IIT ROORKEE

NPTEL ONLINE CERTIFICATION COURSE

# MICROELECTRONICS: DEVICES TO CIRCUITS

## CMOS INVERTER BASICS - II

**SUDEB DASGUPTA**

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING



# Outline

- Propagation Delay
- Inverter capacitances
- Optimum NMOS-PMOS ratio
- Sizing of inverter chain
- Dynamic power dissipation
- Source of leakage currents
- Static power dissipation
- Power-delay product and energy-delay product

# CMOS INVERTER - Propagation delay

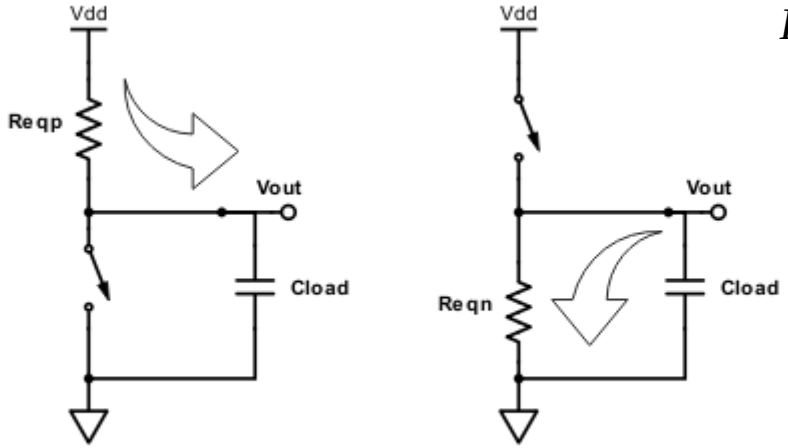



Figure : Transient response

$$R_{eq} = \frac{1}{V_{DD}/2}\int_{V_{DD}/2}^{V_{DD}} \frac{V}{I_{DSAT}(1+\lambda V)}dV \approx \frac{3}{4}\frac{V_{DD}}{I_{DSAT}}\left(1-\frac{7}{9}V_{DD}\right)$$

$$t_{pHL} = \ln(2)\,R_{eqn}\,C_L = 0.69\,R_{eqn}\,C_L$$

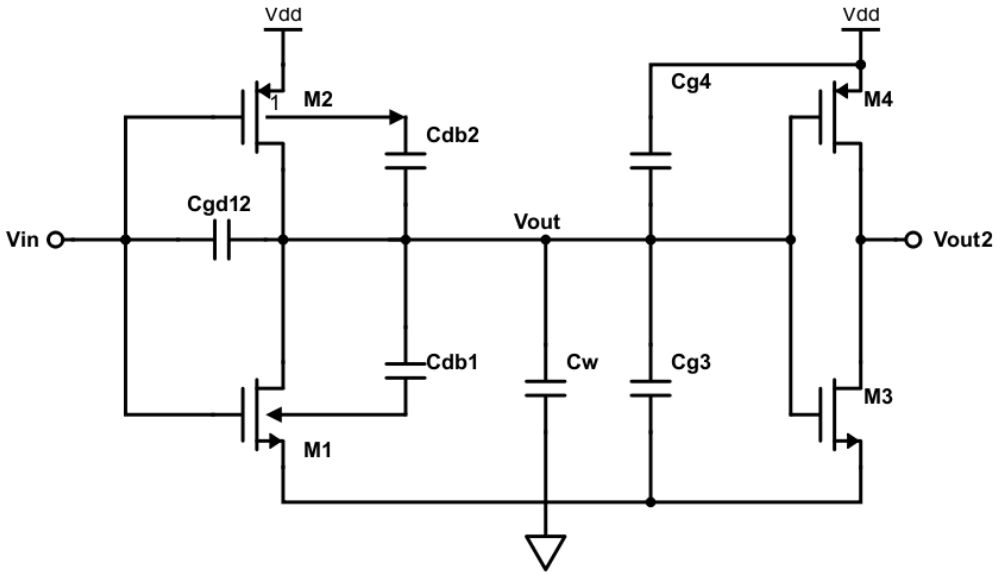
Similarly,

$$t_{pLH} = 0.69(R_{eqp}C_L)$$

Overall propagation delay,

$$t_p = \frac{t_{pHL}+t_{pLH}}{2} = 0.69\,C_L\left(\frac{R_{eqn}+R_{eqp}}{2}\right)$$

