes significan

- <u>Dynamic power consumption</u>

  * will analyze



* First, we analyze a simple case



- Energy supplied from Vdd.

$$\mathcal{E}_{sup}(T) = V_{DD} \cdot \underset{CV_{DD}}{\underline{Q}} = C \cdot V_{DD}^2$$

$$\int_0^{V_{dd}} V(t)\, d(v) = \frac{1}{2} v^2(t) \Big|_0^{V_{dd}}$$

- Energy stored in the capacitor

$q = cv$
$dq = c\,dv$

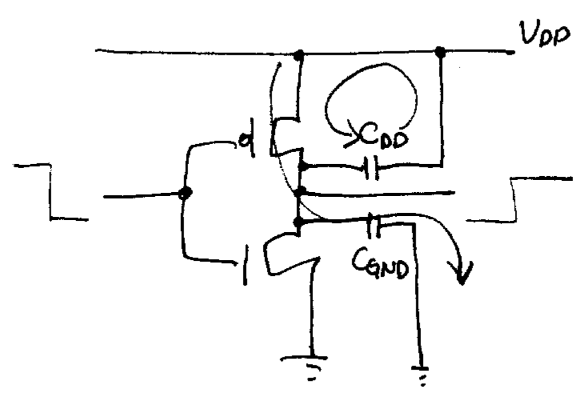
$$\mathcal{E}_{stored}(T) = \int_0^{Q} v(t)\, dq = \int_0^{V_{DD}} V(t)\, C\, dv = \frac{1}{2} C V_{DD}^2$$

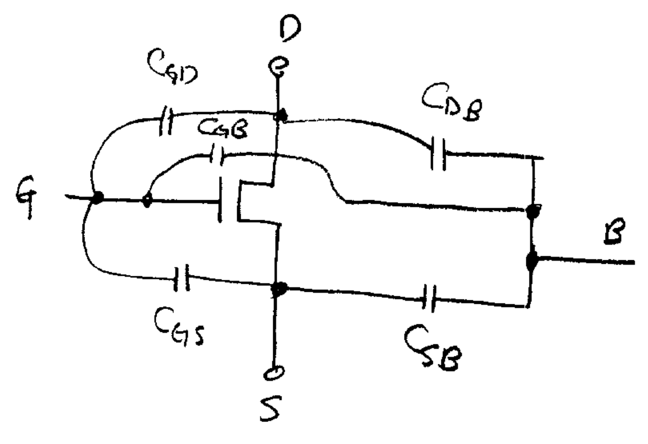
∴ The rest of energy is dissipated at the non-linear non-storing device as "heat"

$$\mathcal{E}_{dis}(T) = C V_{DD}^2 - \frac{1}{2} C V_{DD}^2 = \frac{1}{2} C V_{DD}^2$$

∴ 50% of energy has been lost in "heat"

Load capacitance Model





$v_B = 0$ for nMOS

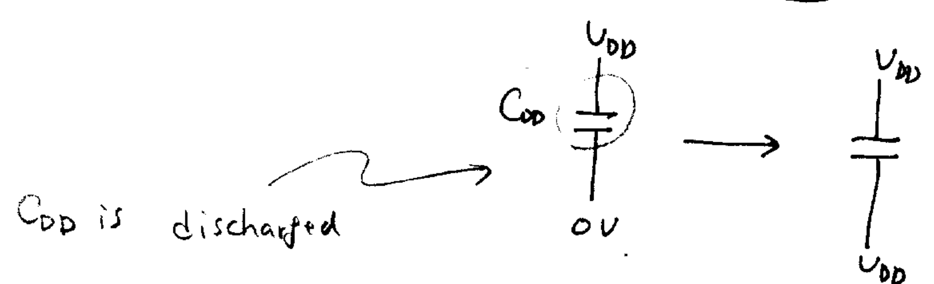
$V_B = 1$ for pMOS

input: ~~0→1~~ 1 → 0

Energy supplied: $C_{GND} V_{DD}^2$

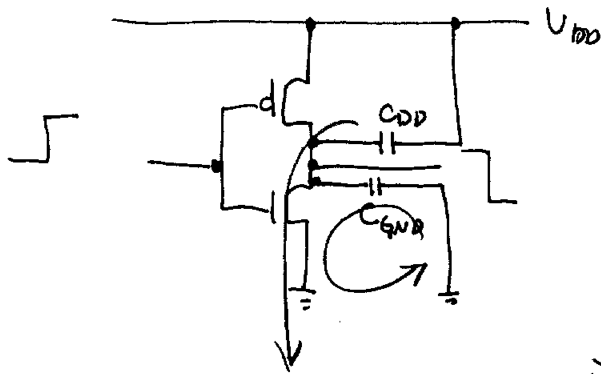
Energy stored in C_{GND}: $\frac{1}{2} C_{GND} V_{DD}^2$

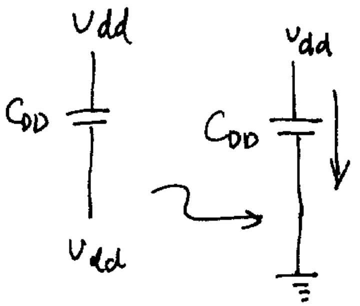
Energy dissipated: $\frac{1}{2} C_{GND} V_{DD}^2 + \frac{1}{2} C_{DD} V_{DD}^2$



C_{DD} is discharged from V_{DD} to 0.

$\therefore \frac{1}{2} C_{DD} V_{DD}^2$ energy saved has been lost!!





