CHAPTER 3

DIGITAL MOS CIRCUITS

3.1 The MOS TRANSISTOR

MOS integrated circuits use MOSFET (Metal Oxide Semiconductor Field Effect Transistor), which are 4-terminal device, as the building block. Two basic types of MOSFET are the NMOS (n-channel device) and PMOS (p-channel device) transistors.

3.1.1 MOS Characteristics

MOSFET is a voltage-controlled device. The gate-source voltage, V_{gs}, controls the channel width and hence the amount of current through the device.

Although MOS transistors are symmetrical device, by convention the majority carriers (electrons for NMOS and holes for PMOS) flow from the source to the drain. Thus, for NMOS, the source terminal is usually of a lower potential than the drain while in PMOS, source is usually of higher potential than the drain.

The input impedance at the gate terminal is also very high thus the gate current is very small, that is, $I_G \sim 0$. As the gate terminal is isolated from the channel/substrate by a layer of oxide, it basically presents itself as a capacitive load to an input signal.

3.1.2 MOS Transistor Symbols

Figure 3-1 shows various symbols used for the different types of MOS transistors. Those on the top row are symbols with 4 terminals. The terminal in the center is the Bulk (B) terminal. As the Bulk terminal is usually connected to the source terminal, simplified symbols as shown in the second row are the more commonly used in circuit representation.

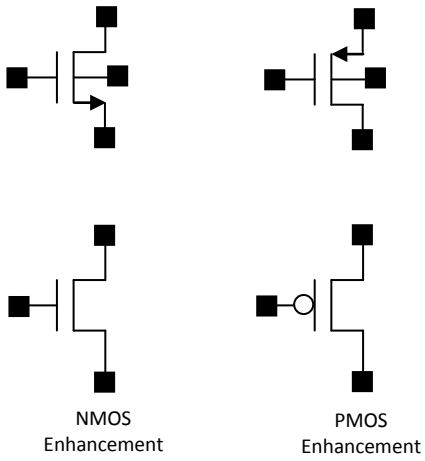
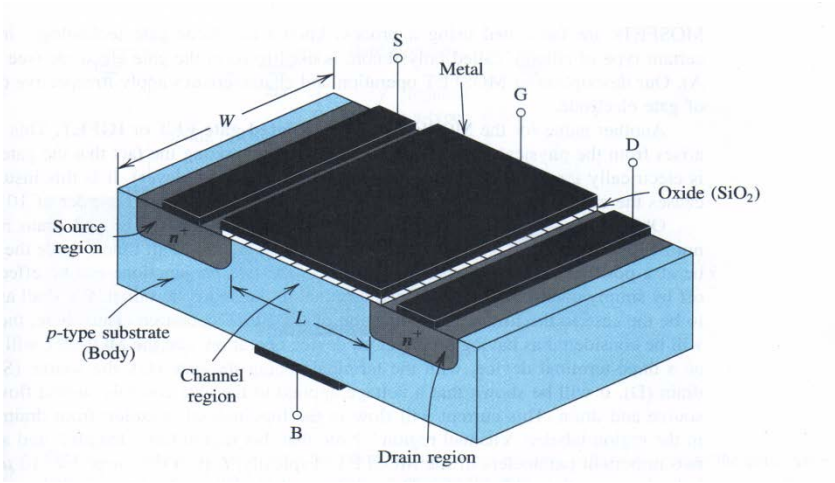


Figure 3-1 MOS Transistor Symbols (4 Terminals & 3 Terminals)

3.1.3 Basic NMOS Device Structure

Figure 3-2 shows the physical and cross sectional structure of a NMOS transistor. It is formed on a p-type substrate. The source and drain regions are formed by introducing impurities of type VI (P or As) into the p-type substrate until the region becomes a n-type semiconductor.

The distance between the source and drain is defined as the channel length and it is parallel to the direction of current flow. Channel width of the transistor is the dimension perpendicular to the direction of current flow. The channel length and width affect the performance of a MOS transistor.



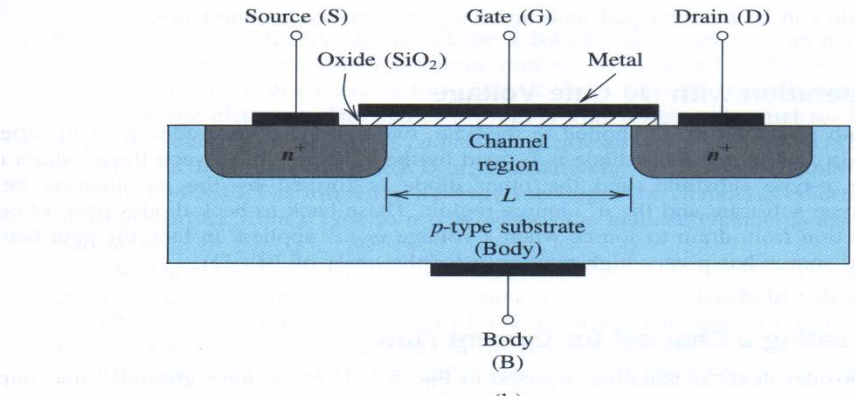


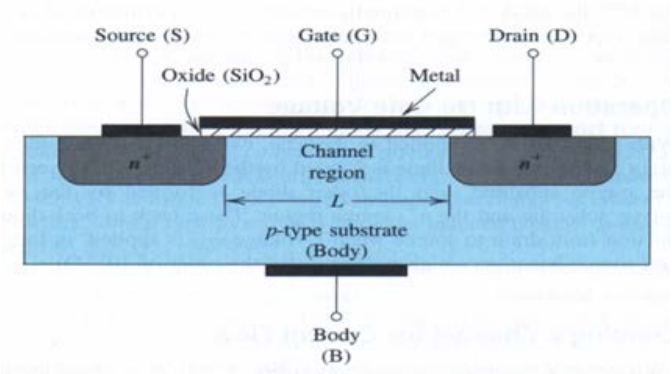
Figure 3-2 Physical & Cross Sectional Structure of NMOS Transistor

3.1.4 Operation of NMOS Enhancement Transistor

A basic enhancement mode MOSFET does not have a channel established at fabrication thus it is a normally OFF device. The transistor have 3 modes of operation, namely :-

(i) Cut-off Mode

When the voltage applied to the gate of the transistor is such that $V_{gs} < V_t$, where $V_t \approx 0.2V_{dd}$, no conducting channel is induced resulting in the path resistance between the source and drain being very high and hence $I_{ds} = 0$.



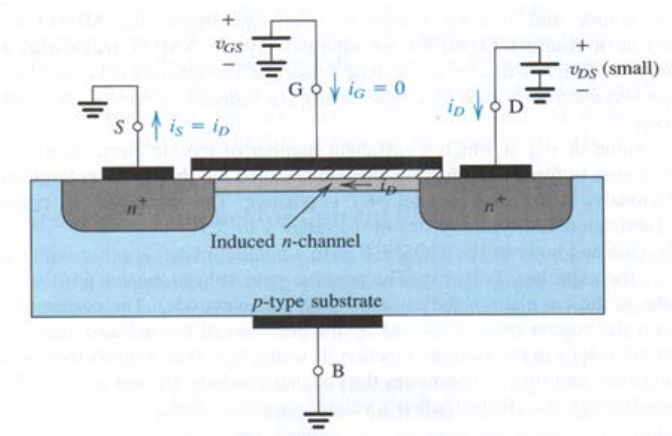
Cutoff Mode

- $V_{gs} < V_t$ and $V_t \approx 0.2V_{dd}$
- $I_{ds} = 0$

Figure 3-3 NMOS transistor in Cutoff Mode Operation

(ii) Triode/Linear Mode

When $V_{gs} \geq V_t$ and $V_{ds} < (V_{gs} - V_t)$, electrons (carriers) accumulated near the surface under the gate region starts to gather to create a conducting channel between the source and the drain. With the channel, current can now flow from the drain to the source (opposite to electron flow). In this region, I_{ds} increases with both the drain and gate voltage.



Linear/Triode Mode

- $V_{gs} \geq V_t$ and $V_t \approx 0.2V_{dd}$
- $V_{ds} < (V_{gs} - V_t)$

$I_{ds} = K_n[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$

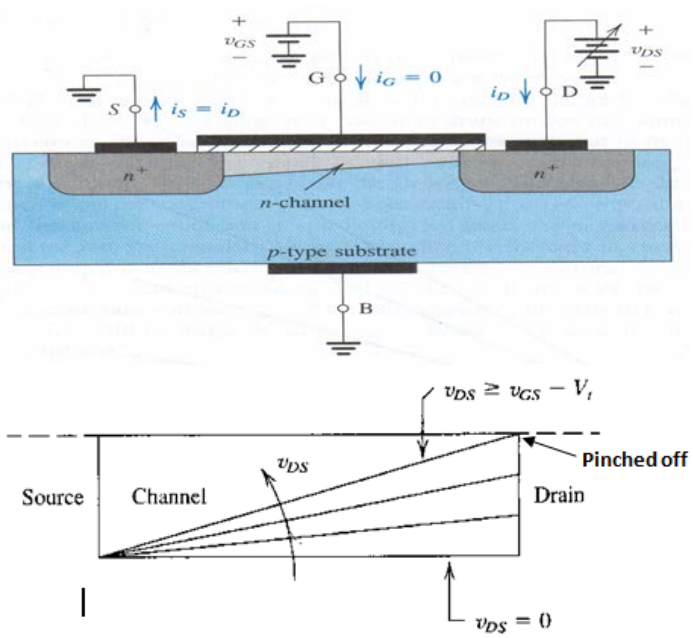
where

$K_n = \mu_n \varepsilon_o \varepsilon_r / t_{ox}$
ε_o = permittivity of free space
ε_r = relative permittivity of SiO_2
t_{ox} = thickness of gate oxide

Figure 3-4 NMOS transistor in Triode/linear Mode Operation

(iii) Saturation Mode

When $V_{gs} \geq V_t$ and $V_{ds} \geq (V_{gs} - V_t)$, the channel will become tapered, as shown in Figure 3-5, because the effective voltage on the channel decreases towards the drain end. In this region, the current I_{ds} ceased to be influenced by the drain voltage. The I_{ds} under this condition is given as :-



Saturation Mode

- $V_{gs} \geq V_t$ and $V_t \approx 0.2V_{dd}$
- $V_{ds} \geq (V_{gs} - V_t)$
- $I_{ds} = \{K_n/2\}[W/L](V_{gs} - V_t)^2$

where

$K_n = \mu_n \varepsilon_o \varepsilon_r / t_{ox}$
ε_o = permittivity of free space
ε_r = relative permittivity of SiO_2
t_{ox} = thickness of gate oxide

Figure 3-5 NMOS Transistor in Saturation Mode

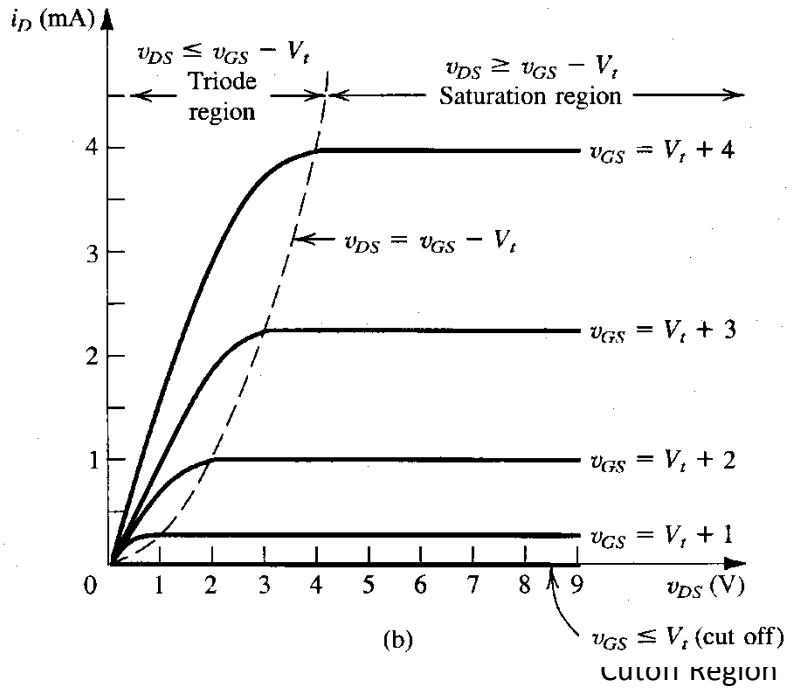


Figure 3-6 NMOS I_{ds} vs V_{ds} Characteristic Curve

In summary, the equations for NMOS enhancement transistor are :-

OFF : $V_{gs} < V_t$ where $V_t \approx 0.2V_{dd}$

ON : $V_{gs} \geq V_t$

(a) Linear Region : $V_{ds} < (V_{gs} - V_t)$
$I_{ds} = K_n[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$

(b) Saturation Region : $V_{ds} \geq (V_{gs} - V_t)$
$I_{ds} = [K_n/2][W/L](V_{gs} - V_t)^2$

3.1.5 PMOS transistor

PMOS transistor needs a N-type substrate with both source and drain doped with type III impurities (Boron) to form p-type regions. The threshold voltage (V_t) is approximately $-0.2V_{dd}$. To turn ON an enhancement PMOS, a gate-source voltage of $V_{gs} \leq V_t$ is required.

The equations for PMOS enhancement transistor are :-

OFF : $V_{gs} > V_t$ where $V_t \approx -0.2V_{dd}$

ON : $V_{gs} \leq V_t$

(b) Linear Region : $V_{ds} > (V_{gs} - V_t)$
$I_{ds} = K_p[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$

(b) Saturation Region : $V_{ds} \leq (V_{gs} - V_t)$
$I_{ds} = [K_p/2][W/L](V_{gs} - V_t)^2$

	NMOS	**PMOS**
OFF	$V_{gs} < V_t$ where $V_t \approx 0.2V_{dd}$	$V_{gs} > V_t$ where $V_t \approx -0.2V_{dd}$
ON : Linear Region	$V_{gs} \geq V_t$ $V_{ds} < (V_{gs} - V_t)$ $I_{ds} = K_n[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$	$V_{gs} \leq V_t$ $V_{ds} > (V_{gs} - V_t)$ $I_{ds} = K_p[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$
ON : Saturation Region	$V_{ds} \geq (V_{gs} - V_t)$ $I_{ds} = [K_n/2][W/L](V_{gs} - V_t)^2$	$V_{ds} \leq (V_{gs} - V_t)$ $I_{ds} = [K_p/2][W/L](V_{gs} - V_t)^2$

3.2 THE MOS SWITCHES

Transistors can be biased to operate as switches. In the case of MOS transistor, a control signal (SN) can be applied to the gate of the device to turn it ON or OFF.

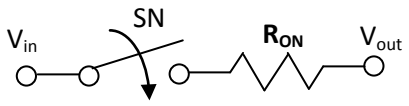


Figure 3-7 A Simple Switch Representation

3.2.1 NMOS Pass Transistor

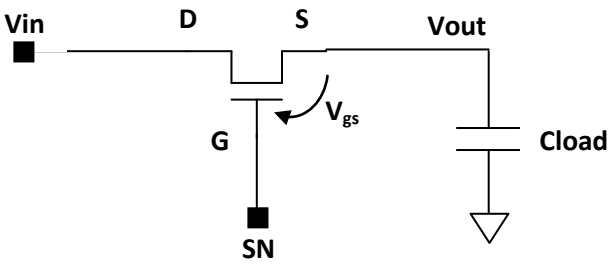


Figure 3-8 NMOS Pass Transistor

Figure 3-8 shows a switch using NMOS transistor. Such switch is known as Pass Transistor. The signal, SN, at the gate of the transistor controls the ON/OFF state of the transistor. When SN = Low, the transistor is OFF and when SN = High (V_{dd}), the transistor is ON. In the case of NMOS enhancement transistor, the signal voltage is actually V_{gs} and it must be greater than V_t, in order for the transistor to turn ON.

Operation

(i) V_{in} = High and C_{load} is initially uncharged

C_{load} will start to charge when the switch is ON, however V_{out} can only reach V_{in}-V_t because exceeding this value, the ON condition will cease to exist and the transistor will turn OFF. Hence, it can be concluded that NMOS transistor transfer logic High with degradation. That is, the V_{out} cannot reach full $V_{in.}$ The direction of the current flow is from left to right.

(ii) V_{in} = Low and C_{load} is initially charged to V_{in}-V_t

With the switch ON, C_{load} will start to discharge and the current flows from right to left. This action can continues until C_{load} discharges to 0V. Thus, NMOS transistor transfers logic Low without degradation.

3.2.2 PMOS Pass Transistor

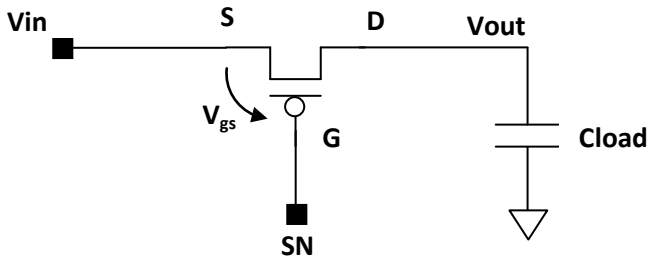


Figure 3-9 PMOS Pass Transistor

In PMOS Pass transistor, a logic low signal is applied at SN to turn ON the transistor. Hence, when SN = High (V_{dd}), the switch will turn OFF. In the case of PMOS enhancement transistor, the signal voltage, V_{gs}, must be less than V_t, in order to turn ON the transistor.

Operation

(i) V_{in} = High and C_{load} is initially uncharged

With SN = Low such that $V_{gs} < V_t$, the transistor is ON and current flows from left to right thus charging C_{load} to V_{in}. For PMOS transistor, the transfer of High logic level is without degradation.

(ii) V_{in} = Low and V_{out} is initially charged to V_{dd}

Keeping the switch ON, C_{load} will starts to discharge and current flows from right to left. However, C_{load} can only discharges until V_{out} is V_t because going beyond this will cause the transistor to turn OFF. Hence, PMOS transistor transfers Low logic with degradation.

3.2.3 CMOS Pass Transistor

Both NMOS and PMOS devices exhibit poor performance when transmitting one of two logic information. A CMOS switch is made up of a NMOS and an PMOS pass transistor connected in parallel. It is controlled by a pair of complementary signals so that both transistors are either “ON” or “OFF” at the same time.