# CMOS INVERTER – Calculation of capacitance




* Parasitic capacitances in cascaded inverters

- For simplicity, we assume all capacitance are lumped together into one single capacitor $C_L$.
- $C_L$ capacitance can be breaks down to following components -
- $C_{gd12}$ = Gate drain capacitances
- $C_{db1,2}$ = Diffusion capacitances
- $C_w$= Wiring Capacitance
- $C_{g3}$, $C_{g4}$ = Gate capacitance of fan out

# CMOS INVERTER – Calculation of capacitance

**Gate drain capacitances $C_{gd12}$:**

- M1 and M2 are either in saturation or in cut off region. Under this condition only the gate drain capacitance is $C_{gd1}$.
- Because the signal swing is opposite in both terminal, effective capacitance $C_{gd}=2C_{GD0}W$. Where, $C_{GD0}$ = overlap capacitance per unit width.

**Diffusion Capacitances $C_{db1,2}$ :**

- **Capacitance between drain and bulk is due to the reverse- biased pn-junction.**
- **Simplified representation of diffusion capacitance , $C_{eq} = K_{eq}C_{j0}$**
- **Where, $C_{j0}$= junction capacitance under zero-bias condition.**
- **$K_{eq}$= multiplication factor.**

# CMOS INVERTER – Calculation of capacitance

## Wiring Capacitance $C_w$:

- This capacitance appears due to the length and width interconnecting wires.
- This is function of distance of the fan-out from the driving gate.

## Gate Capacitance of Fanout $C_{g3}$ and $C_{g4}$

- This is equals to the total gate capacitance of the loading gates $M_3$ and $M_4$.

$$C_{fan-out} = C_{gate}(\text{NMOS}) + C_{gate}(\text{PMOS})$$

$$C_{fan-out} = (C_{GSOn} + C_{GDOn} + W_n L_n C_{ox}) + (C_{GSOp} + C_{GDOp} + W_p L_p C_{ox})$$

# CMOS INVERTER – How to improve design technique

- Reduce $C_L$ : By careful layout design
- Increase transistor sizing: Take care of self-loading as well
- Increase $V_{DD}$ : consider trade off energy dissipation

# CMOS INVERTER – Optimal value of NMOS-to-PMOS Ratio

- While improving the PMOS width improves $t_{pLH}$ of the inverter by increasing the charging current, it also degrade the $t_{pHL}$ by causing a large parasitic capacitance.
- If the optimum ratio $\beta$ , where