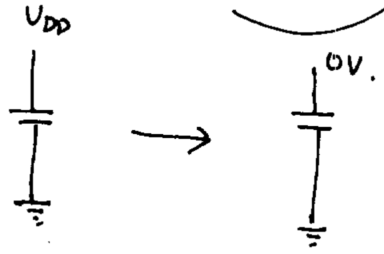
input $0 \to 1$

Energy supplied: $C_{DD} V_{DD}^2$

Energy stored: $\frac{1}{2} C_{DD} V_{DD}^2$

$\updownarrow$

Energy dissipated: $\frac{1}{2} C_D V_{DD}^2 + \frac{1}{2} C_{GND} V_{DD}^2$



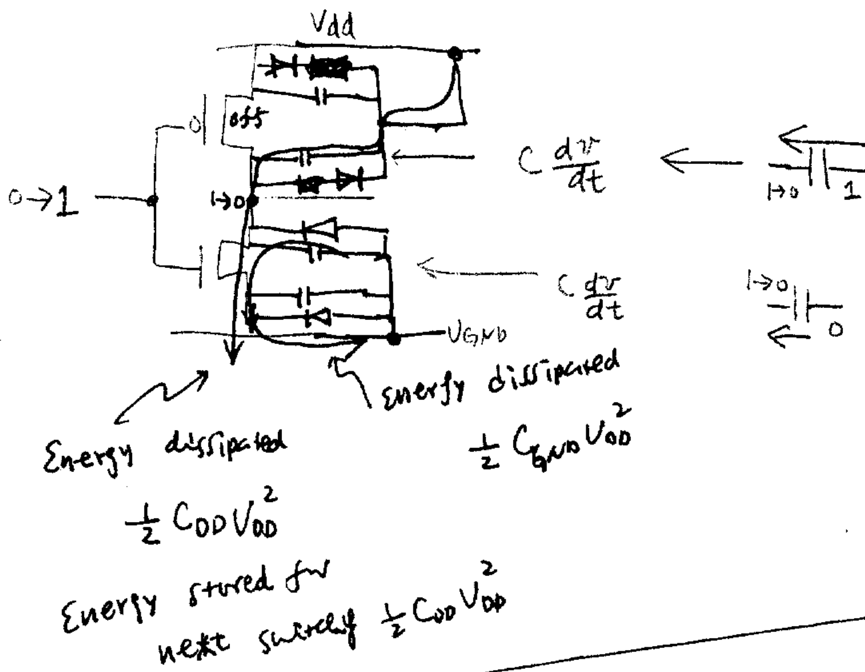
For a Full cycle

$0 \to 1 \to 0$

Energy supplied: $(C_{GND} + C_{DD}) V_{DD}^2$

Energy Stored: 0

Energy dissipated: $(C_{GND} + C_{DD}) V_{DD}^2$



V_{DD}

Energy dissipated

$\frac{1}{2}C_{DD}V_{DD}^2$

on

C_{DD}

$i = C\frac{dv}{dt}$

1→0

0→1

0→1 1 Energy stored discharged

0→1 0 Energy stored

off

C_{GND}

Energy stored for next switching

$\frac{1}{2}V_{GND}^2 C_{GND}$

V_{GND}

Energy dissipated

$\frac{1}{2}C_{GND}V_{DD}^2$

	Input node transitions		
	1→0	0→1	1→0→1
Energy supplied	$C_{GND} V_{DD}^2$	$C_{DD} V_{DD}^2$	$(C_{GND}+C_{DD}) V_{DD}^2$
Energy stored	$\frac{1}{2} C_{GND} V_{DD}^2$	$\frac{1}{2} C_{DD} V_{DD}^2$	0
Energy dissipated	$\frac{1}{2} C_{GND} V_{DD}^2 + \frac{1}{2} C_{DD} V_{DD}^2$	$\frac{1}{2} C_{DD} V_{DD}^2 + \frac{1}{2} C_{GND} V_{DD}^2$	$(C_{GND}+C_{DD}) V_{DD}^2$

borrowed from previous state

consumed by next state

∴ net Energy stored is φ

- **Conclusion:**

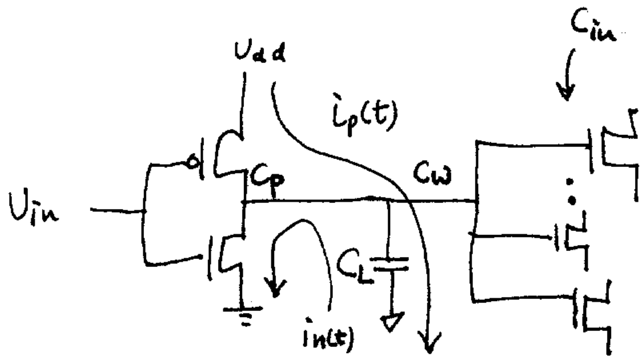
The average dynamic power of a logic node with load capacitance C_L with a complete cycle $0 \rightarrow 1 \rightarrow 0$ (output) transition during time interval T is

$$P_{avg\text{-}dynamic\,C} = E_{sup\text{-}cycle} / T = C_L V_{DD}^2 / T$$

for one period T

includes C_{GND}, C_{DD} & C_{load}.

Dynamic Output Load



$C_L = C_{in} + C_w + C_p$

(Details will be given later)

- C_L is much larger than the parasitic capacitances of the inverter

- $V_{in}(t)$

T

$V_o(t)$

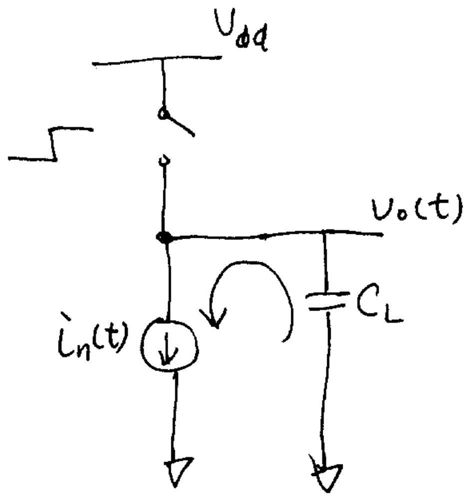
t

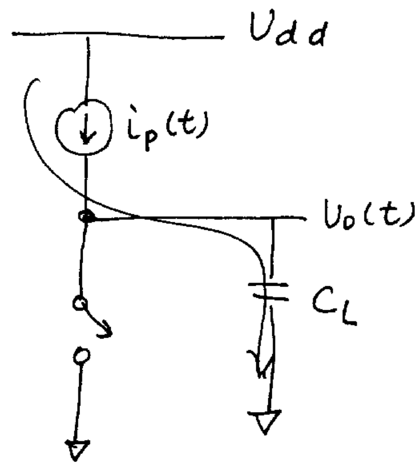
$i(t)$

$i_n(t)$

$i_p(t)$

t





$$P_d = \frac{1}{T}\int_0^T i_o(t)\, v_o(t)\, dt$$

P_d: The average dynamic power required to charge & discharge C_L at frequency $f = \frac{1}{T}$

$\because$ $i_o(t) = i_p(t) = C_L \cdot \frac{dv_o}{dt}$ ← during charging phase

$i_o(t) = i_n(t) = -C_L \frac{dv_o}{dt}$ ← during discharge phase

$$P_d = \frac{1}{T}\left[\underbrace{\int_0^{V_{DD}} C_L v_o\, dv_o}_{\text{charge}} - \underbrace{\int_{V_{dd}}^{0} C_L v_o\, dv_o}_{\text{discharge}}\right]$$

$$= \frac{C_L V_{dd}^2}{T} = C_L V_{dd}^2 f.$$

- Charge phase : output $0 \longrightarrow V_{dd}$

  Energy drawn from power supply

  $$E_d = C_L V_{dd}^2$$

Energy stored:

$$\mathcal{E}_{cap} = \int_0^{V_{dd}} C_L V_0 \, dv_0 = \frac{1}{2} C_L V_{dd}^2$$

Another $\frac{1}{2} C_L V_{dd}^2$ consumed by PMOS transistor

- Discharge phase: output $V_{dd} \longrightarrow 0$

  No Energy drawn from V_{dd}.