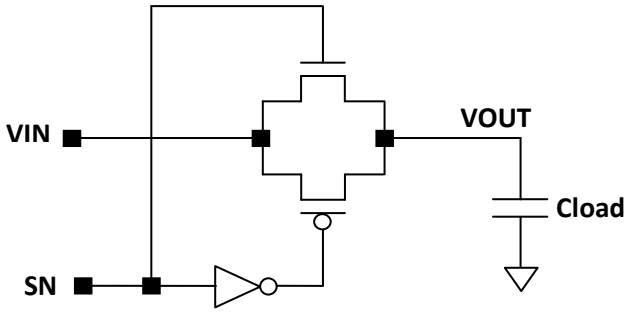


Figure 3-10 CMOS Transmission Gate

The PMOS transistor can charge V_{out} to V_{dd} without degradation while NMOS transistor can discharge V_{out} to V_{ss} (0V) without degradation, thus a CMOS transmission gate gives good logic High and logic Low level transmission.

3.2.4 Power Supply Voltage

V_{dd} has decreased in modern processes, typical V_{dd} for 180nm (0.18um) technology ranges from 1.8V to 3.3V. A high V_{dd} would damage these modern tiny transistors. On the other hand, a lower V_{dd} saves power.

3.3 STATIC CMOS CIRCUITS

3.3.1 CMOS Inverter

Figure 3-11 shows a CMOS inverter. In this inverter, both NMOS and PMOS transistors are used. The PMOS function as the load while the NMOS function as the driver device.

For a V_{in} = High, the PMOS transistor will turn OFF and the NMOS transistor will turn ON. Thus, discharging can take place via the pull-down (driver) device and give a logic low output.

Similarly, for a V_{in} = Low, the PMOS transistor will turn ON while the NMOS transistor will turn OFF, hence the output charges towards V_{dd} via the pull-up (load) device and gives a logic high output.

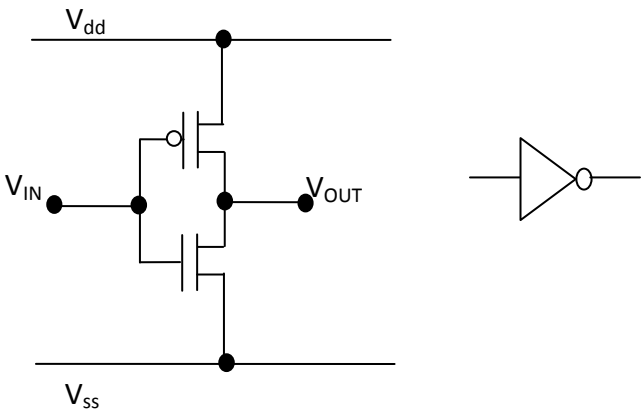
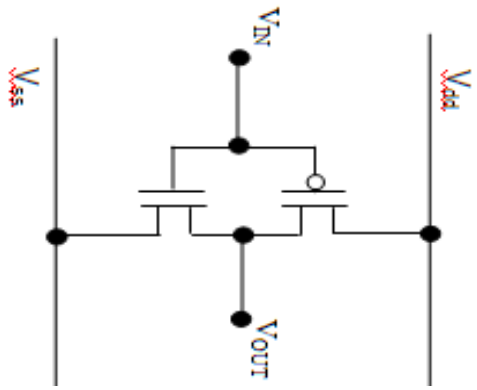


Figure 3-11 CMOS Inverter Circuit



CMOS Inverter Circuit (rotated 90^{o})

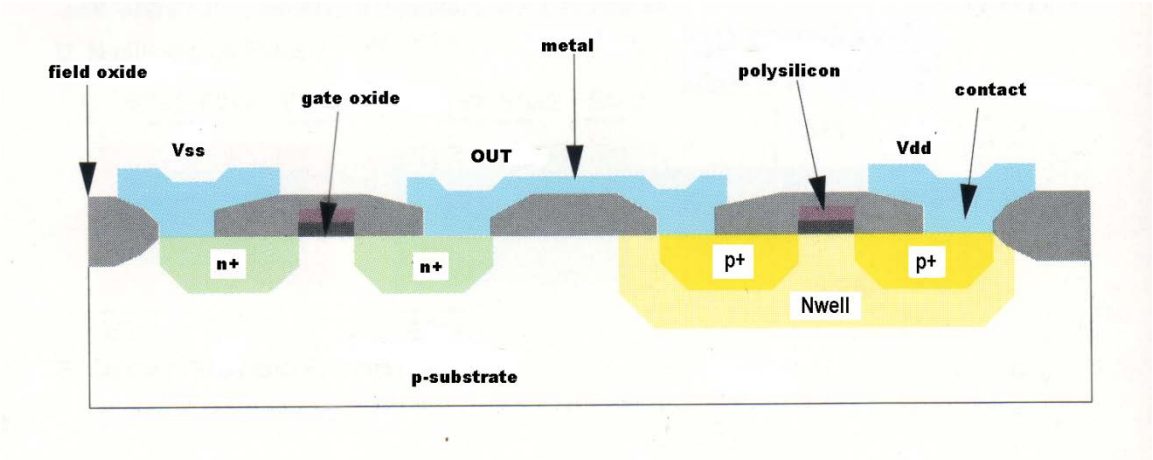


Figure 3-12 Cross sectional view of CMOS inverter

Figure 3-13 shows the amount of current flowing between the supply rails of a CMOS inverter. It can be divided into 5 sub-regions as follows :-

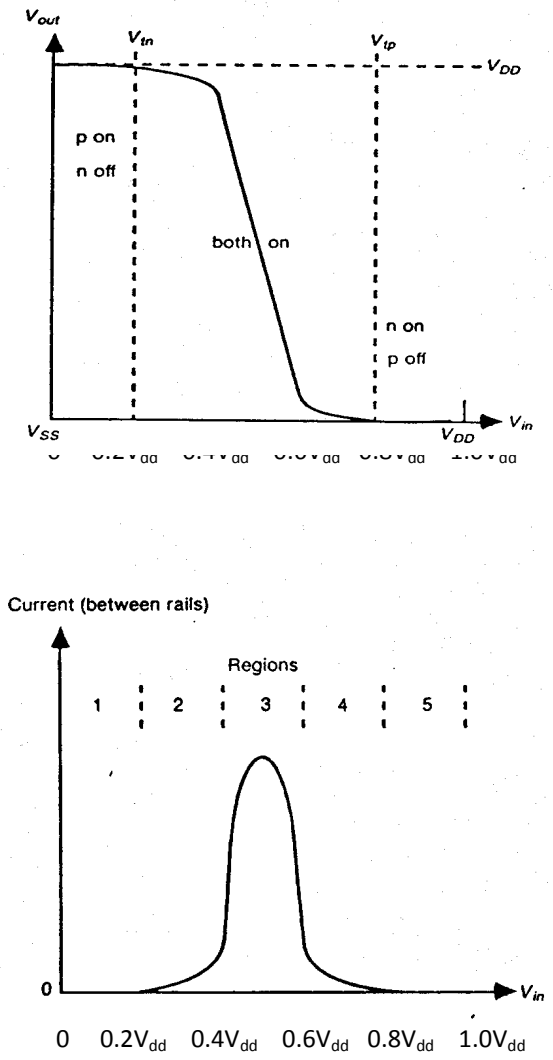


Figure 3-13 CMOS Inverter Current vs V_{in}

Region 1

- V_{in} = 0.1V_{dd} , PMOS fully turned on (Linear) and NMOS is OFF (Cutoff).
- No current flows through the inverter and output is connected to V_{dd}.

Region 2

- V_{in} is increased to a level greater than V_t (eg V_{in}=0.3V_{dd}). This causes the NMOS to turn ON in the Saturation mode and the PMOS remains in the Linear mode.
- Small amount of current flows through the inverter.

Region 3

- With V_{in}= 0.5V_{dd}, both NMOS and PMOS transistor operates in the Saturation mode.

Region 4

- With V_{in} = 0.7V_{dd}, thus $V_{gs} < V_t$ for PMOS resulting in PMOS operating in the Saturation model while NMOS transistor operates in the Linear mode.
- Output voltage drops close to V_{ss}.

Region 5

- V_{in}= 0.9V_{dd}, NMOS is fully turn on (Linear) and PMOS is OFF (Cutoff).
- No current flows through the inverter.

Hence, for CMOS inverter, there is no static current flowing but the inverter draws a high current when the transistors are changing states in Region 3. This current is known as the dynamic current.

3.3.2 Switching Time

How fast a logic gate can switch from one logic state to the next depends on its load capacitance and its wiring path resistance. The gate capacitance of the driven stage and other parasitic capacitance along the way constitute the load capacitance. As for the wiring path resistance, it is dependent on the logic state of the logic gate, which in turn determines the path of current flow.

3.3.2.1 Determination of Path Resistance

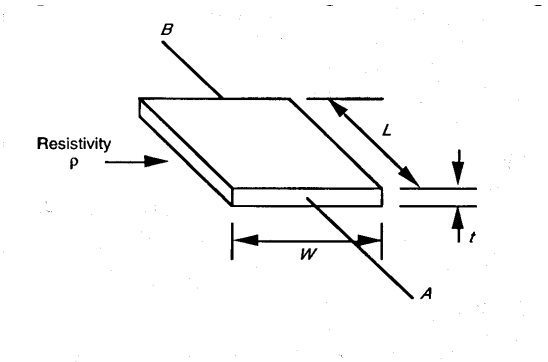


Figure 3-14 Sheet Resistance Model

Resistance of a material is given as:

$$R = \rho L/A \qquad \Omega$$

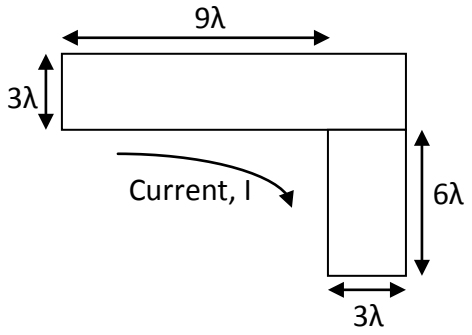
If the material has a sheet resistance of $R_s = \rho/t$ [$\Omega/\square$], then the resistance of the material can be expressed as :

$$R = R_s(L/W) \qquad [\Omega]$$

Table 3-15 gives the sheet resistance of various materials commonly used in the fabrication of MOS integrated circuits.

Layer / Technology	**Rs [Ω/□] (0.18um technology)**
Metal1	0.01
N+ Diffusion	50 ~ 65
P+ Diffusion	155 ~ 173
Polysilicon	7.5 ~ 18
n-channel transistor	R_{sn}
p-channel transistor	$R_{sp} \approx 2.5R_{sn}$

Table 3-15 Typical R_s Of MOS Layers



Square corner rectangle is counted as ½ square of resistance

Figure 3-16 MOS Wire Segment

For polysilicon with R_s = 7.5 Ω/□, the total resistance will be :-

R = 7.5 * (9/3 + ½(3/3) + 6/3)

= **41.25 Ω**

3.3.2.2 Determination of Load Capacitance

Capacitance are found between two conducting layers separated by a dielectric ($Si0_2$) and is given as :-

$$C = (\varepsilon_o * \varepsilon_r * A) / D \quad \text{Farads}$$

where

ε_o = 8.85 x 10^{-12} F/m (permittivity of free space)

ε_r = relative permittivity of $Si0_2$ ~ 4.0

A = Area of overlapping conducting layer

Alternatively, capacitance can also be determined using :-

C = relative area * relative capacitance * □C_g

where

relative area = ratio of actual area (A) to area of a minimum sizes gate

□C_g = gate to channel capacitance of a minimum size MOS transistor of length 2λ and width 2λ (where 2λ is the process technology's minimum feature size)

The advantage of using $\square C_g$ to determine capacitance is that the absolute value of the capacitance can be easily obtained by substituting the standard value of $\square C_g$ of the target process technology into the expression.

Material	Capacitance (aF / μm^2)	Relative Capacitance
Gate to channel	100	1.00
Diffusion to substrate	40	0.40
Polysilicon to substrate	100	1.00
Metal1 to substrate	30	0.30
Metal1 to polysilicon	65	0.65
Metal2 to substrate	15	0.15
Metal2 to polysilicon	18	0.18
Metal2 to Metal1	35	0.35

Note : aF = 1 x 10^{-18} F

Table 3-17 Typical Area Capacitance Values For 0.18μm Technology

(i) Determination of $\square C_g$ for 0.18μm process technology

λ of a MOS transistor using 0.18μm process technology is 0.09μm. Hence, a minimum size MOS transistor of length 2λ and width 2λ for 0.18μm process technology will have a gate area of :-

$$\text{Gate area} = 2\lambda * 2\lambda$$
$$= 0.18 * 0.18 = 0.0324\ \mu m^2$$

For a gate-to-channel capacitance of 100 aF/μm^2, the standard value of $\square C_g$ will be :-

$$\text{Standard value of } \square C_g = 3.24*10^{-2} * 100\ aF/\mu m^2$$
$$= 3.24\ aF$$
$$= 3.24 * 10^{-18}\ F$$

(ii) Multi-layer Capacitance Calculation

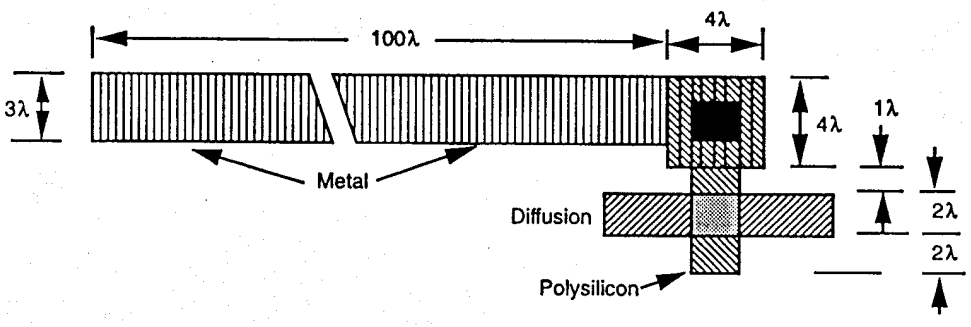
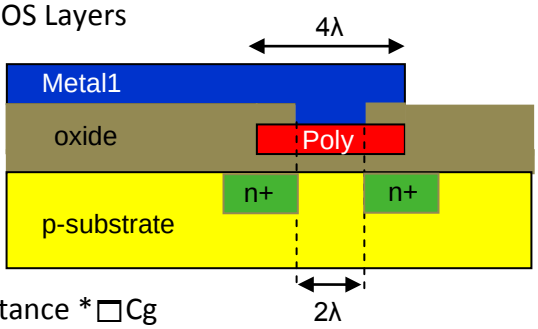


Figure 3-18 Capacitance Between MOS Layers



- Capacitance between Metal and substrate, C_m

  Area of metal = $100\lambda \times 3\lambda = 300\lambda^2$
  Relative area = $300\lambda^2/4\lambda^2 = 75$

  Thus,

  C_m = relative area * relative capacitance * $\square C_g$
  = 75 * 0.3 * $\square C_g$
  = 22.5 $\square C_g$

- Capacitance between Polysilicon and substrate, C_p

  Area of Polysilicon = $(4\lambda \times 4\lambda) + (1\lambda \times 2\lambda) + (2\lambda \times 2\lambda) = 22\lambda^2$
  Relative area = $22\lambda^2 / 4\lambda^2 = 5.5$

  Thus,

  $C_p = 5.5 * 1 * \square C_g = 5.5 \square C_g$

- Capacitance between gate and channel (diffusion and polysilicon overlap region)

  $C_g = 4\lambda^2 / 4\lambda^2 * 1 * \square C_g = 1 \square C_g$

  Thus, the total capacitance is :-

  $C_T = C_m + C_p + C_g$
  = $(22.5 + 5.5 + 1) \square C_g$
  = **$29 \square C_g$**

Depending on the technology used, the actual capacitance value will vary. For 0.18μm technology,

C_T = 29 x 3.24 aF
= 93.96 aF

3.3.3 Delay Calculations

3.3.3.1 τ-Model

In many occasions, a rough timing analysis of circuit is needed before getting into its layout. A simplified model, called the τ-model, can provide such estimation. This model is often used to analyze circuit involving a large capacitance load. In such a case, the load dominates, such that the diffusion and wiring capacitance can be ignored. In circumstances when this condition cannot be satisfied, the τ-model provides a tool for the designer to gain some insight of the timing relationship in the circuit.

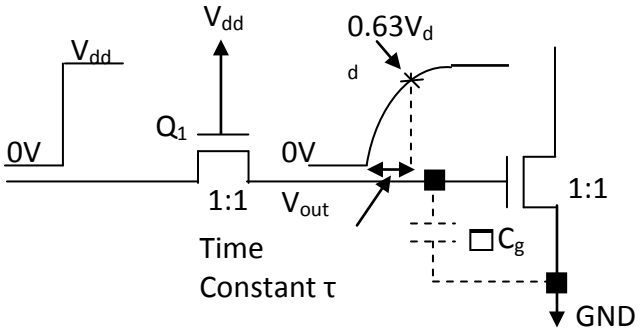


Figure 3-19 Model for Derivation of τ

Consider a load capacitance of 1C_g being charged through a minimum-sized NMOS transistor to 63% of its final value, as shown in Figure 3-19. The time constant,

$$\tau = RC$$

or

$$\tau = R_s\{L/W\} \times 1\ C_g$$

$$\mathbf{\tau = R_{sn} \times C_g}$$

Since τ is calculated in terms of R_{sn} and C_g, which are fixed for each technology, thus for each process technology τ is a constant.