$$\beta = (W/L)_p/(W/L)_n$$

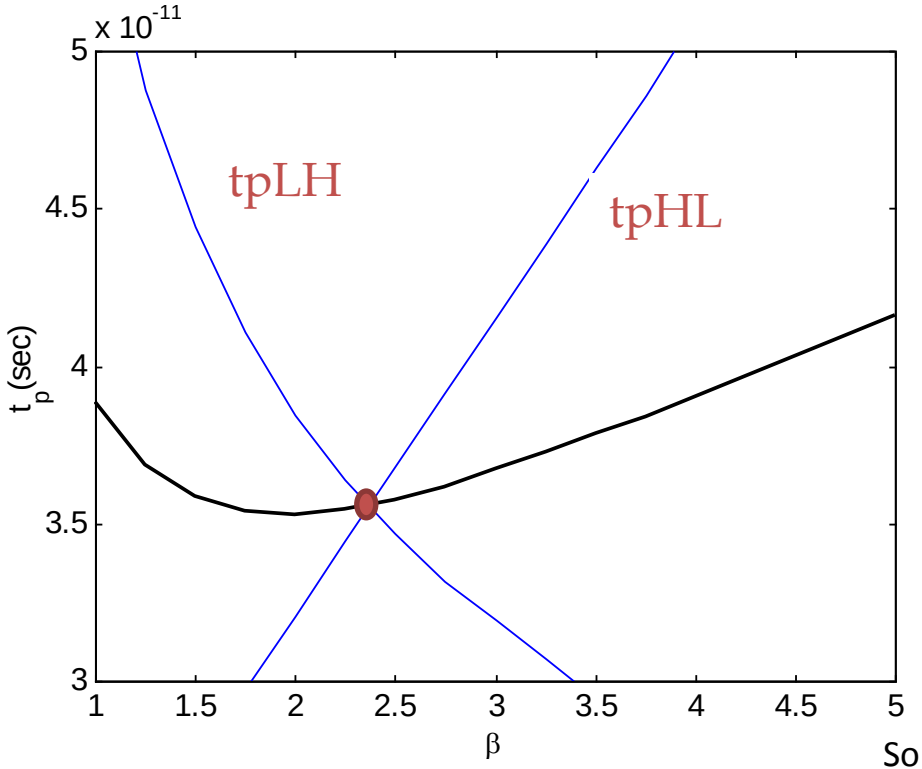
We know that $C_L = \left(C_{dp1} + C_{dn1}\right) + \left(C_{gp2} + C_{gn2}\right) + C_w$

Where, $C_{dp1} \approx \beta C_{dn1}$ and $C_{gp1} \approx \beta C_{gn1}$

So, $C_L = \left((1+\beta)(C_{dn1} + C_{gn2}) + C_w\right)$

$$t_p = \frac{0.69}{2}\left((1+\beta)(C_{dn1} + C_{gn2}) + C_w\right)\left(R_{eqn} + \frac{R_{eqp}}{\beta}\right)$$

# CMOS INVERTER – Optimal value of NMOS-to-PMOS Ratio




Taking $\frac{\delta t_p}{\delta \beta} = 0$

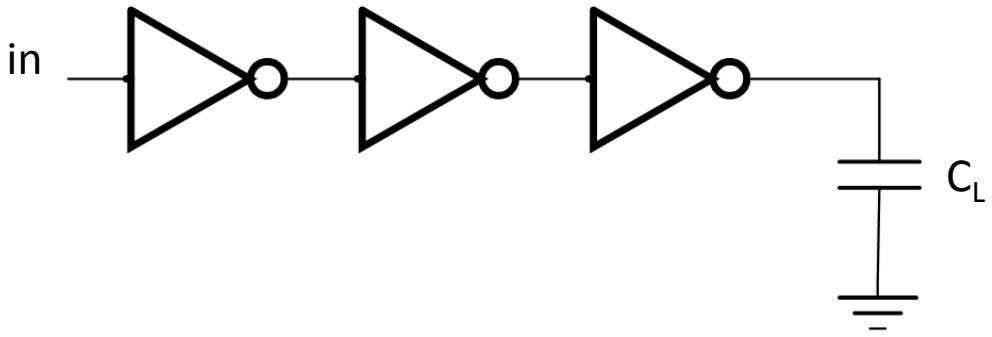
The optimal value of β would be

$$\beta_{opt} = \sqrt{r\left(1 + \frac{C_w}{C_{dn1} + C_{gn2}}\right)}$$

$\beta = W_p / W_n$

Source :Digital Integrated Circuits (2nd Edition)- Jan M. Rabaey

# Sizing of Inverter chain




- For a given $C_L$, what is the minimum propagation delay ?
- What is the number of inverter stages ?
- What is the optimum NMOS-to-PMOS ratio.

- For an inverter, the load capacitance can be divided into an intrinsic and extrinsic capacitance component.

$$C_L = C_{int} + C_{ext}$$

$C_{int}$= intrinsic output capacitance of the inverter, associated with diffusion capacitances.
$C_{ext}$= represents the fan out and wire capacitances.