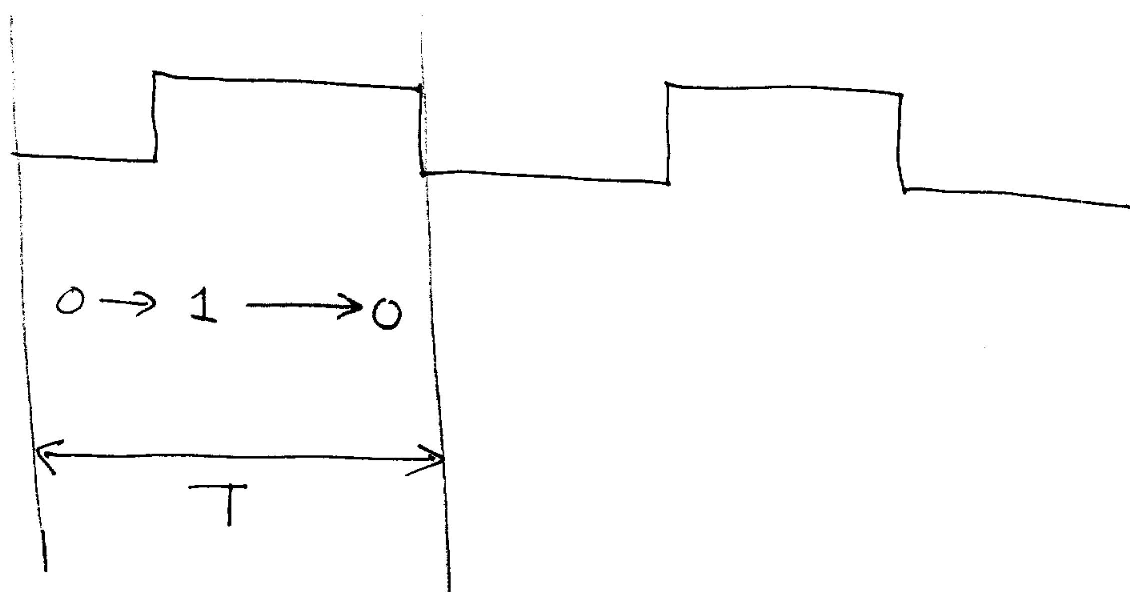
  The energy $\frac{1}{2} C_L V_{dd}^2$ stored in C_L is consumed by pull-down NMOS transistor

∴ ~~power~~ Energy supplied by V_{dd}: $C_L V_{dd}^2$

~~power~~ Energy consumed
- by PMOS: $\frac{1}{2} C_L V_{dd}^2$
- by NMOS: $\frac{1}{2} C_L V_{dd}^2$

∴ ~~power~~ Energy supplied = ~~power~~ Energy consumed.

∴ ~~Energy can't be c~~ Energy Conservation



$\alpha_{0\to1} = P(\text{output has a } 0\to1 \text{ transition})$

node activity

$$P_{ave\text{-}dynamic} = \left(\alpha_{0\to1}\cdot C_L\cdot V_{DD}^2\right)/T$$

for overall period

$$= \alpha_{0\to1}\cdot C_L\cdot V_{DD}^2\cdot f_{clk}.$$

Definition: $E(sw)$: Expected number of node switching $0\to1$ or $1\to0$ during T.

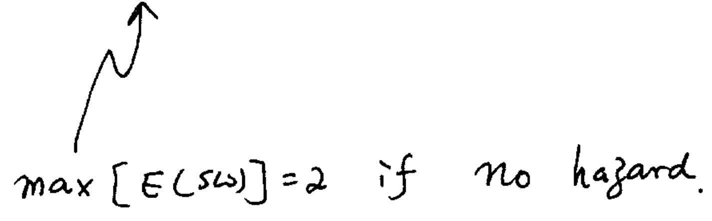
- Node switching ~~E(sw)~~ of $0\to1\to0$ is 2
- ~~the # of complete cycle is 2~~

$\therefore$ ~~the expected # of complete cycle~~ $\alpha_{0\to1} = \frac{1}{2}E(sw)$.

if $P(0\to1) = P(1\to0)$

$$P(0\to1) = \tfrac{1}{2}\left[P(0\to1) + P(1\to0)\right]$$
$$= \tfrac{1}{2}\left[P(0\to1)\cdot1 + P(1\to0)\cdot1\right]$$
$$= \tfrac{1}{2}E(sw)$$

$\therefore$ $P_{ave\text{-}dynamic} = \frac{1}{2} E(sw) C_L V_{DD}^2 f_{clk}$.



$\therefore$ Key factors influence the dynamic power:

$E(sw)$: linear at each node

C_L : linear

V_{DD} : quadratic $\therefore$ trying to reduce the voltage is the most effective way for low-power design

f_{clk} : linear.