3.3.3.2 Propagation Delay

Propagation delay, symbolized by t_p, is the time required for a digital signal to travel from the input of a logic gate to the output. The average time needed for the output to respond to a change in the input logic state is defined as :

$$t_p = \frac{t_{pHL} + t_{pLH}}{2}$$

where t_{pHL} = propagation delay for a *High* to *Low* transition at the output
t_{pLH} = propagation delay for a *Low* to *High* transition at the output

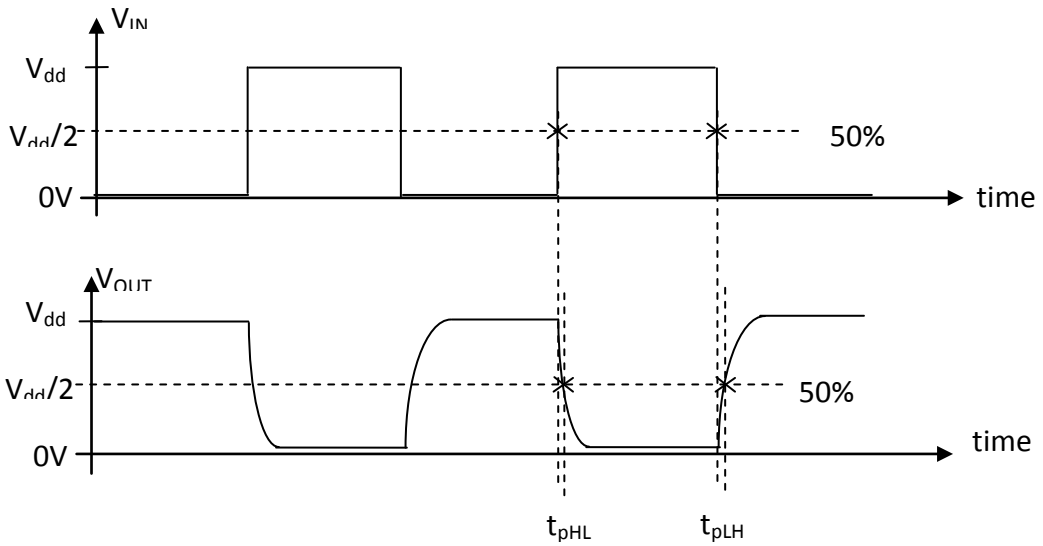


Figure 3-20 Propagation Delay

The timings are measured from the reference voltage of $V_{dd}/2$ at the input to the reference voltage of $V_{dd}/2$ at the output.

Assuming ideal input signal, the high to low propagation delay, t_{pLH}, can be derived by considering the time taken to charge the capacitance, C_L, from 0V to $V_{dd}/2$ and the low to high propagation delay, t_{pHL}, be the time taken to discharge C_L, from V_{dd} to $V_{dd}/2$. Since the transition point of is $V_{dd}/2$, which is close to $0.63V_{dd}$, *it is a common practice to use propagation delay and time constant interchangeably*.

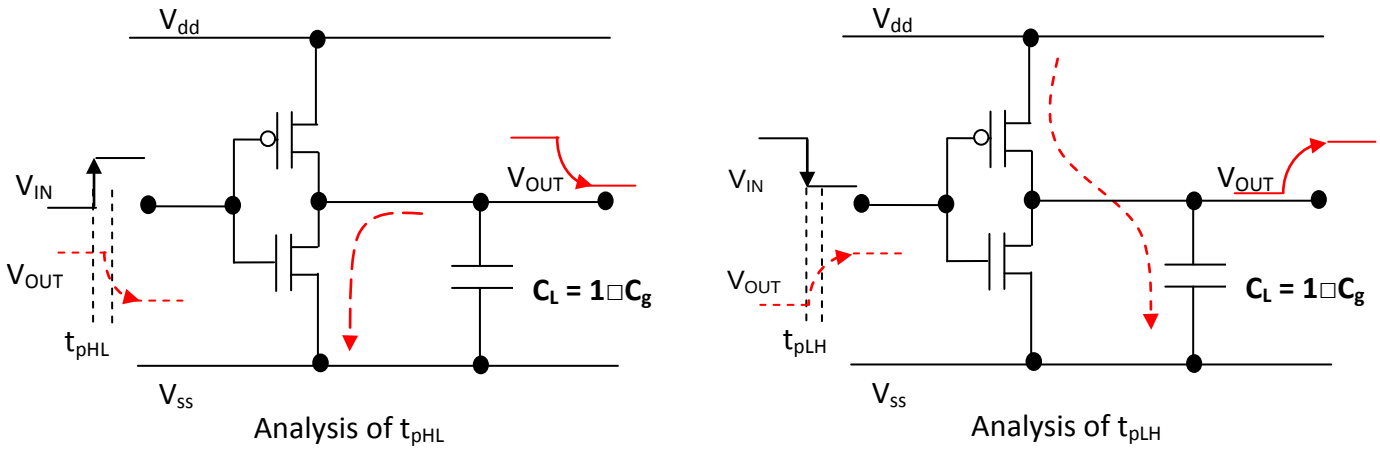


Figure 3-21 t_{pHL} and t_{pLH} Analysis

For a CMOS inverter with transistors of minimum size, that is L:W = 1:1, driving a load of $1\square C_g$, the

$$t_{pLH} \approx R_p C$$

$$t_{pHL} \approx R_n C$$

As the mobility of electrons, μ_n, is ~2.5 times the mobility of holes, μ_p; for PMOS and NMOS of the same W & L ratio, the PMOS transistor will have about 2.5 times lesser

current flowing through it under the same operating conditions, resulting in $R_p \approx 2.5R_n$. Hence,

$$t_{pLH} \approx R_p C$$
$$\approx 2.5 R_{sn} * \square C_g$$
$$\approx 2.5 \tau$$

$$t_{pHL} \approx R_n C$$
$$\approx R_{sn} * \square C_g$$
$$\approx \tau$$

Thus, $t_p = 1.75 \tau$.

To obtain a $t_{pLH} \approx t_{pHL,}$ it is necessary to adjust $W_p \approx 2.5W_n$ so that current passing through the PMOS is about the same as that flowing through NMOS transistor.

Typical electron mobility for Si at room temperature (300 K) is 1400 $cm^2/ (V \cdot s)$ and the hole mobility is around 450 $cm^2/ (V \cdot s)$.

3.3.3.3 Rise and Fall Time

Rise time, t_{rise}, is the time taken to charge V_{out} from 10% to 90% of V_{dd} while fall time, t_{fall}, is the time taken to discharge V_{out} from 90% to 10% of V_{dd}.

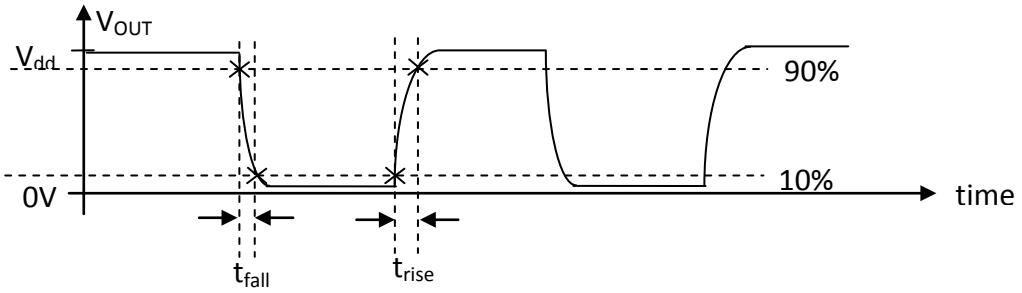


Figure 3-22 Rise and Fall Time

Taking a simplified view where a CMOS inverter charges towards V_{dd} and the p-transistor stayed in saturation while the n-transistor is fully off for the entire charging period of the load capacitor, C_L,

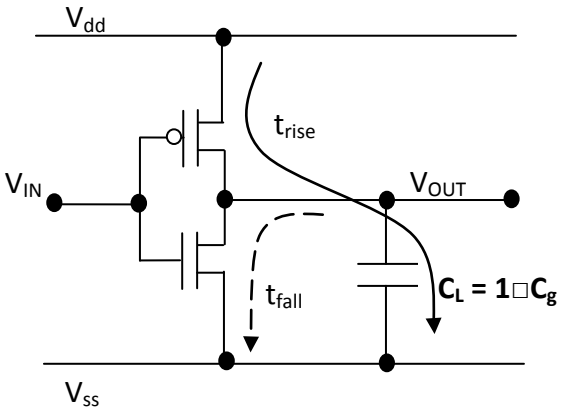


Figure 3-23 Rise and Fall time Analysis

Rise time can be estimated to be :

$$V_{out} = \frac{I_{ds(sat)}\, t}{C_L}$$

and $$I_{ds(sat)} = \frac{K_p}{2}\left(\frac{W_p}{L_p}\right)\left(\left|V_{gs}\right| - \left|V_{tp}\right|\right)^2$$

thus,

$$t_{rise} \approx 3.7 * C_L / (\beta_p V_{dd}) \quad \text{-- (1)}$$

where

$\beta_p = \mu_p \varepsilon_{ox}/t_{ox}\, [W_p/L_p]$
C_L = load capacitance

Similar reasoning can be applied to the discharge of C_L, through the n-transistor, to V_{ss}. Assume the p-transistor is off and n-transistor is on during this period, thus

Fall time can be estimated to be [I=C.dv/dt]:

$$t_{fall} \approx 3.7 * C_L / \beta_n V_{dd}) \quad \text{-- (2)}$$

where

$\beta_n = \mu_n \varepsilon_{ox}/t_{ox}\, [W_n/L_n]$

The ratio of expression (1) and (2) with $\mu_n \approx 2.5\mu_p$, results in $\beta_n \approx 2.5\beta_p$. That is, when a gate uses transistors of the same size (L & W) for both NMOS and PMOS, it will experience a $\boldsymbol{t_{rise} \approx 2.5\, t_{fall}}$. In order to achieve symmetrical operation using minimum channel length, make $W_p \approx 2.5W_n$.

3.3.3.4 Maximum Switching Frequency

The sum of transient time (t_{rise} + t_{fall}) represents the minimum time needed for a gate to undergo a complete cycle, ie. for the output to change from a logic 1 to logic 0 voltage and then back to logic 1 value. Thus, maximum switching frequency can be defined as :-

$$f_{max} = \frac{1}{t_{rise} + t_{fall}}$$

This frequency also represents the maximum rate of data transfer for the gate. In system design, the working value of f_{max} is set by the slowest gate or datapath element in the network.

3.3.3.5 CMOS Inverter Characteristics Summary

(i) Power Consumption

When the CMOS inverter is in a steady state, regardless of the output state, only half of the circuit is conducting and there is no closed path between V_{dd} and V_{ss}. Since no current flows, the steady state power is 0W. That is, there is no static power dissipation.

When the CMOS inverter is in transition from one state to another, there is a window where both the transistors are turned on, so a current flows between V_{dd} and V_{ss} through them. This is called the short-circuit current, which contributes to the dynamic power dissipation in the CMOS inverter.

A well-designed circuit operating with well-behaved signals of reasonably fast rise and fall time would go through the transition quickly. This dynamic power dissipation is thus less significant compared to the power caused by current that flows through the transistors to charge or discharge a load capacitance, C_L.

(ii) Ratio-less Device

V_{OUT} of the CMOS inverter is able to give a full voltage swing from V_{dd} to V_{ss} regardless of the relative dimensions of the PMOS and NMOS transistors. Thus CMOS devices are sometimes known as ratio-less devices.

(iii) Timing Performance

Ratios introduced in CMOS gates are for the purpose of improving the timing aspects of the gates. Timing performance is affected by the size of the transistors and the number of transistor turned ON, to achieve the desired output state.

3.3.4 CMOS NAND Gate

For CMOS, the pull-up PMOS transistors and the pull-down NMOS transistors form a complementary set. The pull-up transistors act as loads and are complementary in connections to the pull-down NMOS devices. The pull-down devices, in turn, determine the output function.

When the pull-down devices are connected in series, an ANDed function is achieved and when they are connected in parallel, an ORed function is achieved. The connection shown in Figure 3-24 gives a NAND output.

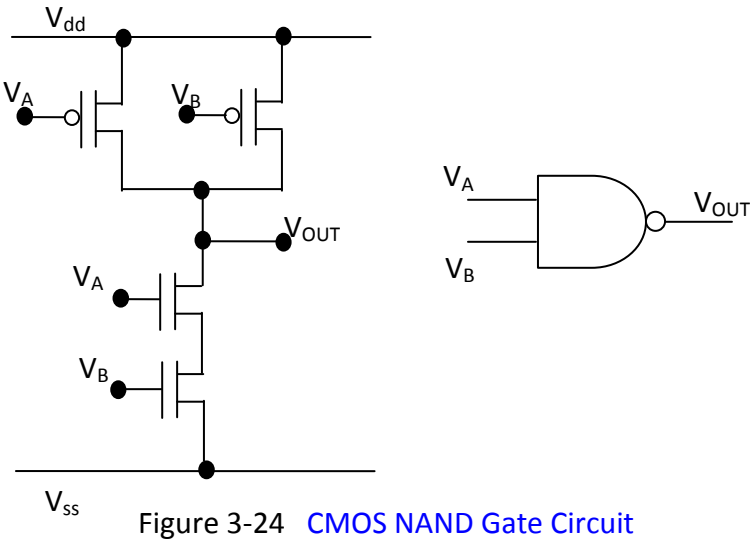


Figure 3-24 CMOS NAND Gate Circuit

3.3.5 CMOS NOR Gate

In the case of a NOR gate, the pull-down transistors are connected in parallel. Hence, the pull-up transistors will be connected in series. With the pull-down transistors in parallel, the effect is that of an ORed function. The connection shown in Figure 3-25 gives a NOR output.

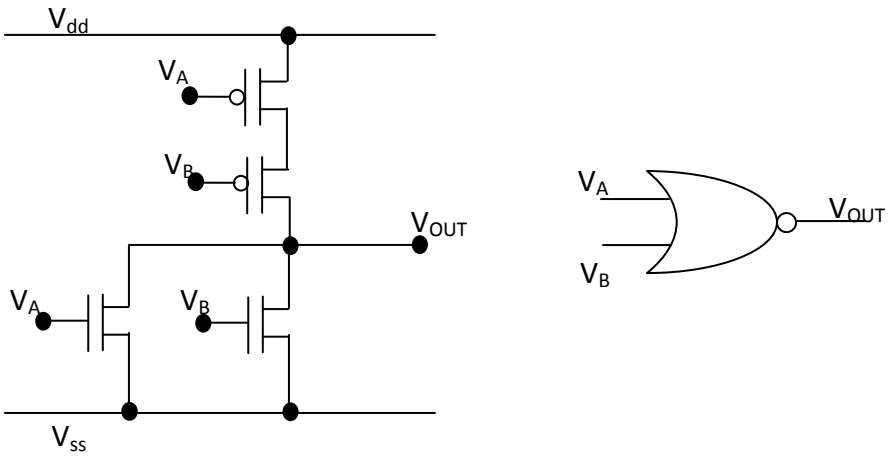


Figure 3-25 CMOS NOR Gate Circuit

3.3.6 NAND and NOR Gate Delays

(i) Low-to-High Delay (t_{pLH})

The t_{pLH} of the NAND gate depends on the number of p-transistors that are connected in parallel (number of inputs) while that of a NOR gate depends on the number of p-transistors that are connected in series.

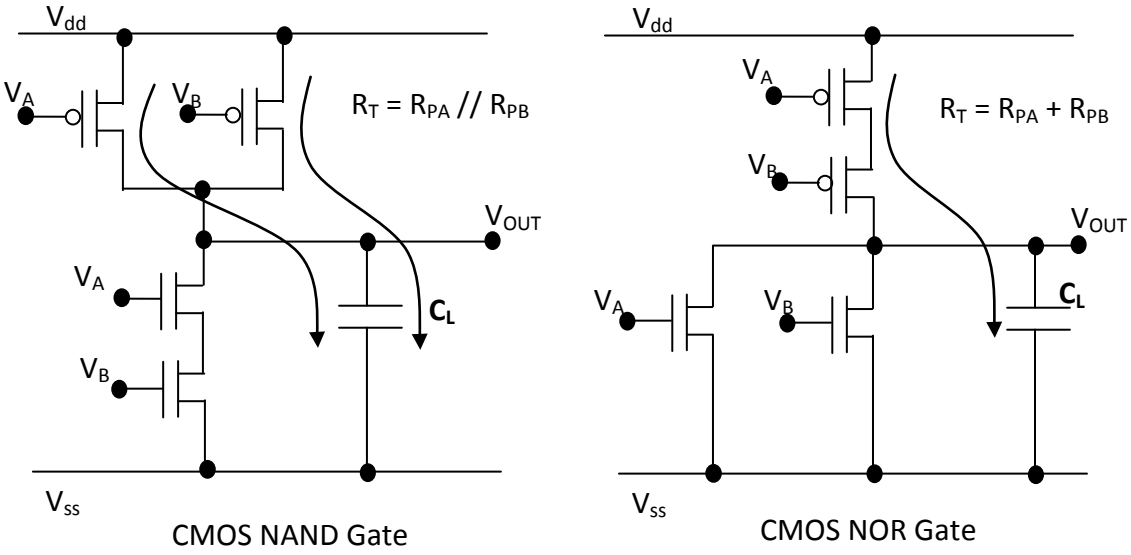


Figure 3-26 Direction of current flow during t_{pLH}

Figure 3-26 shows a 2-input NAND gate and a 2-input NOR gate. Assuming that all the transistors are minimum-sized transistors, the total resistance experienced by the NAND gate during the low-to-high transition is R_{sp} [L/W] or less while that experienced by the NOR gate is 2 * R_{sp} [L/W]. Hence, the t_{pLH} of the NOR gate will be at least 2 times greater than that of the NAND gate.

(ii) High-to-Low Delay (t_{pHL})

The t_{pHL} of the NAND gate depends on the number of n-transistors that are connected in series (number of inputs) while that of a NOR gate depends on the number of n-transistors that are ON.

As shown in Figure 3-27, assuming that all the transistors are minimum-sized transistors, the total resistance experienced by the NAND gate during the high-to-low transition is 2 * R_{sn} [L/W]. For NOR gate, the t_{pHL} resistance depends on the number of n-transistors that are ON, which is R_{sn}[L/W] or less.