# Sizing of Inverter chain

Propagation delay, $t_p = 0.69 R_{eq} \left( C_{int} + C_{ext} \right)$

$$t_p = 0.69 R_{eq} C_{int} \left( 1 + C_{ext} / C_{int} \right)$$

$R_{eq}$= equivalent resistance of gate

The delay of a inverter it self $t_{p0} = 0.69 R_{eq} C_{int}$

$$\boxed{t_p = t_{p0} \left( 1 + C_{ext} / C_{int} \right)}$$

# Sizing of Inverter chain

- Sizing up an inverter reduces its delay, but it also increases its input capacitance.
- The input gate capacitance and intrinsic output capacitance are function of transistor gate size.

$$C_{int} = \gamma C_g$$

Where, $\gamma$ = proportionality factor and function of technology

$$t_p = t_{p0}\left(1 + f/\gamma\right)$$

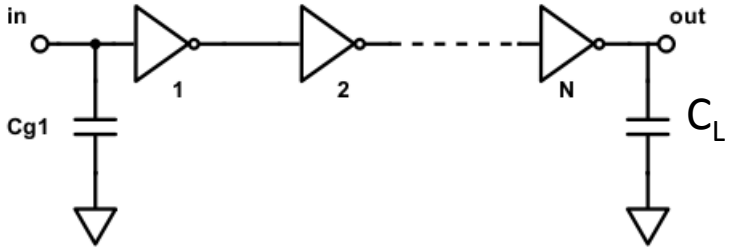
Where, $f = C_{ext}/C_{int}$ Is defined as effective fan-out.

The delay of the inverter is a function of the ratio between external load capacitance and its intrinsic capacitance.

# Sizing of Inverter chain

$t_p = t_{p1} + t_{p2} + \ldots + t_{pN}$

$$t_p = \sum_{j=1}^{N} t_{p,j} = t_{p0} \sum_{i=1}^{N} \left( 1 + \frac{C_{gin,j+1}}{\gamma C_{gin,j}} \right), \quad C_{gin,N+1} = C_L$$




- Delay equation has N - 1 unknowns
- Minimize the delay, find N - 1 partial derivatives and solve for $\delta t_p / \delta C_{g,i} = 0$

$$C_{g,j+1} / C_{g,i} = C_{g,i} / C_{g,j-1}$$

- Size of each stage is the geometric mean of two neighbors

# Sizing of Inverter chain

The optimum size of each inverter is the geometric mean of its neighbors sizes :

$$C_{gin,j} = \sqrt{C_{gin,j-1} C_{gin,j+1}}$$

This means that each inverter is sized up by the factor f with respect to the preceding gate, has the same effective fan-out ($f_j$=f),

$$f = \sqrt[N]{C_L / C_{g,1}} = \sqrt[N]{F}$$

The minimum delay through the chain as

$$t_p = N t_{p0} (1 + \sqrt[N]{F} / \gamma)$$

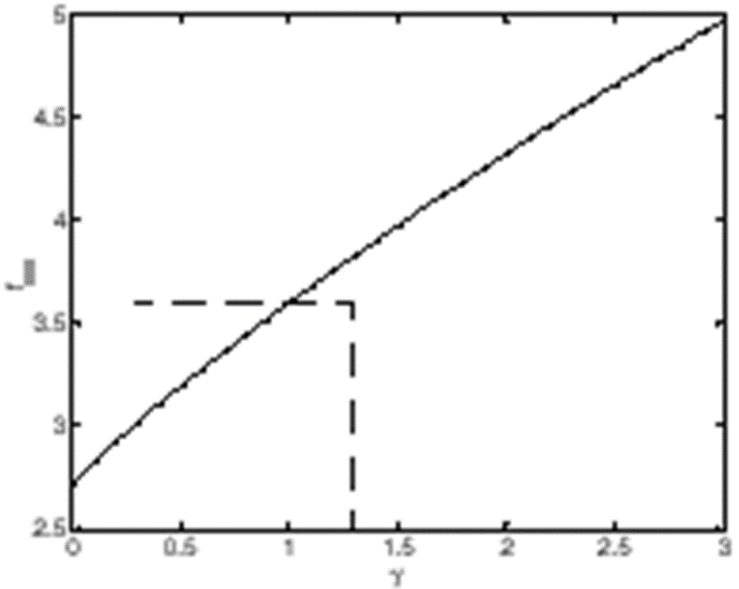
F represents the over all effective fan-out of the circuit.