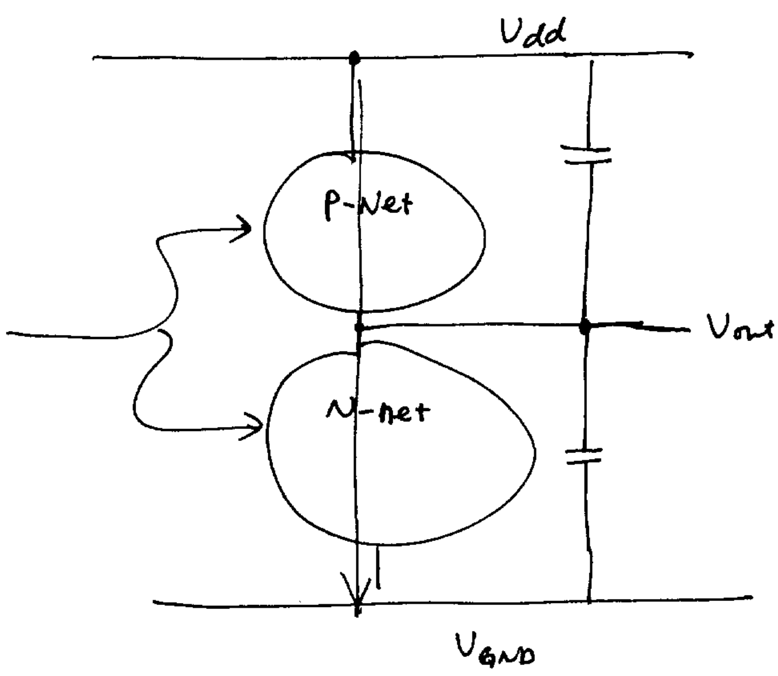
- $P_{ave\text{-}dynamic} = \left(\frac{1}{2} E(sw) C_L\right) V_{DD}^2 f_{clk}$

$= \underline{C_{eff}} \cdot V_{DD}^2 f_{clk}$

Called effective capacitance.

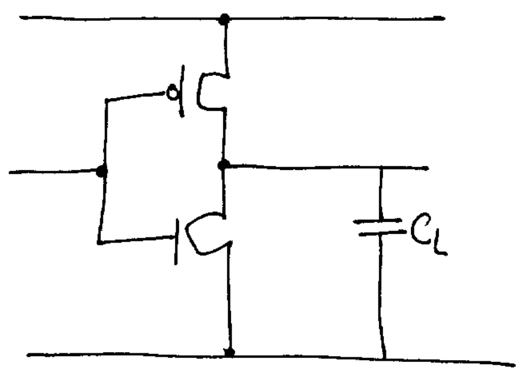
- Short circuit power: Not useful at all.
  produces heat in the conducting path.

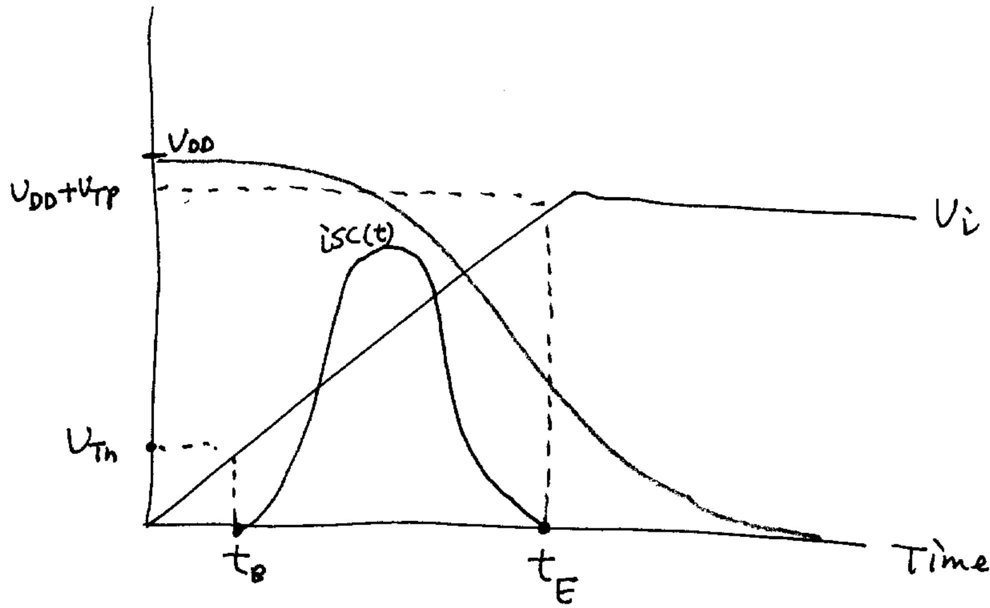
- short circuit current



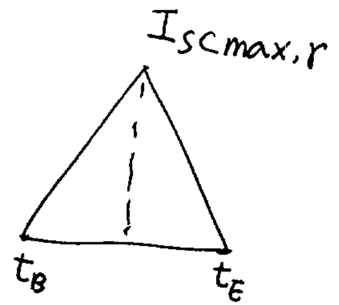
The current flowing from Vdd to V_{GND} through the the pMOS & nMOS transistor nets for input transitions.