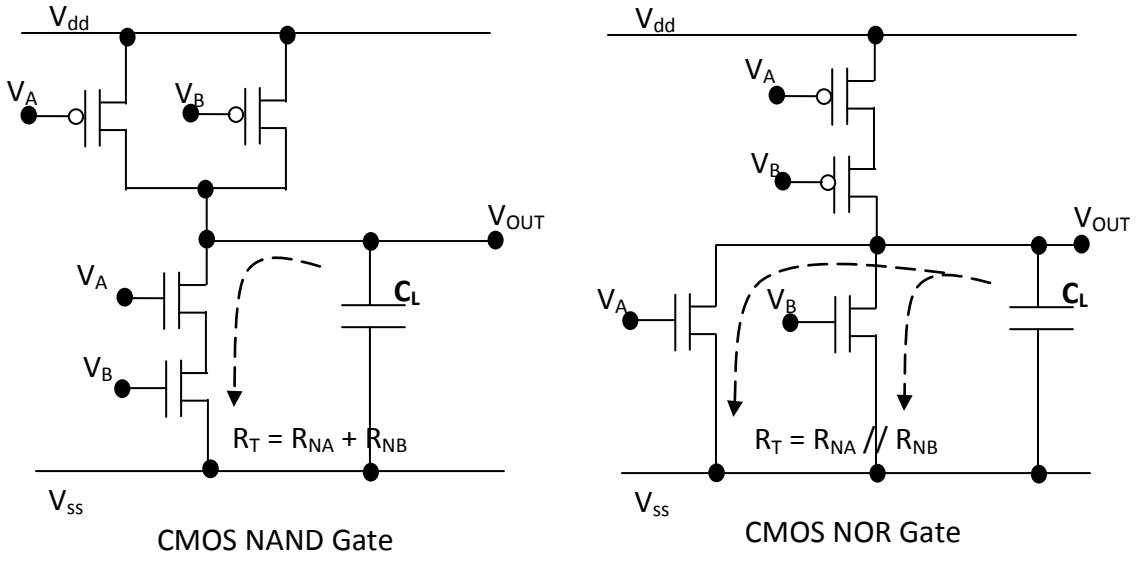


Figure 3-27 Direction of current flow during t_{pHL}

3.3.7 CMOS Complex Implementation

For the circuit shown in Figure 3-28, pull-up PMOS transistors and the pull-down NMOS transistors form a complementary set. The pull-up transistors act as loads and are complementary in connections to pull-down NMOS devices. The pull-down devices, in turn, determine the output function.

Input *A* and *B* are in series, hence they are *AND* together. Similarly, input C and D are also ANDed together. As the *AB* branch and *CD* branch are in parallel, the two branches are hence *OR*ed together. Thus, the output function of the complex implementation is :-

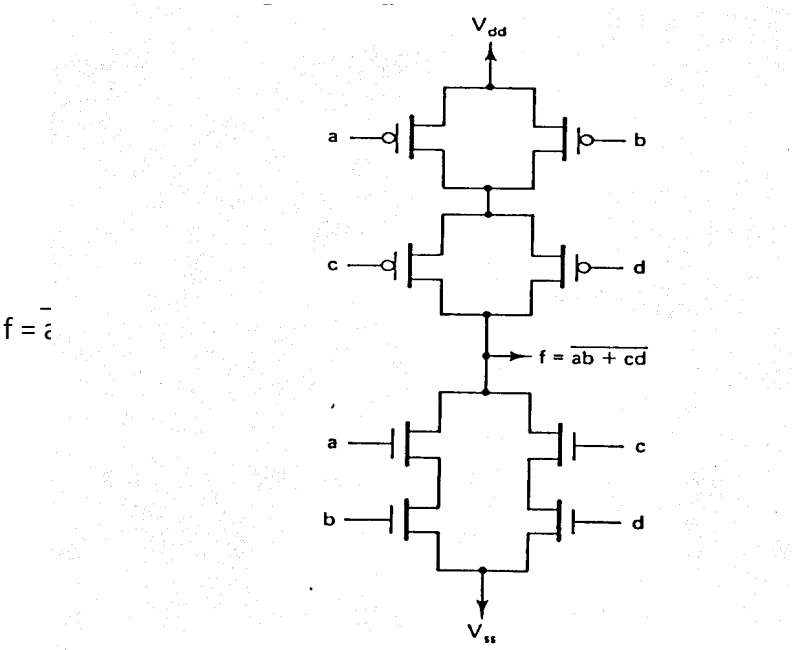


Figure 3-28 Complex Gate CMOS Implementation

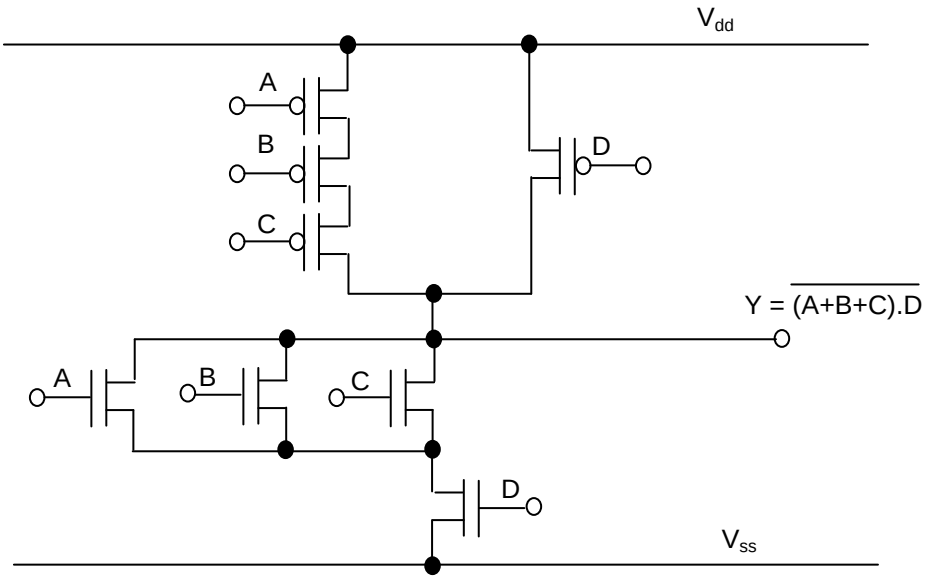


Figure 3-29 Complex Gate CMOS Implementation

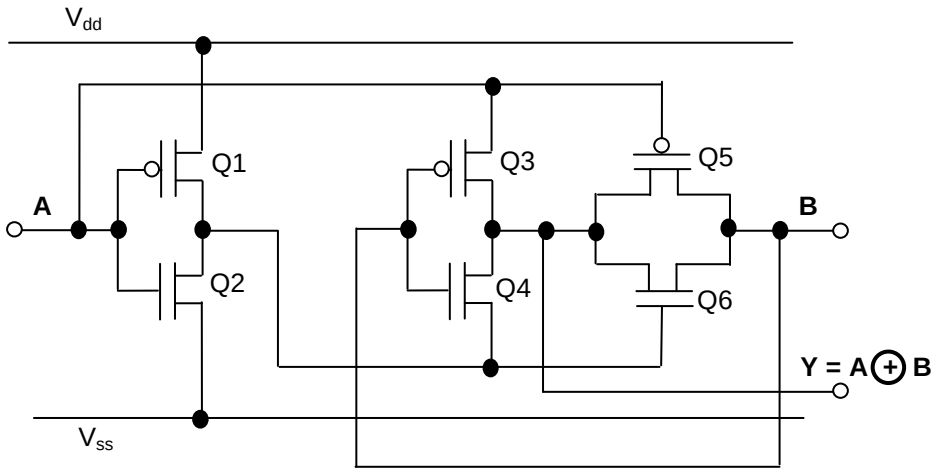


Figure 3-30 Complex Gate Implementation (XOR)

3.4 DYNAMIC MOS CIRCUIT

Dynamic MOS circuits work on the principle of turning transistors ON or OFF under clock control. Basically, a p-transistor is used for the non-time-critical precharging of the output line so that the output capacitance is charged to V_{dd} during the OFF period of the clock signal ϕ. During the same period (next half cycle), the inputs are applied to the n-block and the state of the logic is then evaluated.

Such circuit can operates at a higher speed compared to the static circuits as input capacitance of dynamic gate is smaller. With precharging (discharging) of the output line, the amount of time needed for the output to change from one state to another is also reduced significantly.

3.4.1 P-E (Precharge-Evaluation) Logic

This logic works on one clock signal and two separate devices, the p-transistor and the n-transistor. Figure 3-31(a) shows a PE logic using n-channel transistor as the evaluation transistors while Figure 3-31(b) shows a PE logic using p-channel transistors as the evaluation transistors.

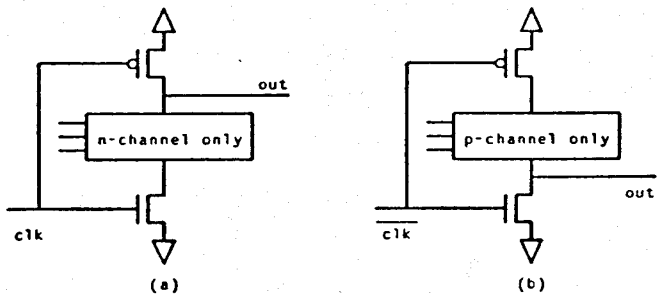


Figure 3-31 P-E Logic Gate

There are two phases to their operation, namely the precharge phase and the evaluation phase. With reference to Figure 3-31(a),

Precharge Phase

Precharging is done when the *clk* signal is Low. During which the n-transistor is OFF and the p-transistor is ON. This causes the output to be precharged HIGH. For logic using n-channel evaluation gates, its output must be connected only to p-channel based gates for correct precharging.

Evaluation Phase

Evaluation of logic is done when the *clk* signal goes High. When the *clk* signal goes High, the p-transistor turns OFF while the n-transistor goes ON. This provides a path to V_{ss}. Depending on the logic of the evaluation transistors, the charges on the output line may either be discharged or remains.

When using a single phase clock, dynamic logic structure of the same type cannot be cascaded since, owing to circuit delays, an incorrect input to the next stage may be present when evaluation begins. One remedy is to alternate the structure.

3.4.2 Domino Logic

This logic consists of two basic elements, namely the PE gate and an inverter at the output. The inverter at the output acts as a buffer to increase the driving capability of the gate. This logic supports single phase clock with cascading of logic structures of the same type. That is, it enables its output to be connected to similar type of (n-type or p-type) gates, yet achieve proper precharging of the output lines.

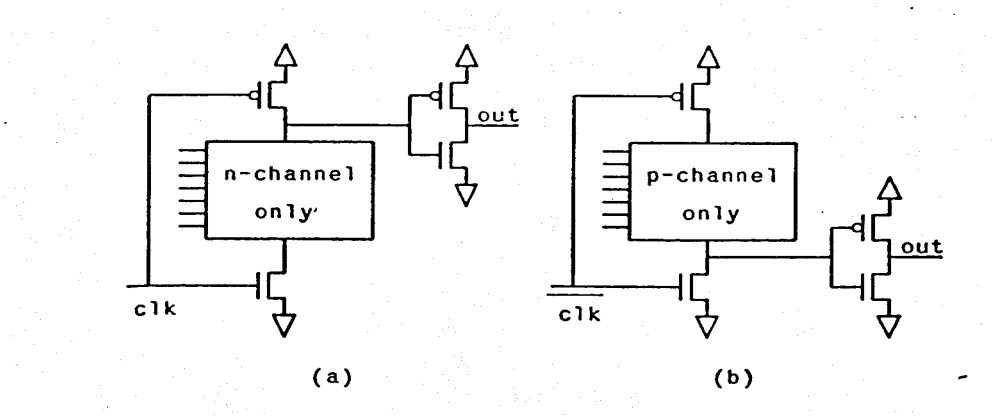


Figure 3-32 Domino Logic Gate

3.5 MOS CIRCUIT PERFORMANCE

The electrical behavior of a MOS transistor is primarily determined by its gain factor, β, its threshold voltage V_T and its body factor, K. Generally, the values of these parameters are largely dependent on the width, W and length L of a transistor. The influence of these dependencies is particularly significant in submicron and deep submicron MOS processes.

3.5.1 Scaling of MOS Devices

Over the past decades, scaling of CMOS devices has yield tremendous gains in chip performance and functionality. Scaling refers to the systematic reduction of transistor dimensions from one generation to the next, reducing the parasitic capacitances and also the carrier transit times in the devices. One result of this scaling is an improvement in circuit speed; this improvement is quite dramatic in the case of MOS-based circuits. Another result of scaling is that it narrows the performance gap between CMOS and logic gates based on bipolar transistors. Finally, reduction of transistor dimensions improves the packing density of CMOS.

3.5.1.1 Full Scaling of MOS devices

There are many possible schemes for scaling CMOS. Of these, the most obvious approach is "full scaling" which involves scaling all dimensions and voltages by the same factor, 1/s, where s is greater than one. The key dimensions to scale are gate dimensions and oxide thickness; the key voltage to scale is the supply voltage. The motivation for scaling the voltage is that this will leave the electric field intensities unchanged, thus avoiding breakdown effects.

3.5.1.2 Constant Voltage Scaling of MOS devices

Another approach is called "constant voltage scaling". This involves scaling all dimensions by the factor, 1/s while keeping all voltages constant. Although constant voltage scaling increases the electric filed intensities in the devices, its advantages includes allowing the same supply voltage from one generation to the next and better speed performance. However, the disadvantage is that the electric field increases as the minimum feature length is reduced. This leads to velocity saturation, mobility degradation, increased leakage currents and lower breakdown voltages.

Parameter	Symbol	Constant Field Scaling	Constant Voltage Scaling	Constant Voltage Scaling with velocity saturation
Gate length	L	$1/\alpha$	$1/\alpha$	$1/\alpha$
Gate width	W	$1/\alpha$	$1/\alpha$	$1/\alpha$
Field	$\mathcal{E}$	1	α	α
Oxide thickness	t_{ox}	$1/\alpha$	$1/\alpha$	$1/\alpha$
Substrate doping	N_a	α^2	α^2	α^2
Gate capacitance	C_G	$1/\alpha$	$1/\alpha$	$1/\alpha$
Oxide capacitance	C_{ox}	α	α	α
Transit time	t_r	$1/\alpha^2$	$1/\alpha^2$	$1/\alpha$
Transit frequency	f_T	α	α^2	α
Voltage	V	$1/\alpha$	1	1
Current	I	$1/\alpha$	α	1
Power	P	$1/\alpha^2$	α	1
Power-delay	$P \triangle t$	$1/\alpha^3$	$1/\alpha$	$1/\alpha$

Figure 3-33 Comparison of the effect of scaling on MOSFET device parameters

3.5.2 Channel Length Modulation

Channel length modulation is the shortening of the length of the inverted channel region, beyond the pinch-off point, with increase in drain bias. As the drain voltage increases, the length of the channel shortens further, causing the channel resistance to decrease, hence an increase in current. The effect is more pronounced in transistors of small feature size.

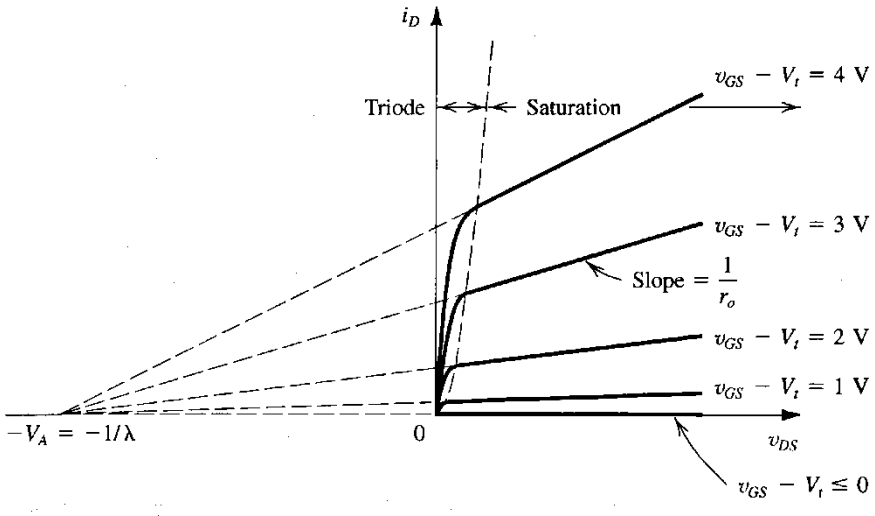


Figure 3-34 Effects of V_{ds} on I_{ds} in Saturation Region

The saturation current, I_{ds}, for devices experiencing channel length modulation is given as :-

$$I_{ds} = \{K_n/2\}[W/L](V_{gs}-V_t)^2(1 + \boldsymbol{\lambda} V_{ds})$$

where

λ is the empirical channel length modulation factor; λ~ 0.03/V

3.5.3 Threshold Voltage

One of the parameter that determines the amount of drain current in a MOS transistor is the threshold voltage. It can be seen that higher the threshold voltage, lower will be the device current and hence slower will be the performance.

This voltage is not a constant but is affected by :-

(i) body effect.

(ii) hot carrier effect.

3.5.3.1 Body Effect

Normally, the substrate (B) of a NMOS device is biased to V_{ss} while substrate of a PMOS device is normally biased to V_{dd}. Hence, the back-bias voltage,

$$V_{sb} = V_s - V_{substrate} = 0$$

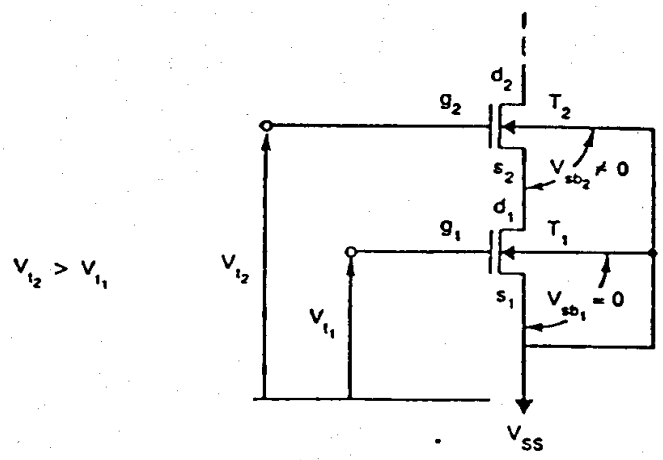


Figure 3-35 NMOS in Series Resulting in Different V_{sb} Biasing

When two devices are connected is series (e.g. NAND gate), their V_{sb} will be different, that is, $V_{sb1} = 0$ but $V_{sb2} \neq 0$. With $V_{sb2} \neq 0$, an increase in the thickness of the depletion region under the channel of T_2 will occur. The increased depletion layer requires additional charge. The channel charge therefore decreases if V_{gs} is held constant. The channel conductivity can only be maintained if V_{gs} is increased. That is, threshold voltages of T_2 to T_1 is such that $V_{t2} > V_{t1}$ so as to get same amount of current to flow through T_1 and T_2. In other words, the threshold voltage is related to the back-bias voltage, V_{sb} and T_2 is said to be experiencing body effect. It causes MOS transistors of the same type and dimensions to have different threshold voltages.

$$\boldsymbol{V_T = V_x + K\sqrt{V_{sb} + 2\emptyset_f}}$$

$$\boldsymbol{V_{T0} = V_x + K\sqrt{2\emptyset_f}}$$

where V_x = process-related constant threshold voltage

$V_T = V_T|_{V_{sb}=0V}$

K = process parameter, also known as body factor

V_{sb} = source-bulk (back-bias) voltage

$2\emptyset_f$ = band bending where inversion first occurs

If a device was biased without the bulk node connected then a change in operating point could take the device out of its saturation region and significantly change the circuit performance. The bulk voltage is thus a very important parameter in circuit applications and therefore it is best to connect the bulk to the device source connection.