# Choosing right number of stages

The optimum value of right number of stages can be found by differentiating the minimum delay expression be the number of stages and setting the result to 0.

We get $\gamma + \sqrt[N]{F} - \frac{\sqrt[N]{F}}{N} = 0$

In common practice, optimum fan-out could be selected as 4




Source :Digital Integrated Circuits (2nd Edition)- Jan M. Rabaey

# Power Dissipation –Dynamic power

- Power dissipation during switching activity
- Energy taken from supply voltage = $E_{VDD}$

$$E_{VDD} = \int_0^{\infty} i_{VDD}(t)\, V_{DD}\, dt = C_L V_{DD} \int_0^{VDD} dv_{out} = C_L V_{DD}^2$$

- Energy stored/removed on the load capacitor = $E_C$

$$E_C = \int_0^{\infty} i_{VDD}(t) v_{out} dt = C_L \int_0^{VDD} v_{out} dv_{out} = \frac{C_L V_{DD}^2}{2}$$

This is independent of transistor size.

If the switching activity is $f_{0\to1}$ times per second

$$P_{dyn} = C_L V_{DD}^2 f_{0\to1}$$

# Power Dissipation – Direct path currents

- Finite value of rise time and fall time of input signal cause a direct path between $V_{DD}$ and GND, while PMOS and NMOS are conducting simultaneously.

$$E_{dp} = V_{DD}\frac{I_{peak}t_{sc}}{2} + V_{DD}\frac{I_{peak}t_{sc}}{2} = t_{sc}V_{DD}I_{peak}$$

- Average power consumption

$$\boxed{P_{dp} = t_{sc}V_{DD}I_{peak}f}$$

Where, $t_{sc}$= time both devices are conducting

# Power Dissipation – Static power dissipation

This is the steady state power dissipation, when no switching activity is present

$$P_{stat} = I_{stat} V_{DD}$$

Source of leakage current :

- Reverse bias diode junction of the transistors, located between source or drain and body. Thermally generated carriers may affect the junction leakage current.
- Sub-threshold leakage current when $V_{GS} < V_T$
- Choice of threshold voltage is a trade off between performance and static power consumption

$$P_{tot} = P_{dyn} + P_{dp} + P_{stat} = (C_L V_{DD}^2 + V_{DD} I_{peak} t_s) f_{0\rightarrow 1} + V_{DD} I_{leak}$$

# Power Delay Product and Energy Delay Product

Power-delay Product : It represents the average energy consumption per switching events.

$$PDP = C_L V_{DD}^2 f_{max} t_p = \frac{C_L V_{DD}^2}{2}$$

Because, $f_{max} = \frac{1}{2t_p}$

This value can be arbitrary by changing the $V_{DD}$ .

Energy-Delay Product : This performance matrix combines the performance and energy.

$$EDP = PDP \times t_p = \frac{C_L V_{DD}^2}{2} t_p$$

# Techniques for Reducing Power Dissipation

- Reduce supply voltage
- Reduce switching activity
- Reduce physical capacitances

# Recapitulation

- Propagation delay is the time to charge or discharge of the load capacitance.
- Optimum ratio of NMOS-to-PMOS is required to find out minimum propagation delay.
- The intrinsic delay is independent of transistor sizing and depends on the technology parameter and physical layout
- In very charging or discharging of inverter, only one half of the energy stored or removed in load capacitance respectively. Other half is dissipated by PMOS or NMOS.
- Power consumption is dissipated by the dynamic power during the switching period.

# Thank You