- Show by Inverter





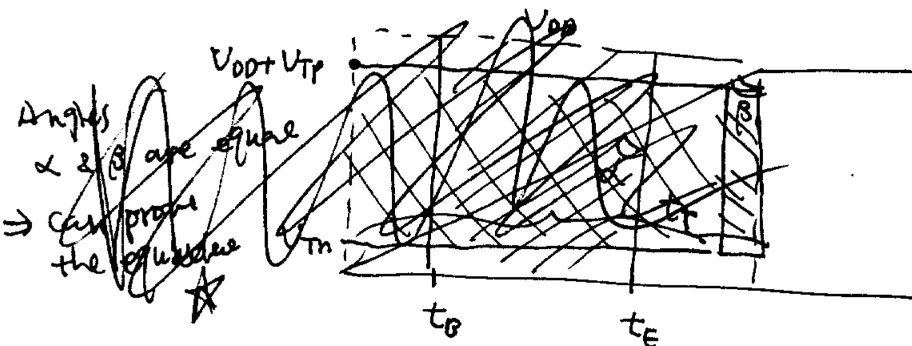
· Rising short circuit Energy

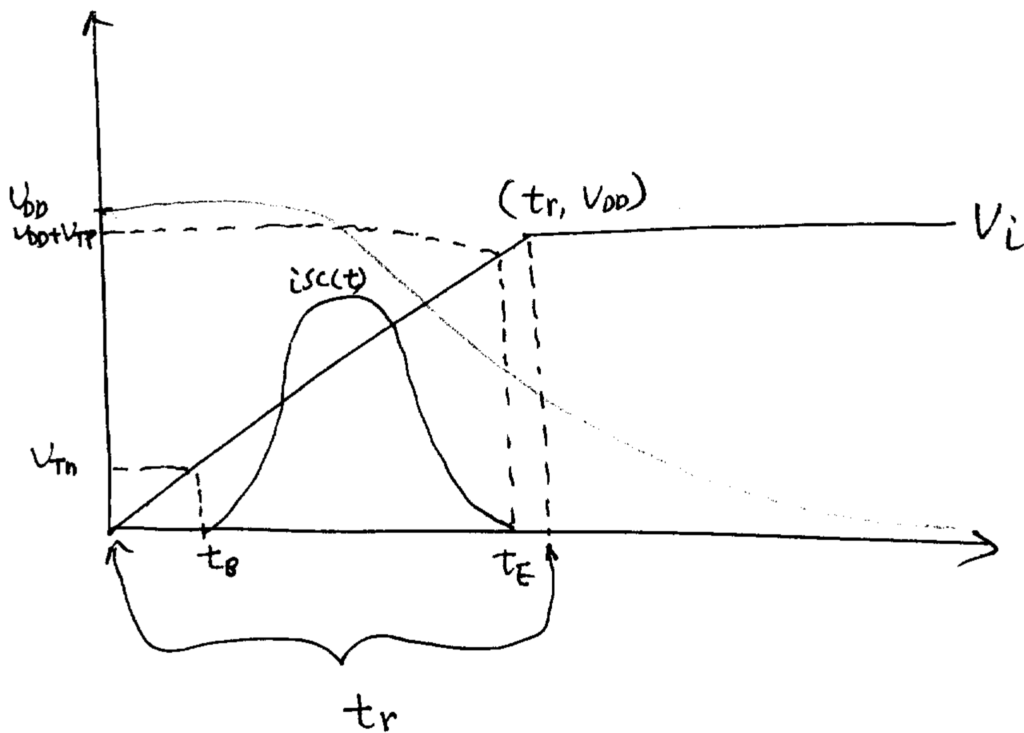


$$E_{scr} = \int_{t_B}^{t_E} V_{DD} \cdot i_{sc}(t)\, dt \approx \left[\frac{(t_E - t_B)}{2} I_{scmaxr} \right] \cdot V_{DD}$$

$$= \frac{V_{DD}}{2} (t_E - t_B) I_{scmaxr} = \frac{1}{2} t_r (V_{DD} + V_{Tp} - V_{Tn}) \cdot I_{scm}$$

t_r: time required for V_{in} to rise from 0 to V_{DD}





$$\frac{V_{DD}}{t_r} = \frac{V_{DD}+V_{TP}}{t_E} = \frac{V_{Tn}}{t_B}$$

$$\therefore \quad \frac{V_{DD}}{t_r} = \frac{V_{DD}+V_{TP}-V_{Tn}}{t_E - t_B}$$

Exmple: $\frac{4}{3} = \frac{8}{6} = \frac{20}{15}$

$$\frac{4}{3} = \frac{12}{9}$$

$$\Rightarrow V_{DD}(t_E - t_B) = (V_{DD}+V_{TP}-V_{Tn})\, t_r$$

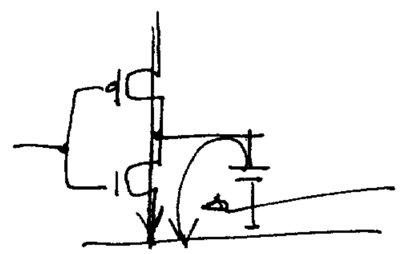
I_{scmaxr} depends on

① The size of the transistors

Larger size ↑ $\quad$ I_{scmaxr} ↑

② Load capacitor C_L:

$$C_L \uparrow \Rightarrow I_{scmaxr} \downarrow$$



See next page

larger C_L will induce larger dynamic capacitive current, reduce the current flow for I_{scmaxr}.

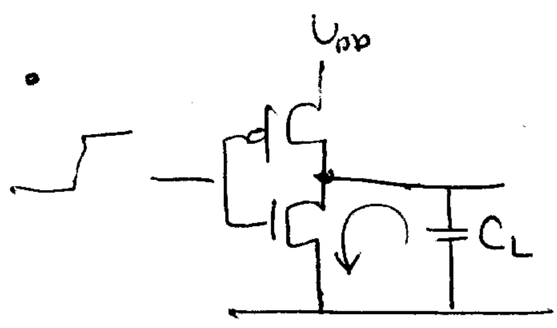
∴ I_{scmaxr} is max when $C_L = 0$.

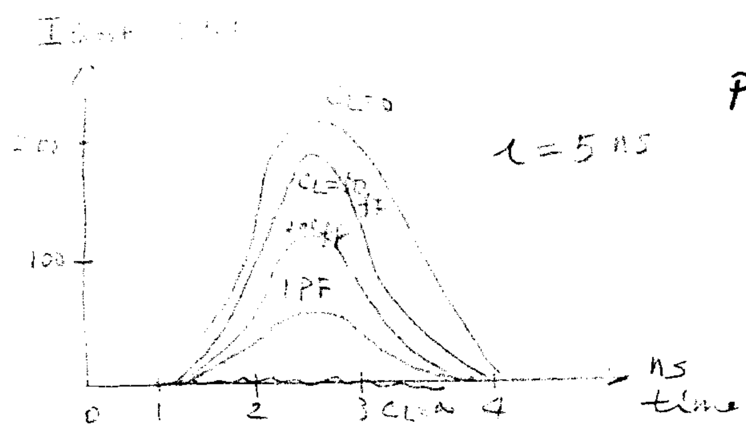
- Falling short-circuit energy

$$\mathcal{E}_{scf} = \frac{1}{2} t_f \left(V_{DD} + V_{Tp} - V_{Tn} \right) \cdot I_{scmaxf}.$$

- Short circuit energy per cycle.

$$E_{sc} = E_{scr} + E_{scf} = \left(V_{DD} + V_{Tp} - V_{Tn} \right) \left(t_r \cdot I_{scmaxr} + t_f I_{scmaxf} \right) \frac{1}{2}$$





$C_L \uparrow \Rightarrow I_{SC\,max} \downarrow$ why?

- If C_L is very large, → output fall time is much larger than input rise time.

  ∴ V_{ds} of PMOS is essentially 0 during the period

  ∴ Very small I_{sc}.

- If C_L is very small → output fall time is much smaller than input rise time.

  ∴ V_{ds} of PMOS is close to V_{dd} during most of the transition time.

  ∴ Very large I_{SC}.

Summary for short ckt energy:

① $t_r, t_f \uparrow \Rightarrow E_{sc} \uparrow$

② $C_L \uparrow \Rightarrow E_{sc} \downarrow$

③ when V_{DD} is scaled down to $\leq |V_{TP}| + V_{Tn}$, we have $E_{sc} = 0$. (~~No short circuit current~~).