3.5.3.2 Hot-Carrier Effect

Hot carrier effect is the trapping of high energy (hot) carriers in the gate oxide. Such carriers may be injected from the substrate or the channel into the gate oxide. This effect is more likely to occur when the drain to source (V_{ds}) voltage is large or when the substrate is heavily doped (excess carriers). These trapped electrons cause shifts in the threshold voltage and reliability problems in the IC chip.

Lightly doped drain (LDD) structure is used to suppress hot carrier effect, particularly in modern devices where the channel length is getting shorter (<2µm). Reduction in power supply voltage also reduces hot carrier effect.

In LDD structure, the presence of spacer nitride results in the source & drain regions being further away from the gate, effectively reduces the source-drain bias induced electric field, thus suppressing the hot carrier effect.

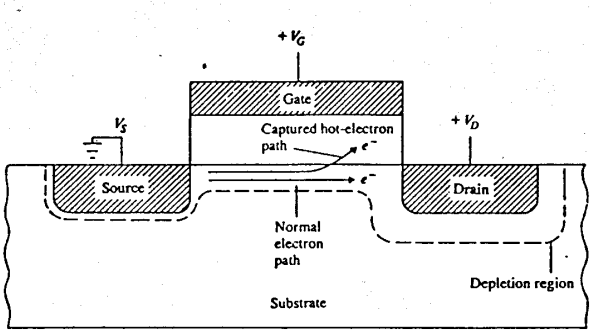


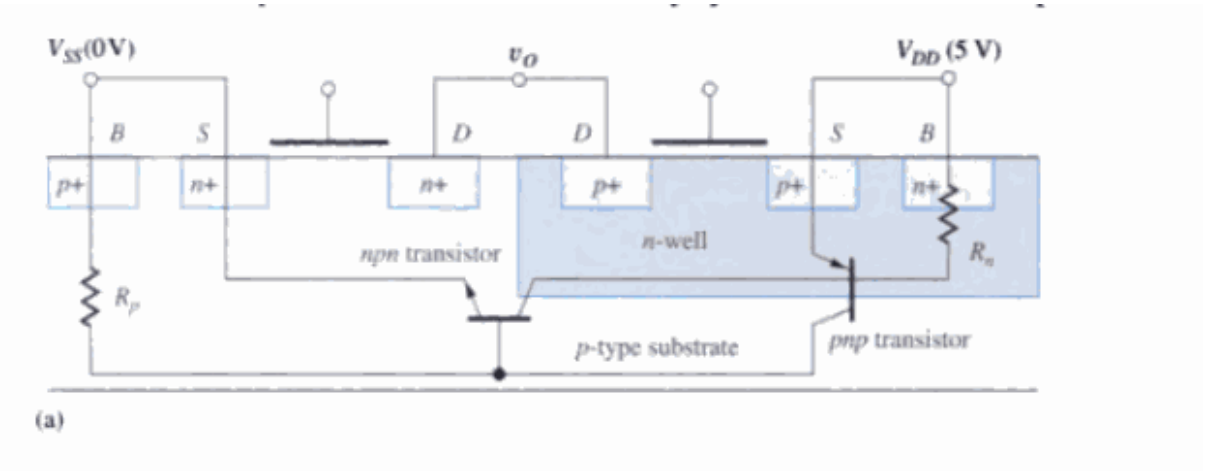
Figure 3-36 Channel Electrons Injected Into the Oxide

3.5.4 CMOS Latchup

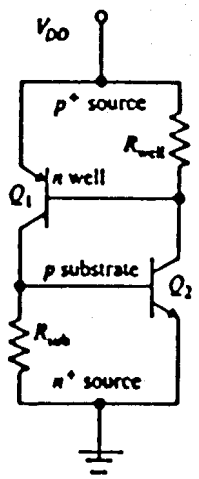
Latch up is the inadvertent creation of a low-impedance path between the power supply rails of a CMOS circuit. It occurs when the parasitic structure within the CMOS circuit is activated thus disrupting proper functioning of the part and possibly even leading to circuit destruction due to over-current. A power cycle is required to correct this situation.

Figure 3.37 shows the parasitic structure. Its silicon-controlled rectifier (SCR) equivalent is shown in Figure 3.38. In the structure, the collector of each BJT is connected to the Base of the other transistor in a positive feedback structure. The parasitic structure is normally dormant under normal operating condition. However, it can be activated when transients occur.

Transients may occur internally during power up or externally due to operating the device beyond its normal operating range. When these transients occur, transistors Q_1 and Q_2 may become forward biased and hence forming a closed loop. A high current can now flow between V_{dd} and V_{ss}, and destroy the device in the process.



A scenario of latch-up is when sufficient current flows through R_{sub}, this in turn causes Q_2 to turn ON. With Q_2 ON, current will be drawn through R_{well}. If sufficient voltage develops across R_{well}, Q_1 can then turn ON, causing more current to flow



R_{sub} is due to the p-substrate
R_{well} is due to the n-well

Figure 3-38 Latch-up Equivalent Circuit

3.5.4.1 Methods to reduce possibility of latch-up

Latch up in CMOS circuits can be reduced by :-

(i) increasing the distance between N-well and N+ source/drain diffusion of the NMOS device $\Rightarrow$ reduction of β_2.

(ii) using guard rings around the transistors as shown in Figure 3.39. For an n^+ guard ring, it provides a low resistance path to collect holes before they interfere with the operation of other circuits outside the guard ring, thus reduces parasitic resistance.

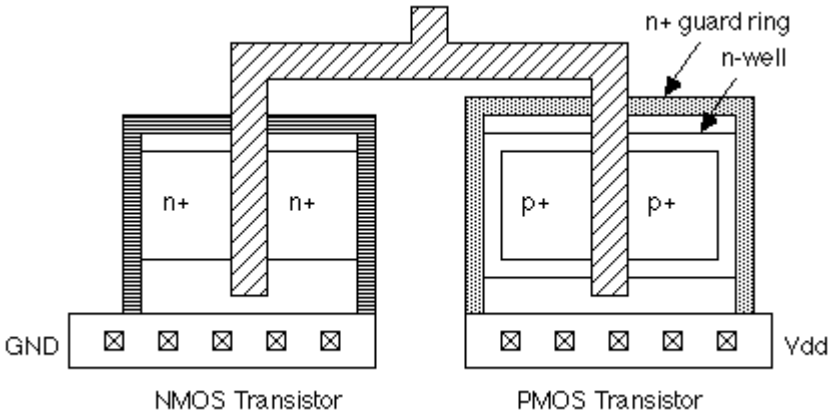
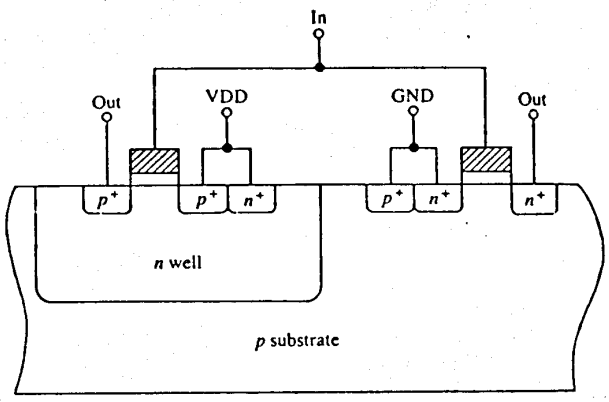


Figure 3-39 CMOS circuit with guard rings

(iii) placing a well tie and a substrate tie close to the supplies. These will in turn reduce the value of R_{sub} and R_{well} $\Rightarrow$ higher accidental current is required to cause latch-up



well tie substrate tie

Figure 3-40 CMOS structure with well tie and substrate tie

(iv) increase substrate doping level $\Rightarrow$ reduction in R_{sub}

3.6 BiCMOS DIGITAL DESIGN

BiCMOS technology combines the advantages of Bipolar and CMOS technology. It retains the low power consumption advantage of CMOS technology and the high drive capability of Bipolar technology.

The timing of such technology is also less sensitive to temperature variations. However, the process complexity is increased.

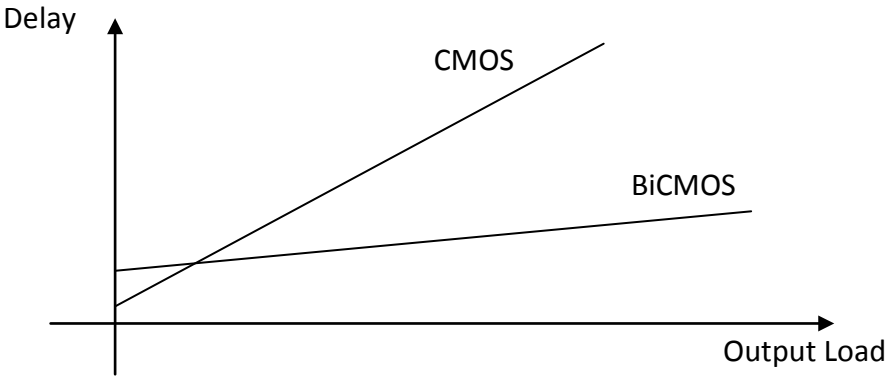


Figure 3-41 Delay vs Output Load

Figure 3-42 shows a BiCMOS inverter circuit and Figure 3-43 shows its cross sectional view on a wafer.

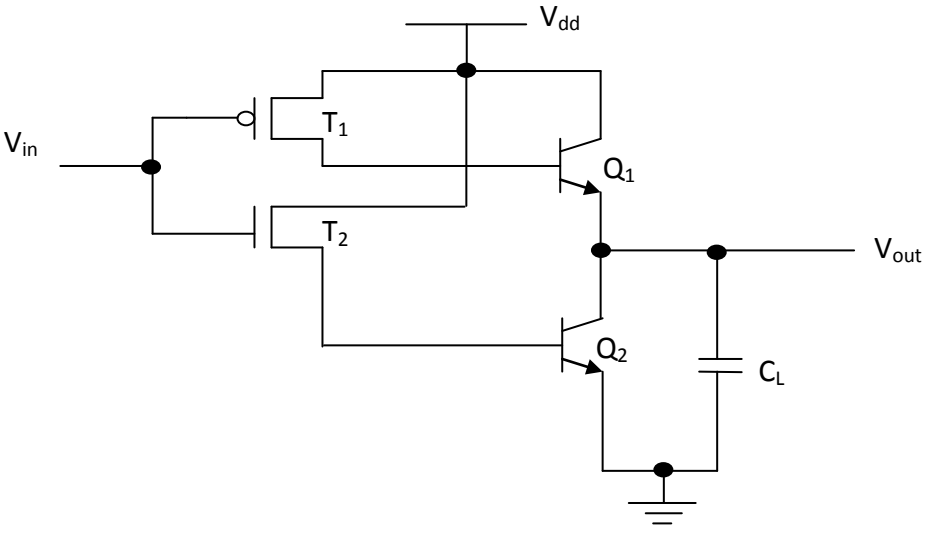


Figure 3-42 Example of a BiCMOS Inverter Circuit

Operation of the BiCMOS inverter

(i) When V_{in} = High,
- T_2 turns ON, T_1 is OFF
- With T_2 ON, Q_2 also comes ON thus allowing C_L to discharge to a low value and hence V_{out} = Low. As in bipolar technology, V_{OL} is in the range of 0.2V to 0.3V.

(ii) When V_{in} = Low
- T_1 turns ON, and T_2 turns OFF
- With T_1 ON, Q_1 also comes ON thus allowing C_L to charge to a high value of

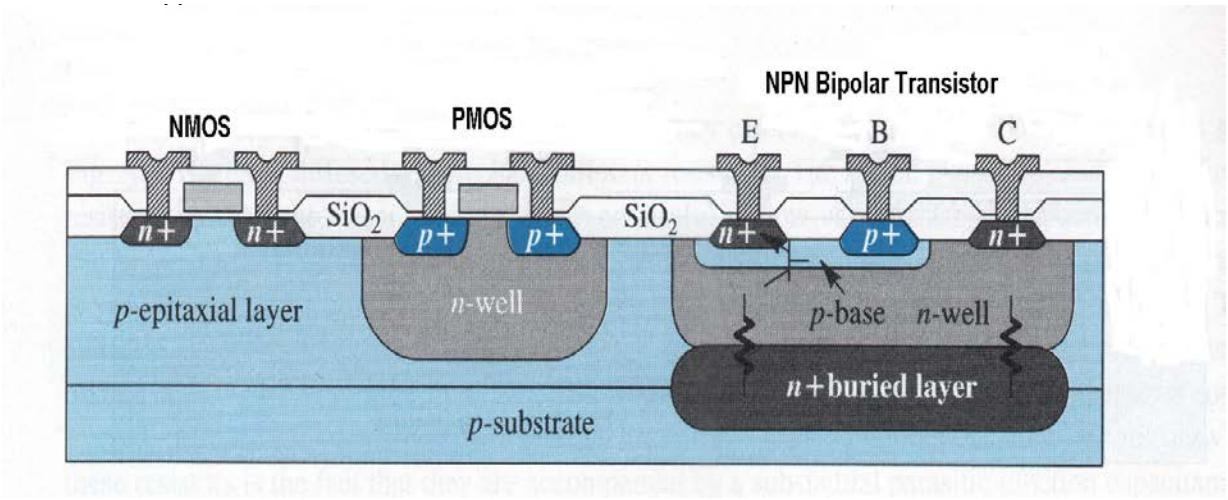


Figure 3-43 Cross-Sectional structure of Bi-CMOS Process Wafer

3.7 CMOS FABRICATION PROCESS

There are many processes available for the fabrication of CMOS circuitries.

3.7.1 N-well Process

The substrate material is of p-type. P-channel (PMOS) transistor is found inside the N-Well region. This method enables sharing of process steps with NMOS fabrication process.

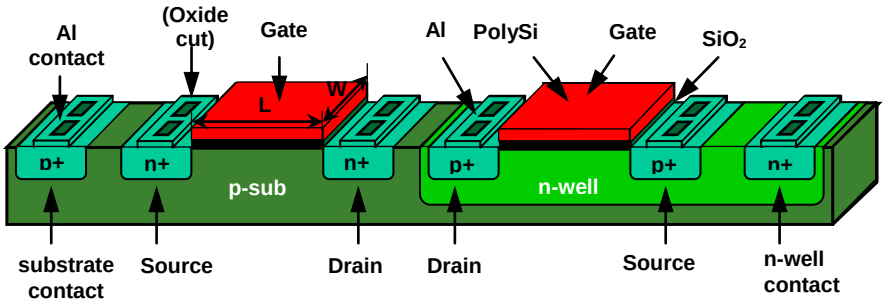


Figure 3-44 Cross-sectional structure of N-Well CMOS Inverter

3.7.2 P-well Process

The substrate is of n-type. For this process, n-channel (NMOS) transistors are found in the P-Well regions.

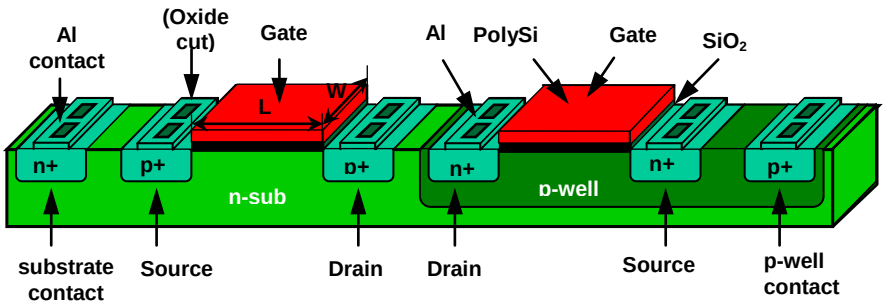


Figure 3-45 Cross-sectional structure of P-Well CMOS Inverter

3.7.3 Twin-tub Process

For this process, both P-well and N-well are formed on the substrate, thus the substrate can be of p-type or n-type. This process requires more processing steps but it provides better optimisation of performance.

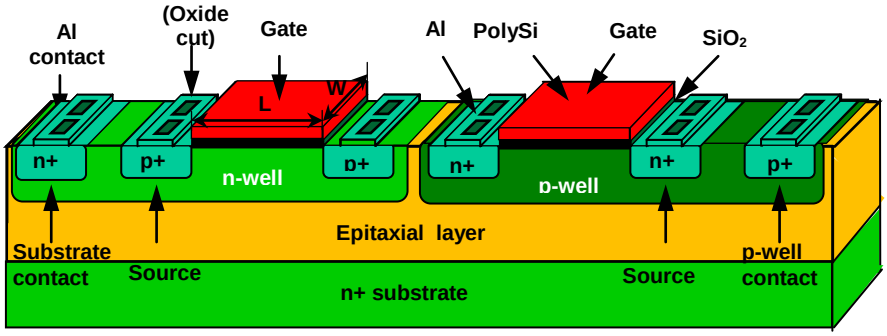


Figure 3-46 Cross-sectional structure of twin-tub CMOS Inverter

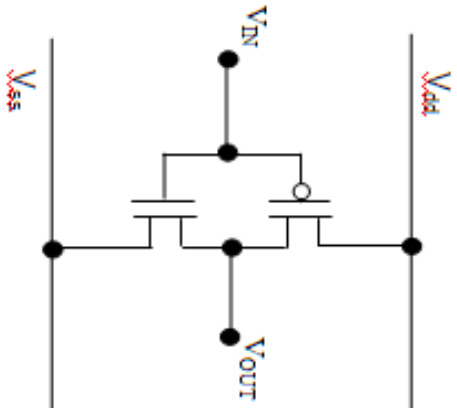
3.7.4 Silicon on Insulator (SOI) CMOS Technology

This technology uses sapphire, which is basically an insulator, as the substrate. The advantage of this technology is that problem such as latch-up or body effect is not an issue, hence it can operate at higher speed. However, the process is more expensive, poorer in heat conduction and relatively low in yield.