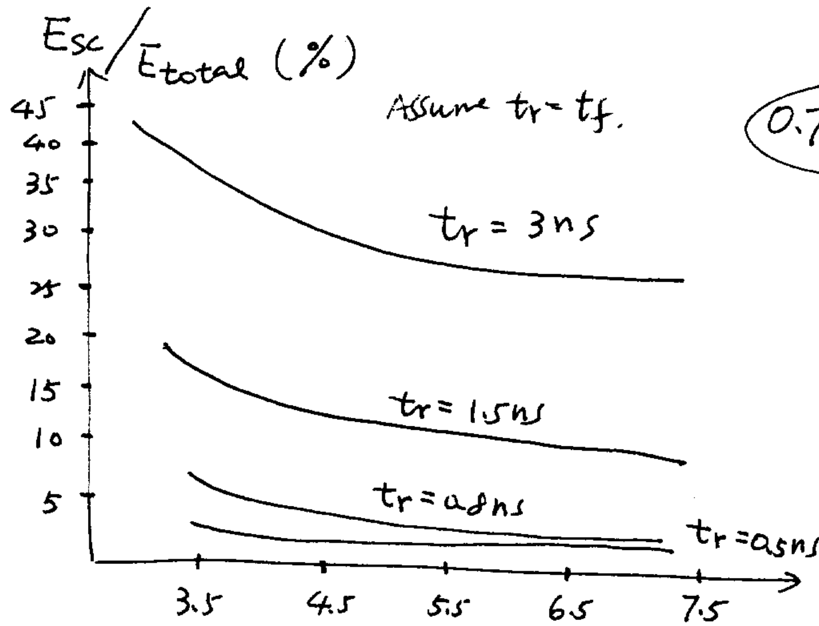
$\Downarrow$

$(V_{DD} + V_{TP} - V_{Tn} = 0)$

$\therefore E_{sc} = 0$.

But, the threshold values are not scaled proportionally.

• How important is short circuit energy?

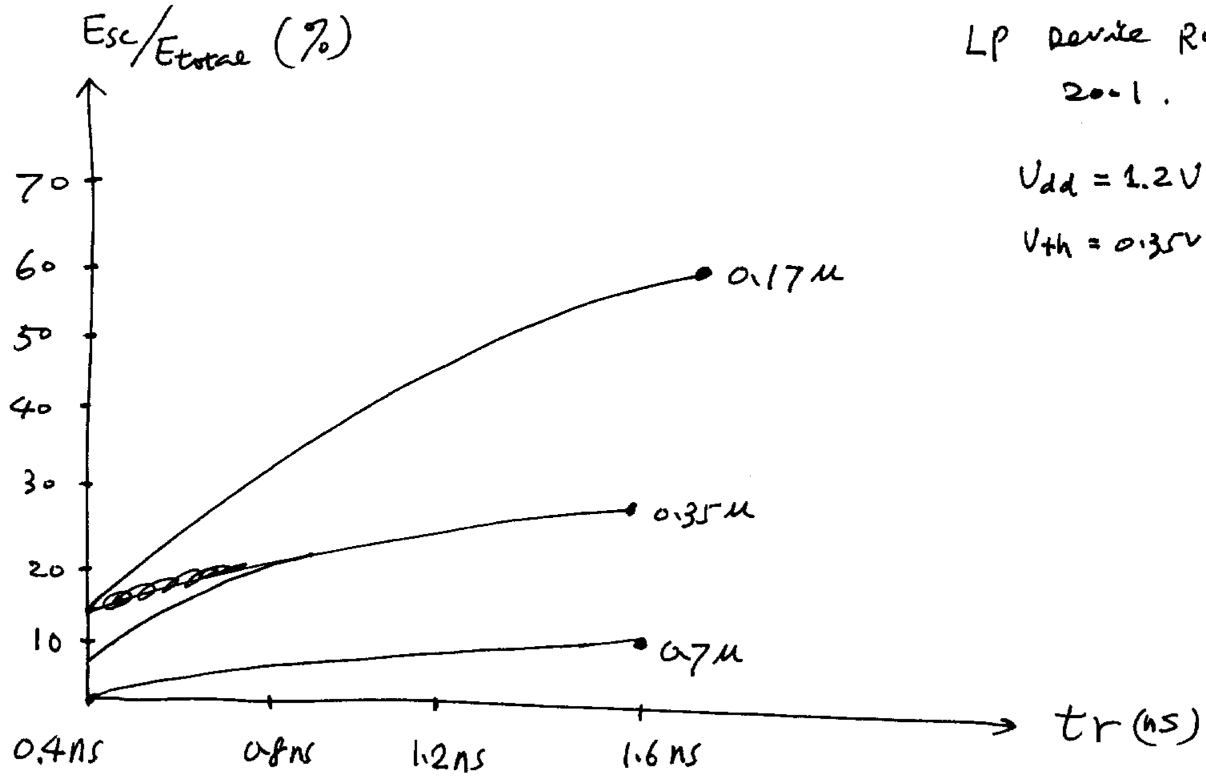


0.7μ technology

below 20% for typical transition time (0.5ns ~ 1.5ns)

$$E_{sc} \propto f(t_r, C_L)$$

- Effect of scaling



LP Device Roadmap 2001.

$V_{dd} = 1.2V$

$V_{th} = 0.35V$.

$$E_{sc} \propto f(t_r, \text{technology})$$

0.7μ ⟶ E_{sc}/E_{total} : 1% ~ 16%

0.35μ ⟶ 4% ~ 37%

0.17μ ⟶ 12% ~ 60%

Main reason:

The threshold Voltage does not scale down proportionally to Supply Voltage!!!

- Short circuit power model

Average short circuit power,

$$P_{avgsc} = (a_{0\to1} E_{sc})/T = a_{0\to1} E_{sc} \cdot f_{clk}$$

$a_{0\to1}$: Node activity

P(0→1 transition in a single period T).

$$P_{avgsc} = \frac{1}{2} E(sw) \cdot E_{sc} \cdot f_{clk}$$

where E(sw): average # of transitions 0→1 or 1→0

$$= \frac{1}{2} E(sw) \left[(V_{DD} + V_{Tp} - V_{Tn}) (t_r \cdot I_{scmaxr} + t_f I_{scmaxf}) \frac{1}{2} \right] \cdot f_{clk}$$

$$= \frac{1}{4} E(sw) \left[(V_{DD} + V_{Tp} - V_{Tn}) (t_r \cdot I_{scmaxr} + t_f I_{scmaxf}) \right] \cdot f_{clk}$$

Summary: P_{avgsc}

① proportional to f_{clk}, the clock frequency

② linear to the sum of the average transition energy of each node $\frac{E_{sc}}{2}$

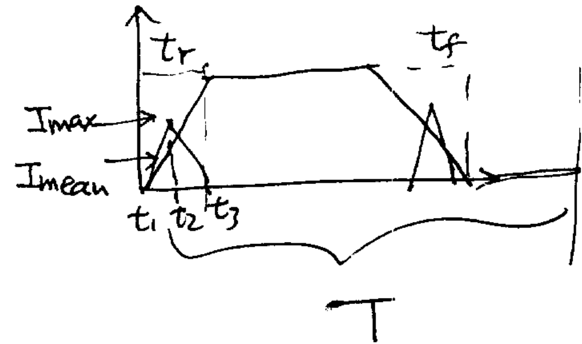
③ linear to the expected # of switchings.

• Analysis can be done by further extending $I_{sc\,max}$.

Assume $C_L = 0$ in which I_{sc} is maximum

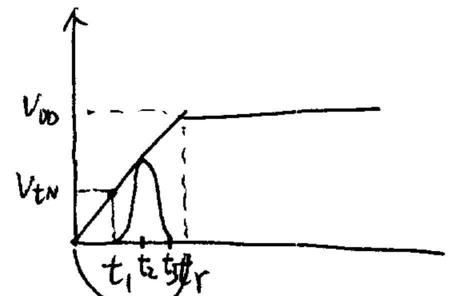
$$I_{sc}^{(max)} = \frac{\beta_N}{2}\left(V_{GS} - V_{TN}\right)^2$$

when saturate



$$\beta_N = \frac{\mu_n C_{ox}}{2}\,\frac{W}{L}$$

assume $\beta_N = \beta_P$ and $V_{TN} = |V_{TP}|$



we have

$$I_{sc}(t) = \frac{\beta_N}{2}\left(V_{in}(t) - V_{TN}\right)^2$$

$$\frac{V_{tN}}{t_1} = \frac{V_{DD}}{t_r}$$

$$\therefore t_1 = \frac{V_{TN}\cdot t_r}{V_{DD}}$$

$$\frac{V_i(t)}{t} = \frac{V_{DD}}{t_r = \tau} \quad \therefore V_i(t) = \frac{V_{DD}}{\tau}\cdot t$$

$$I_{mean} = \frac{2}{T}\cdot 2\int_{t_1}^{t_2} I_{sc}(t)\,dt$$

$$= \frac{4}{T}\int_{t_1}^{t_2} \frac{\beta_N}{2}\left(V_i(t) - V_{TN}\right)^2 dt$$

input rise time, input fall time

$\tau = \tau_r = \tau_f$

$$V_i(t) = \frac{V_{DD}}{\tau}\cdot t$$

$$t_1 = \frac{V_{TN}}{V_{DD}}\cdot\tau \qquad t_2 = \frac{\tau}{2}$$

$$= \frac{\beta_N \tau}{12 V_{DD} T}\left(V_{DD} - V_{TN}\right)^3$$

$$\therefore P_{sc} = I_{mean}\cdot V_{DD} = \frac{\beta_N \tau f}{12}\left(V_{DD} - V_{TN}\right)^3$$

- $P_{sc} \propto f$
- $P_{sc} \propto V_{DD}^3$
- P_{sc} : Designer can just control β and τ.

∴ Try to keep rising & falling edges fast!!!

$\tau \downarrow$

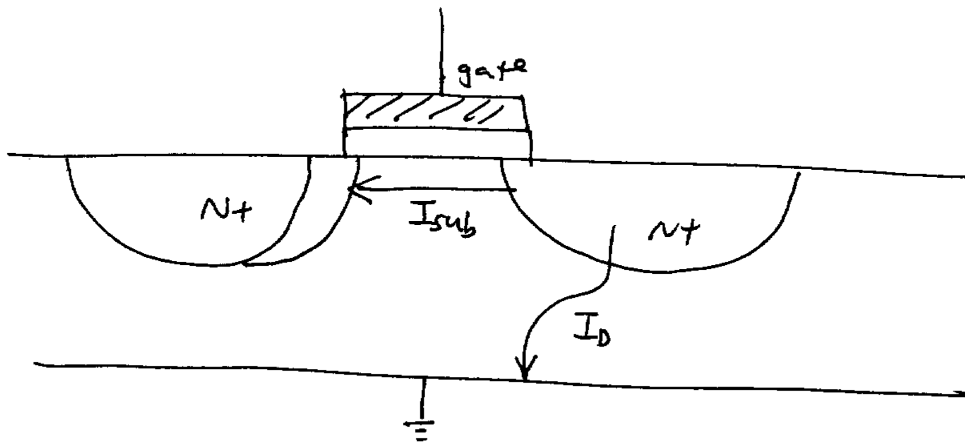
· Leakage power

* A small quiescent leakage current flow in CMOS from Vdd to Vgnd in static condition (inputs & internal nodes have constant values).

* I_{leak} ⟶ leakage power consumption (static)

* P_{leak} was negligible in old technologies

* With deep submicron scaling, I_{leak} & P_{leak} are more and more significant.

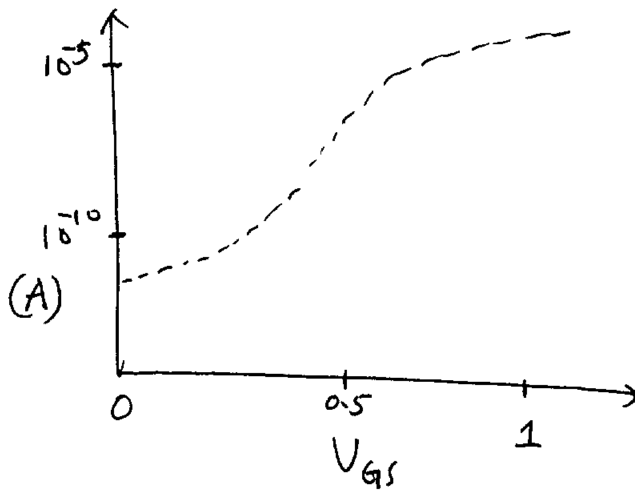


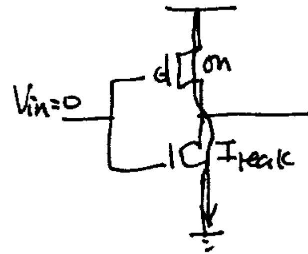
will discuss two major leakage sources:

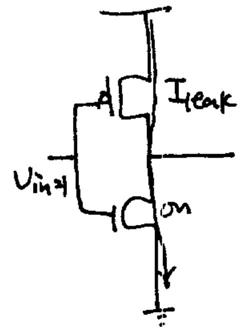
① subthreshold leakage current

② Reverse biased leakage current.