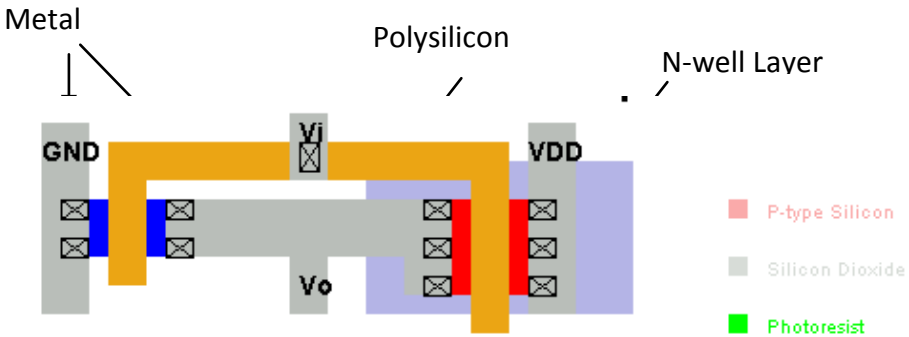
3.7.5 Fabrication of a CMOS Inverter

Figure 3-47 shows the layout diagram of a CMOS inverter using NWELL process.

(a) Transistor Level schematic



(b) Plan View



(c) Cross-section View

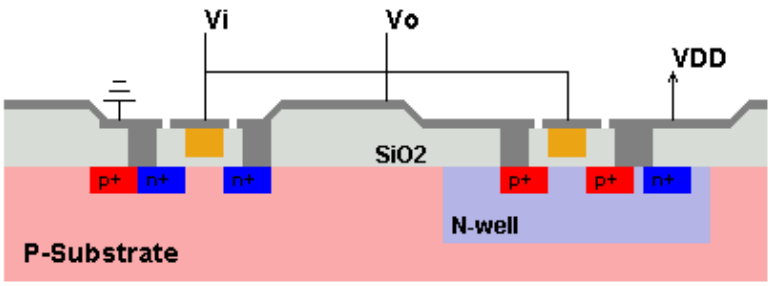


Figure 3-47 Translation of CMOS Inverter To Mask Design

Figure 3-48 shows the fabrication sequences of the CMOS inverter.

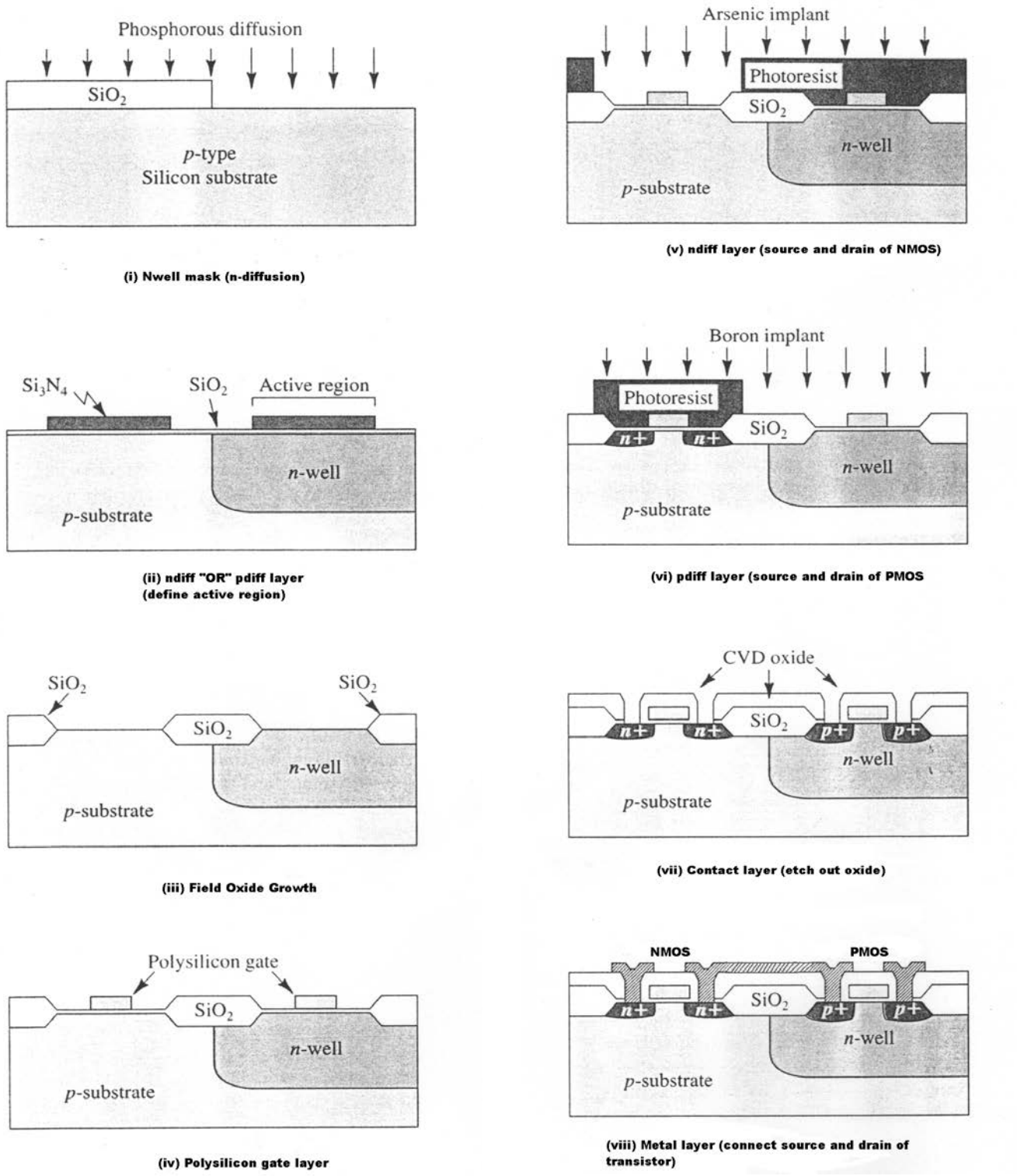


Figure 3-48 CMOS Inverter Fabrication Process Sequence

Review Questions

Question 1

(a) From the diagram given in Figure 3-49, identify the :-
(i) type of the MOS transistor.
(ii) gate, source and drain terminals.

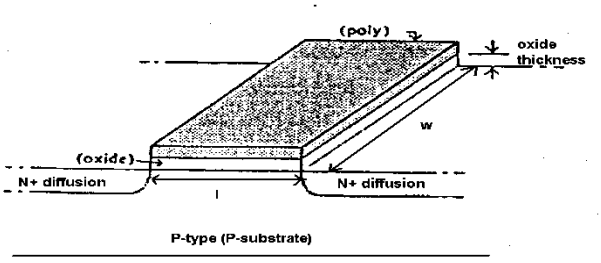
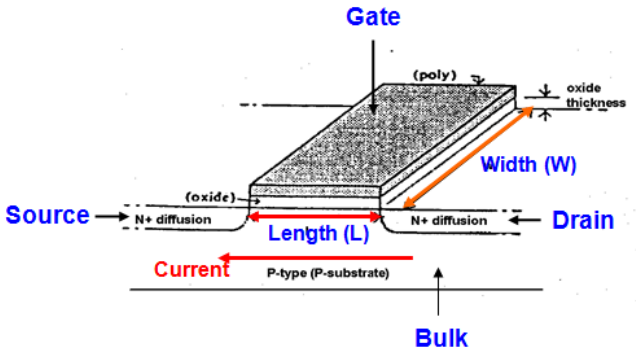


Figure 3-49 MOS transistor

(b) State the current equation of the MOS transistor under Linear Mode. Based on the current equation, explain the effects of having a long channel length or a wide channel width.

Solution to Q1

(a)



An NMOS transistor
(Both Source and Drain are of n+ diffusion)

(b) A MOS transistor is in linear mode when $V_{gs} \geq V_{tn}$ and $V_{ds} < (V_{gs} - V_{tn})$. The associated current equation is :-

$$I_{ds} = K_n[W/L][(V_{gs} - V_{tn})V_{ds} - V_{ds}^2/2]$$

$$\rightarrow \quad I_{ds} \ \alpha \ (W/L)$$

Thus, if length (L) increases, I_{ds} decreases.
if width(W) increases, I_{ds} increases.

Question 2

Explain what does each of the following statement means when referring to a MOS transistor.

(a) It is a symmetrical device.

(b) It is a voltage-controlled device.

(c) It is a unipolar device.

Solution to Q2

(a) A MOS transistor is a **symmetrical device** because source and drain terminal are interchangeable. It depends on the voltage potential at the terminal with respect to one another to determine its characteristic.

For NMOS transistor, the source terminal is always at a lower potential than the drain terminal and vice versa for PMOS transistor.

(b) A MOS transistor is a **voltage-controlled device** because a bias voltage needs to be applied to the gate terminal to turn 'ON' or 'OFF' the device.

A gate voltage of $|V_{gs}| > V_{th}$ is required to turn on transistor. For NMOS transistor, $V_{th} \approx 0.2V_{dd}$ while for PMOS transistor, $V_{th} \approx -0.2V_{dd}$.

(c) A MOS transistor is a **unipolar device** because conduction (current flow) depends only on the majority carriers. For an NMOS transistor, electrons are the majority carriers while for PMOS transistor, holes are the majority carriers.

Question 3

(a) Explain the three modes of operations of an NMOS enhancement transistor by :-

(i) drawing and labelling the various cross sectional structures.

(ii) providing the voltage conditions and current equations for each mode of operation.

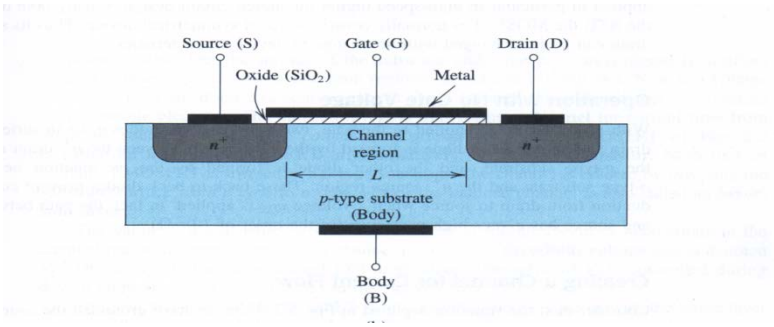
(b) Sketching the current (I_{ds}) versus voltage (V_{ds}) characteristic curve. Indicate the 3 regions of operation.

Solution to Q3

(a) A basic enhancement mode MOSFET does not have a channel established at fabrication thus it is a normally OFF device. The 3 modes of operation are :-

Cut-off Mode

When the voltage applied to the gate of the transistor is such that $V_{gs} < V_{th}$, where $V_{th} \approx 0.2V_{dd}$, no conducting channel is induced resulting in the path resistance between the source and drain being very high and hence $I_{ds} = 0$.

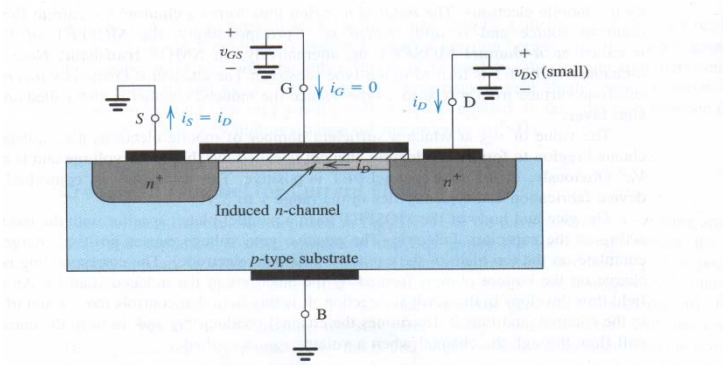


$_s < V_{tn}$, $I_{ds} = 0$

NMOS transistor in Cutoff Mode Operation

Triode/Linear Mode

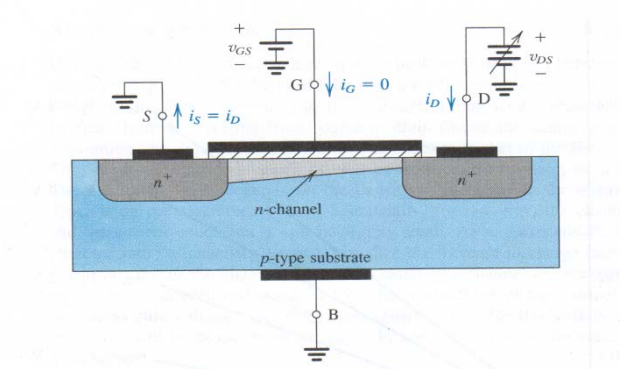
When $V_{gs} \geq V_{th}$ and $V_{ds} < (V_{gs} - V_{th})$, electrons (carriers) accumulated near the surface under the gate region starts to gather to create a conducting channel between the source and the drain. With the channel, current can now flow from the drain to the source (opposite to electron flow).



NMOS transistor in Triode/linear Mode

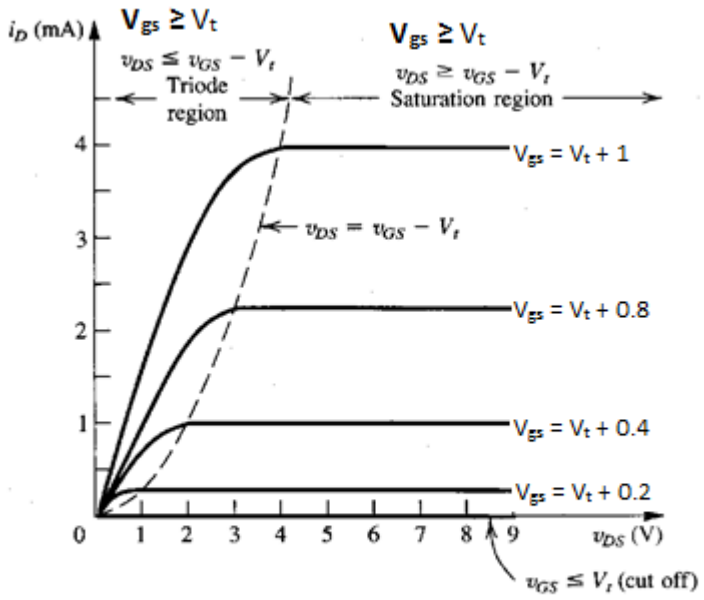
Saturation Mode

When $V_{gs} \geq V_{th}$ and $V_{ds} \geq (V_{gs} - V_{th})$, the channel will become tapered, as shown in the figure below, because the effective voltage on the channel decreases towards the drain end. The current I_{ds} is also affected. The I_{ds} under this condition is given as :-



NMOS Transistor in Saturation Mode

(b) NMOS Current (I_{ds}) versus voltage (V_{ds}) characteristic curve



NMOS VI Characteristics Curve

Question 4

For an NMOS enhancement transistor, with V_t = 0.4V, V_{gs} = 1V, L = 0.18μm, W = 5μm, $K_n = \mu_n \varepsilon_o \varepsilon_r / t_{ox}$ = 100μA/V^2 = 100μF/V-sec, calculate:-

(a) its current when V_{ds} = 1.0V.

(b) its current when V_{ds} = 0.5V.

Solution to Q4

(a) V_{ds} = 1.0 V

$V_{gs} - V_t = 1 - 0.4 = 0.6$ V

Since

$V_{ds} > (V_{gs} - V_{th})$ ➔ transistor is in Saturation Mode.

Thus, $I_{ds} = \{K_n/2\}[W/L](V_{gs} - V_{tn})^2$

$= 100*10^{-6}/2 * (5/0.18) * (0.6)^2$

$=$ 0.5 mA

(b) V_{ds} = 0.5 V

$V_{gs} - V_{th} = 1 - 0.4 = 0.6$ V

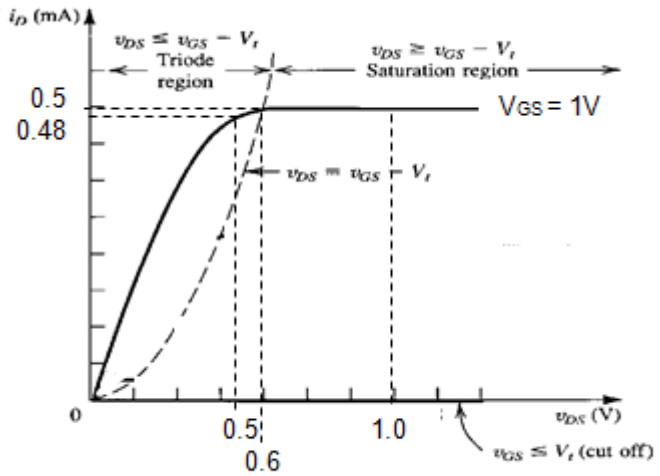
Since

$V_{ds} < (V_{gs} - V_{th})$ ➔ transistor is in Triode Mode.

Thus, $I_{ds} = K_n[W/L][(V_{gs} - V_{tn})V_{ds} - V_{ds}^2/2]$

$= 100*10^{-6} * (5/0.18) * [0.6*0.5 - 0.5^2/2]$

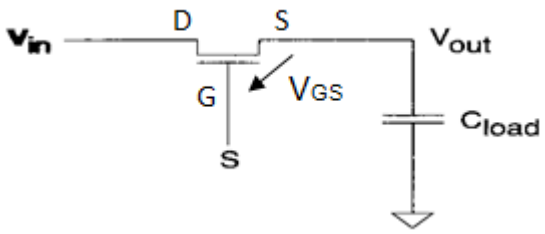
$=$ 0.486 mA



Question 5

Explain with the aid of diagrams why a N-channel (NMOS) pass transistor transfers :-
(a) logic "High" with degradation.
(b) logic "Low" without degradation.