- Subthreshold leakage current — unexpected current flow from drain to source.







$$I_{sub} = \mu_0 C_{ox} \cdot \frac{W}{L} \cdot v_t^2 \, e^{\frac{V_{GS} - V_{TH}}{n \cdot v_t}}$$

μ_0: carrier surface mobility

C_{ox}: gate oxide capacitance / unit area

W.

L.

v_t: thermal voltage. $\propto$ (T: (absolute temperature)

$v_t = \frac{kT}{q} = 26mV$ at 300K k: Boltzmann Constant. ~~n: a technology~~ q: charge of electron).

n: A technology-dependent parameter

V_{GS}:

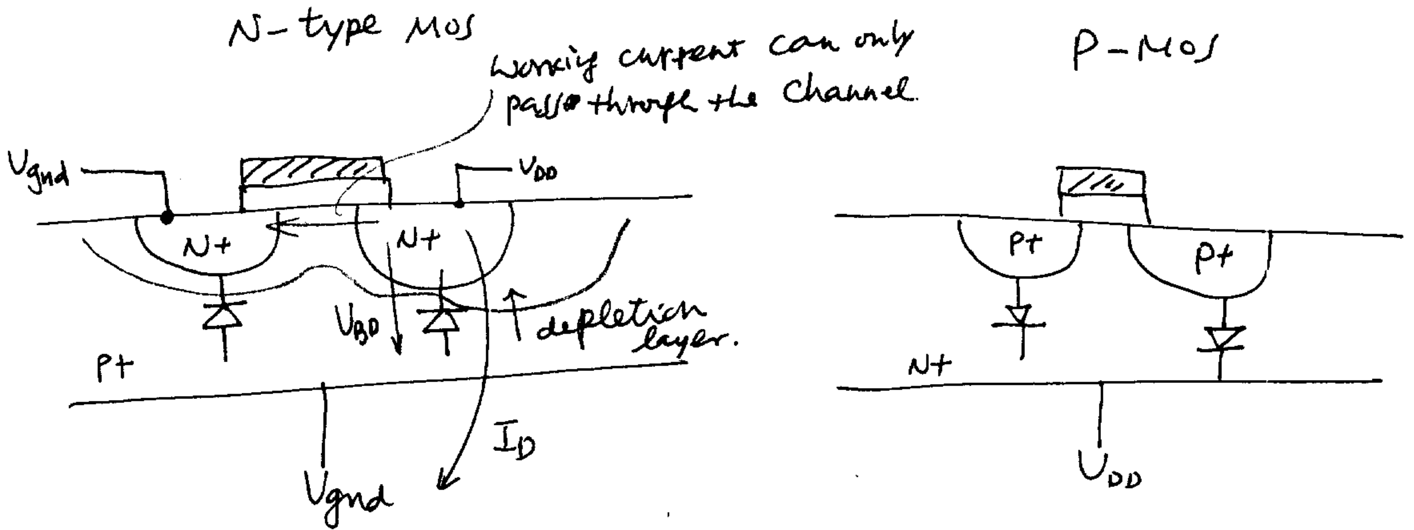
V_{TH}: threshold voltage.

Technogy advanced ⇒ V_{dd} ↓ ⇒ V_{TH} ↓ (for performance)

⇒ I_{sub} ↑

∴ for deep submicron VLSI, I_{sub} contribution to the leakage power grows in importance.

- Reversed-Biased leakage current



- Diodes are reversely biased

* I_D is related to temperature

$$T \uparrow \rightarrow I_D \uparrow$$

$$I_{BD} = I_S \cdot \left[e^{\frac{q|V_{BD}|}{KT}} - 1 \right]$$

I_S ↑ Saturation Current of p-n junction.

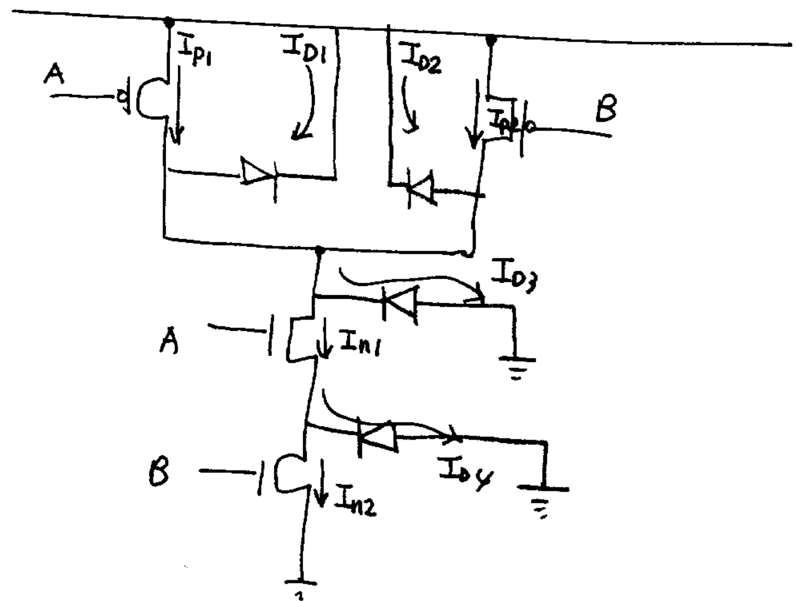
q: charge of electron

K: Boltzmann constant

T: Temperature (in terms of absolute temp.)

* In the level of fA (10^{-15} ampere) for 0.35 μ technology.

Exaple:



• The state and input of a ckt have influence on the leakage power

$\because$ $I_{sub} \propto V_{GS}$

$I_{BD} \propto V_{BD}$

$I_{BS} \propto V_{BS}$

$\therefore$ different circuit state gives different leakage current

Exaple: ISCAS C1355