Solution to Q5



NMOS Pass Transistor

An NMOS transistor can act like a switch. A signal voltage, S, is applied at the gate (with respect the source terminal; V_{gs}) to control the ON and OFF of the transistor. When S is low, the transistor is OFF and when S is high (V_{dd}), the transistor is ON.

(a) V_{in} = High and C_{load} is initially uncharged

C_{load} will starts to charge when the switch is ON, however V_{out} can only reach V_{in}-V_{tn} because exceeding this value, the ON condition will cease to exist and the transistor will turn OFF. Hence, it can be concluded that NMOS transistor *transfers logic High with degradation*. That is, the V_{out} cannot reach full $V_{in.}$ The direction of the current flow is from left to right.

(b) V_{in} = Low and C_{load} is initially charged to V_{in}-V_{tn}

With the switch is closed, C_{load} will starts to discharge and the current flows from right to left. This action can continues until C_{load} discharges to 0V. Thus, NMOS transistor *transfers Low logic without degradation*.

Question 6

Figure 3-50 shows a polysilicon wire with sheet resistance, R_s = 10 Ω/□. If the relative capacitance for polysilicon to substrate is 1.0 □Cg, compute the followings :-

(a) resistance of the polysilicon wire in ohms.

(b) capacitance of the polysilicon to substrate in terms of □Cg.

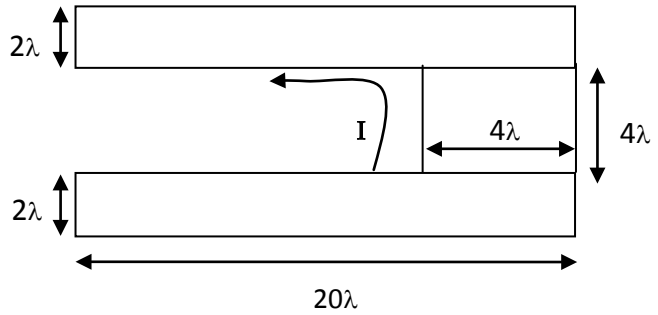
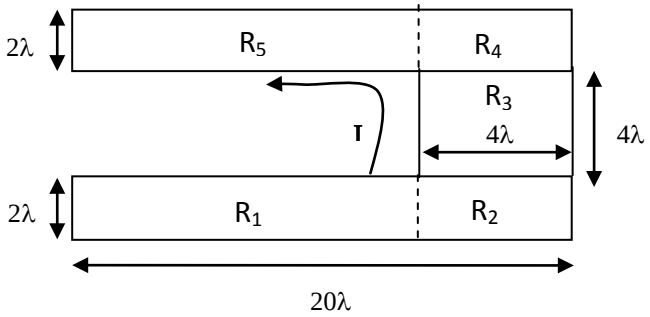


Figure 3-50

Solution to Q6



(i) resistance of wire = $R_1 + R_2 + R_3 + R_4 + R_5$
= $R_s (L_1/W_1 + ½ L_2/W_2 + L_3/W_3 + ½ L_4/W_4 + L_5/W_5)$
= 10 (16/2 + ½(4/2) + 4/4 + ½(2/4) + 16/2)
= **182.5 Ω**

(i) capacitance of wire = relative area * relative capacitance * □Cg
= $(A_1 + A_2 + A_3)/4$ * 1 * □Cg
= (2*20 + 4*4 + 2*20)/4 * 1 * □Cg
= **96 □Cg**

Note: relative area = ratio of actual area (A) to area of a minimum sizes gate

Question 7

Figure 3-51 shows two CMOS inverters cascaded together to drive a capacitive load of C_L = 16 □Cg. Assume $W_p = 2.5W_n$, for V_{in} = Low and V_{in} = High, calculate:-

(a) the delay experienced by the pair (from V_{in} to V_{out}) in terms of τ.

(b) average propagation delay, t_p.

Assume minimum-sized transistors for NMOS.

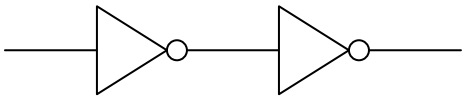
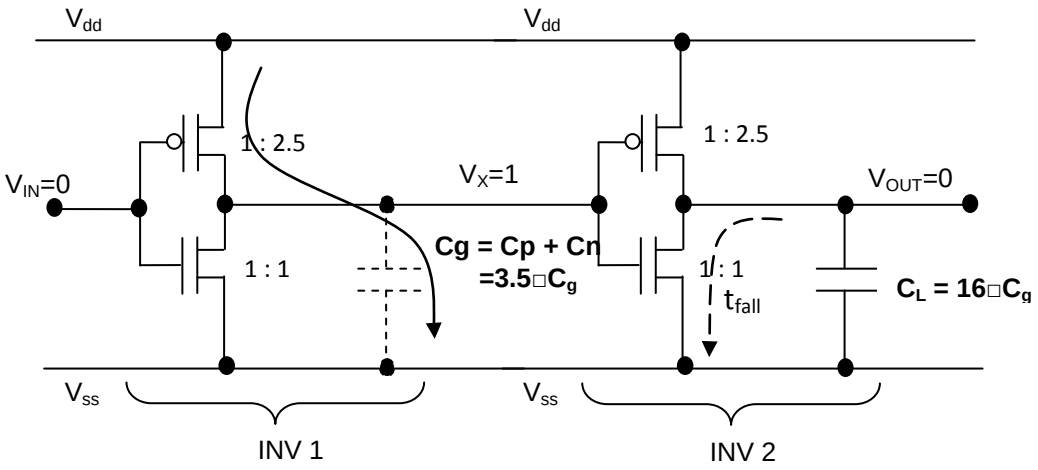


Figure 3-51

Solution to Q7



Transistor Level Circuitry

(a)(i) When V_{IN} = LOW

$T_{INV1}(OFF) = R_pC$

$= R_{SP}\ (1/2.5)^* \ 3.5□C_g$

$= (2.5\ R_{SN})(1/2.5) * 3.5□C_g$

$= 3.5\ \tau$

$T_{INV2}(ON) = R_nC$

$= R_{SN}\ (1/1) * 6□C_g$

$= 16\ \tau$

$t_{pHL} = 3.5\ \tau + 16\ \tau$

$= 19.5\ \tau$

(a)(ii) When V_{IN} = HIGH

$T_{INV1}(ON) = R_nC$
$= R_{SN}\ (1/1\ )* 3.5\square C_g$
$= 3.5\ \tau$

$T_{INV2}(OFF) = R_pC$
$= R_{SP}\ (1/2.5) * 16\square C_g$
$=(\ 2.5\ R_{SN}\ )(1/2.5) * 16\square C_g$
$= 16\ \tau$

$t_{pLH} = 3.5\ \tau + 16\ \tau$
$= 19.5\ \tau$

(b) Average propagation delay = $(t_{pHL} + t_{pHL})/2$
$= (19.5\ \tau + 19.5\ \tau) / 2$
= **$19.5\ \tau$**

Question 8

For the CMOS circuit shown in Figure 3-52, the PMOS transistor size of the NOR gate is W=16λ, L=2λ and that of the inverter is W=8λ, L=2λ respectively. The NMOS transistor size in both the inverter and NOR gate is L=2λ and W=2λ.

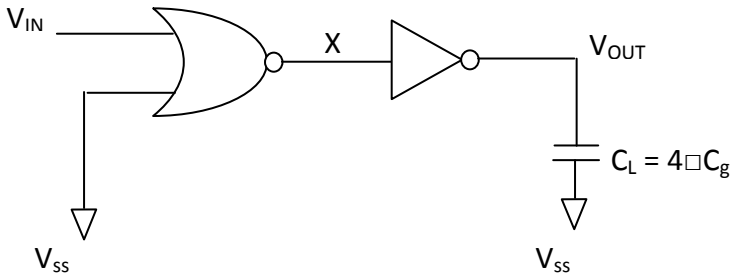


Figure 3-52

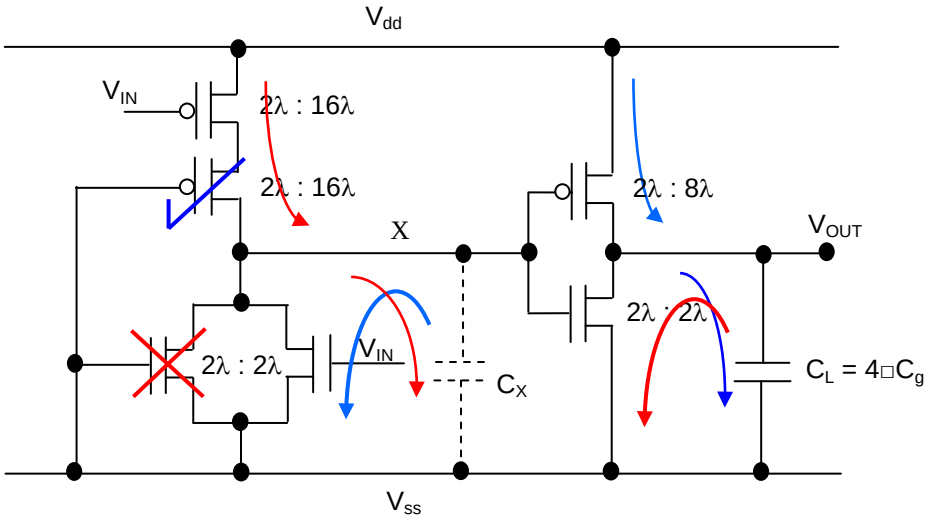
(a) Draw the transistor level circuitry for the circuit shown.

(b) Calculate the V_{IN} to V_{OUT} delay in terms of τ and indicate the direction of current flow if V_{IN} = High.

(c) Calculate the V_{IN} to V_{OUT} delay in terms of τ and indicate the direction of current flow if V_{IN} = Low.

(d) If the sizes of the PMOS transistors are the same as those of the NMOS transistors, comment on its effect on symmetricity of timing and performance of the circuit.

Solution to Q8



(a) $C_X = C_p + C_n$ where C = rel area * rel cap * $\square C_g$

$= \frac{(2*8)}{(2*2)} \square C_g + \frac{(2*2)}{(2*2)} \square C_g$

$= \mathbf{5 \square C_g}$

(b) When V_{IN} = "H", time taken to discharge V_X to "L" is :

$T_{NOR}(ON) = RC$

$= R_{Sn} (\frac{2}{2})$ x $5 \square C_g$

$= 5\ \tau$

With V_X = "L", time taken to charge V_{OUT} to "H" is :

$T_{INV}(OFF) = R_{Sp} (\frac{2}{8})$ x $4 \square C_g$

$= 2.5\ R_{Sn} (\frac{2}{8})$ x $4 \square C_g$ where $R_{Sp} \approx 2.5\ R_{Sn}$

$= 2.5\ \tau$

Total Delay = $T_{NOR}(ON) + T_{INV}(OFF)$

$= (5 + 2.5)\ \tau$

$= \mathbf{7.5\ \tau}$

(c) When V_{IN} = "L", time taken to charge V_X to "H" is :

$T_{NOR}(OFF) = RC$

$= R_{Sp} (\frac{2}{16} + \frac{2}{16})$ x $5\square C_g$

$= 2.5R_{Sn} (\frac{2}{16} + \frac{2}{16})$ x $5\square C_g$ where $R_{Sp} \approx 2.5\ R_{Sn}$

$= 3.125\ \tau$

With V_X = "H", time taken to discharge V_{OUT} to "L" is :

$T_{INV}(ON) = R_{Sn} (\frac{2}{2})$ x $4\square C_g$

$= 4\ \tau$

Total Delay = $T_{NOR}(OFF) + T_{INV}(ON)$

$= (3.125 + 4)\ \tau$

= **7.125 τ**

(d) The difference in sizing between the two types of transistors allows a more symmetrical propagation delay to be obtained.

However, if the two sets of transistors are made the same size, then the taken by the NOR gate when V_{IN} = "L" will be much longer, approx 25τ, while that taken by the INV when V_{IN} ="H" will be approx 10τ. These differences not only affect the symmetricity of the timing but also significantly reduce the performance of the circuit.

Question 9

For the circuit shown in Figure 3-53 determine the boolean expression at X.

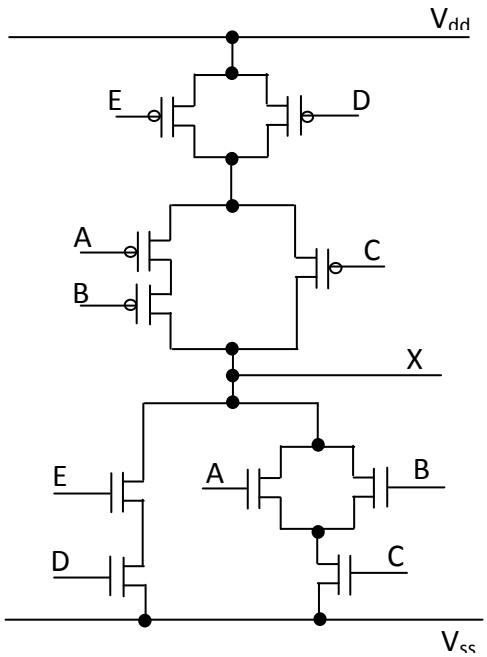


Figure 3-53

Solution to Q9

$X = \overline{DE + C(A+B)}$

Question 10

(a) Show the CMOS transistor level implementation of $S = \overline{A.B}.(C + D)$ using:
(i) three 2-inputs NAND gates only.
(ii) complex gate concept.
Assume the *negated* signals are readily available.

(b) With reference to part(a), which implementation is the preferred choice. Why?

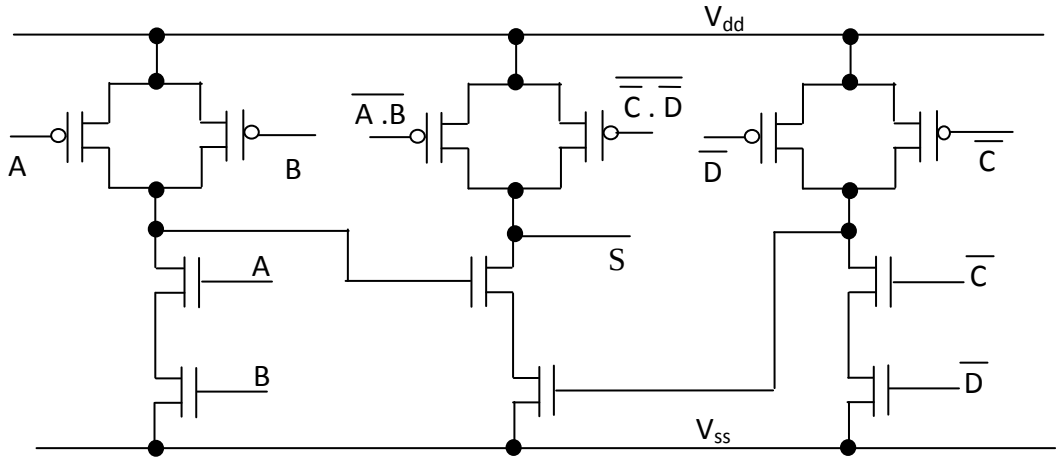
(c) With reference to a 3-inputs NAND gate, if the p-transistors and n-transistors are of the same size, will the rise time be shorter than, longer than or equal to the fall time? Why?

(d) Explain briefly why CMOS circuitry is known to be ratioless.

(e) Discuss briefly the power dissipation in CMOS circuitry.

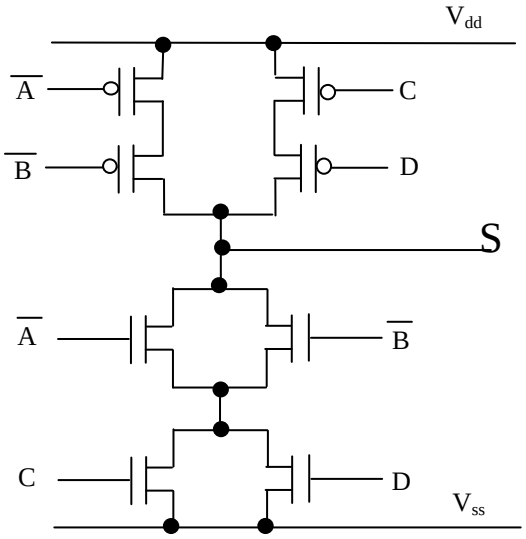
Solution to Q10

(a)(i) $S = A \,.\, B \,.\, (\overline{C} \,.\, \overline{D})$



Function Implemented Using Standard NAND Gates Concept

(ii) $S = \overline{(\overline{A} + \overline{B}) \cdot (C + D)}$



Function Implemented Using Complex Gates Concepts

(b) The Complex Gates implementation is preferred as it requires fewer transistors for implementation of the same function.

(c) As electrons has a mobility of about 2.5x ~ 3x that of holes, for transistors of equal sizes, $R_p \approx 3R_n$.

During T_{ON}, all 3 n-transistors will be ON, hence the total resistance experienced during this time is $3R_n$. Thus, $\mathbf{T_{fall} \approx 3\ R_nC}$.

For $T_{OFF,}$ 3 possibilities exist namely:

- 1 PMOS is ON:- the total resistance experienced during this time is R_p ($\approx 3R_n$). Thus, $\mathbf{T_{rise} \approx\ 3R_nC}$.

- 2 PMOS are ON: the total resistance experienced during this time is $\frac{1}{2}R_p$ ($\approx 1.5R_n$). Thus, $\mathbf{T_{rise} \approx\ 1.5R_nC}$.

- 3 PMOS are ON: the total resistance experienced during this time is $^1/_3R_p$ ($\approx R_n$). Thus, $\mathbf{T_{rise} \approx R_nC}$.

Hence, the T_{rise} is generally shorter than the T_{fall} except under worst case scenario. The shorter T_{rise} is due to the parallel combination of the PMOS.

(d) CMOS circuits are ratioless because $V_{OL} = V_{ss}$ (ground) and does not depends on the ratio of PMOS to NMOS transistors.

(e) At V_{OL}, only NMOS is 'ON' while PMOS is 'OFF' and vice versa when output is at V_{OH}. Hence, no current flows from V_{dd} rail to V_{ss} rail.

Question 11

Given the following Boolean equations, show their implementation using complex gate concept in CMOS technology.

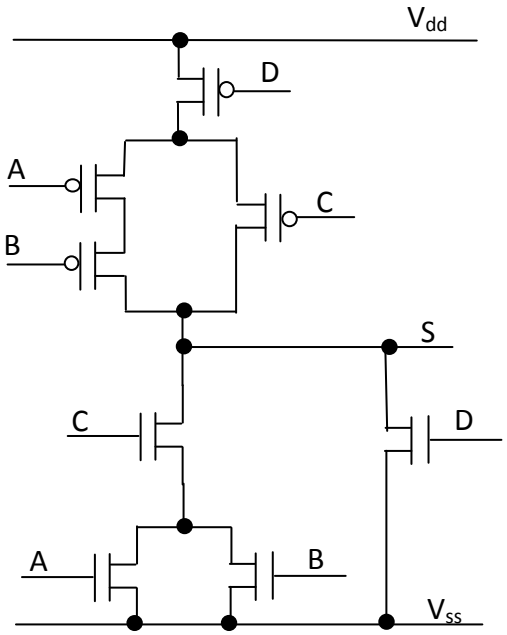
(a) $X = \overline{(A + B)\, C + D}$

(b) $Y = \overline{A + D\,(B + \overline{C})}$

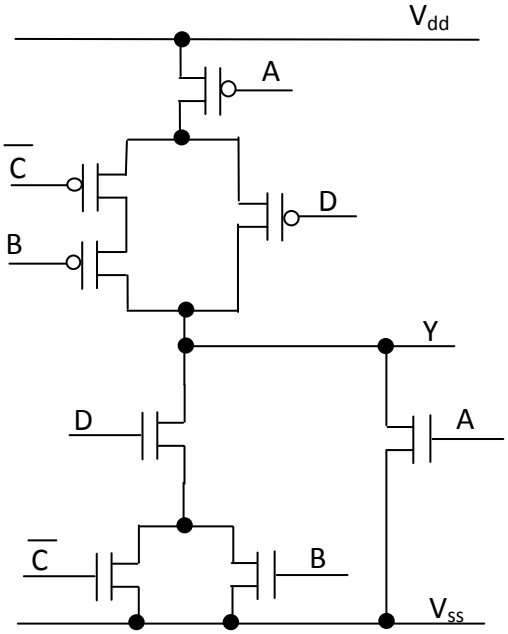
(c) $Z = A + B\,C$

Solution to Q11

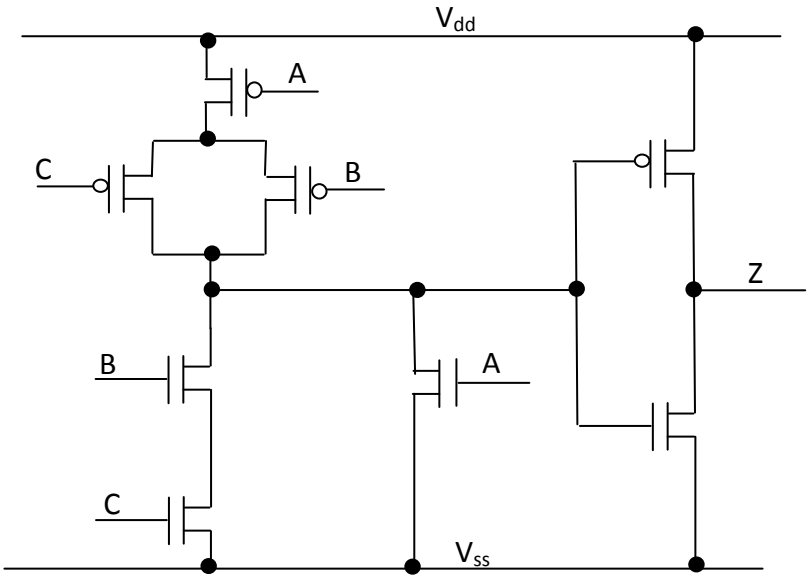
(a) $X = \overline{(A + B)\, C + D}$



(b) $Y = \overline{A + D\,(B + \overline{C})}$



(c) Z = A + B C



Question 12

For the logic function F = AB + C + D, show how it can be implemented in CMOS technology using:

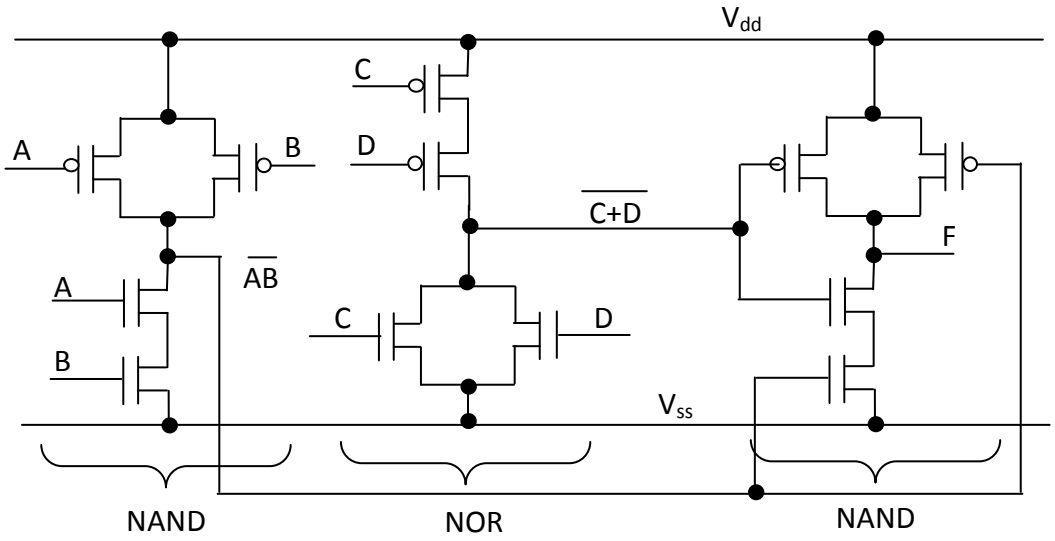
(a) standard NAND/NOR/INVERTER gates.

(b) complex gate concept.

Solution to Q12

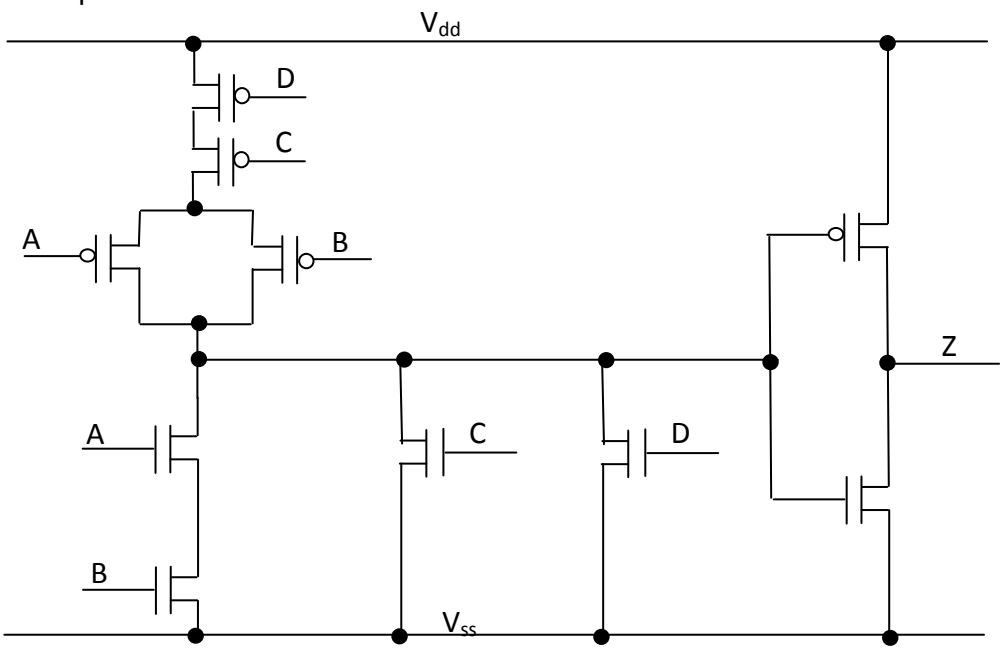
(a) standard NAND/NOR/INVERTER gates.

$F = AB + C + D = \overline{\overline{AB} \cdot \overline{C + D}}$



(b) Complex gate concept

F = AB + C + D



Question 13

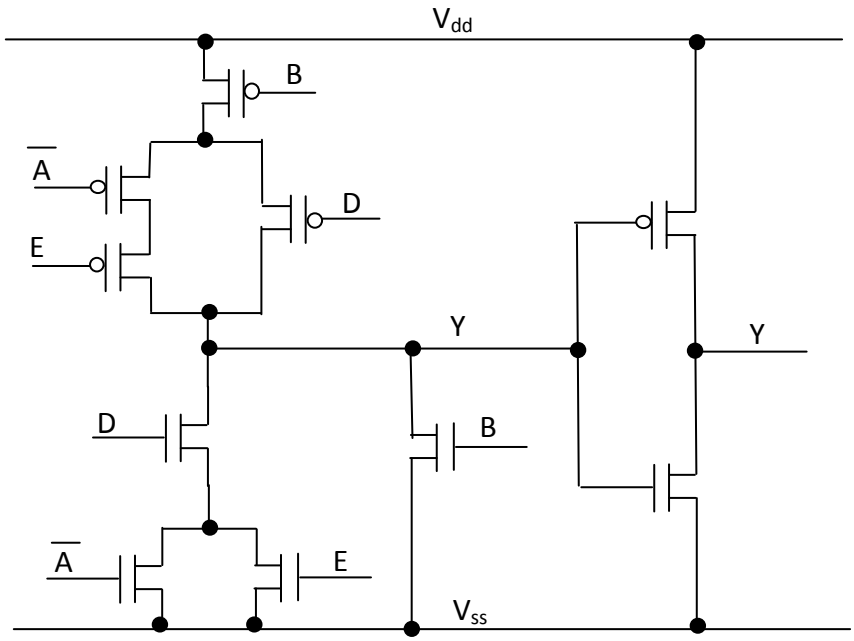
For the logic function $X = B + D(E + \overline{A})$,

(a) show its implementation using CMOS complex gate.

(b) indicate and explain which transistors in the circuit is affected by body effect.

Solution to Q13

(a) $Y = B + D(E + \overline{A})$



(b) When two devices are connected is series (e.g. NAND gate), their V_{sb} will be different, that is, $V_{sb1} = 0$ but $V_{sb2} \neq 0$. With $V_{sb2} \neq 0$, an increase in the thickness of the depletion region under the channel of will occur. The increased depletion layer requires additional charge to be supplied. If V_{gs} is held constant, the charge in the channel will decrease thereby reducing conductivity. To maintain the channel conductivity, V_{gs} need to increase.

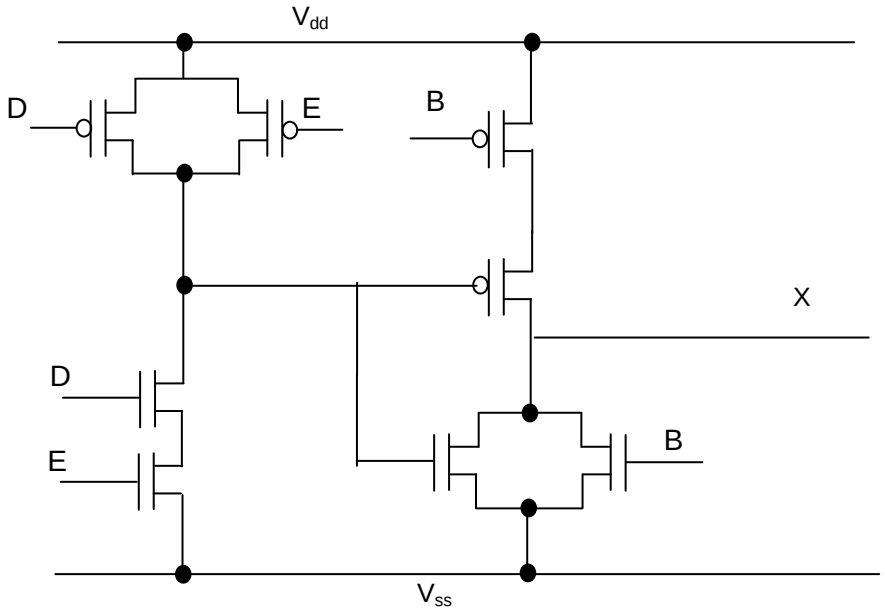
Thus, transistors whose $V_{sb2} \neq 0$ are said to experience body effect.

Transistors experiencing body effect are :

NMOS - D

PMOS - $\overline{A}$, E, D

Question 14



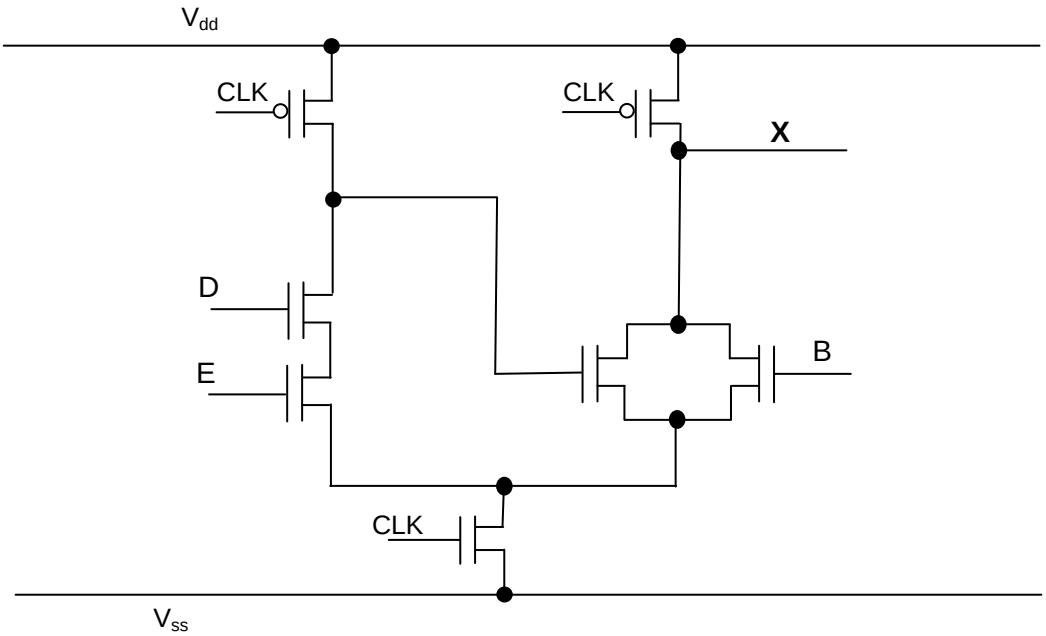
(a) Determine the boolean expression at X.

(b) Redraw using CMOS **dynamic** gate concept.

Solution to Q14

(a) $X = \overline{\overline{DE} + B}$

(b)



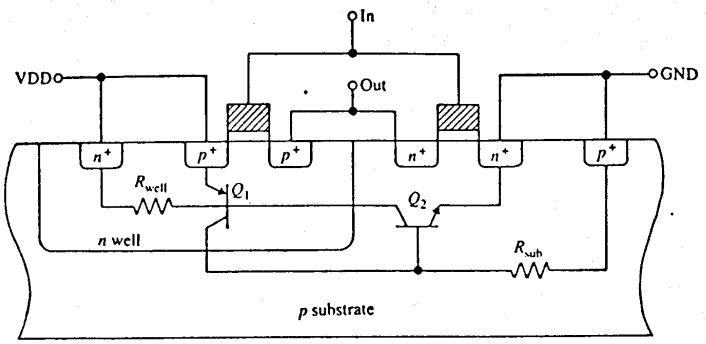
Question 15

(a) Sketch the cross-sectional structure of a N-well process CMOS inverter showing all parasitic elements that constitute the occurrence of latch-up.

(b) Explain how latch-up occurs and how can it be overcome.

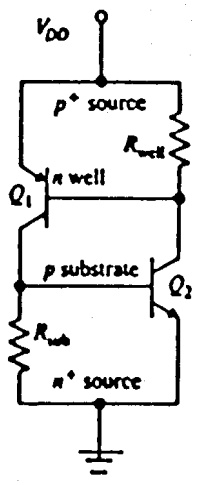
Solution to Q15

(a)



(b) Latch up is the inadvertent creation of a low-impedance path between the power supply rails of a CMOS circuit. It occurs when the parasitic structure within the CMOS circuit is activated thus disrupting proper functioning of the part and possibly even leading to circuit destruction due to over-current.

A scenario of latch-up is when sufficient current flows through R_{sub}, which in turn causes Q_2 to turn ON. With Q_2 ON, current will be drawn through R_{well}. If sufficient voltage develops across R_{well}, Q_1 can then turn ON, causing more current to flow throug



R_{sub} is due to the p-substrate
R_{well} is due to the n-well

Latch up in CMOS circuits can be reduced by :-

(i) increasing the distance between N-well and N+ source/drain diffusion of the NMOS device $\Rightarrow$ reduction of β_2.

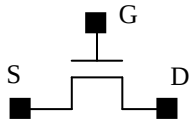
(ii) using guard rings around the transistors to provide a low resistance path to collect holes before they interfere with the operation of other circuits outside the guard ring.

(iii) placing a well tie and a substrate tie close to the supplies. These will in turn reduce the value of R_{sub} and R_{well} $\Rightarrow$ higher accidental current is required to cause latch-up.

(iv) increase substrate doping level $\Rightarrow$ reduction in R_{sub}

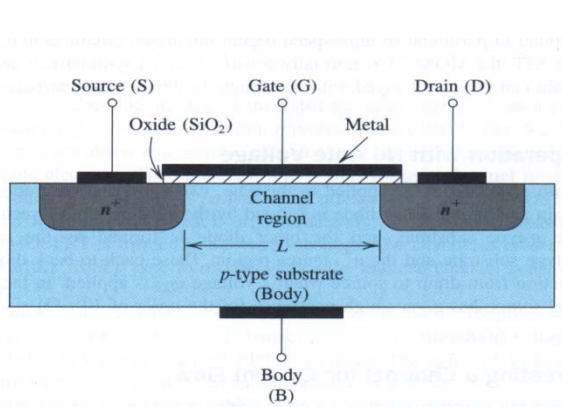
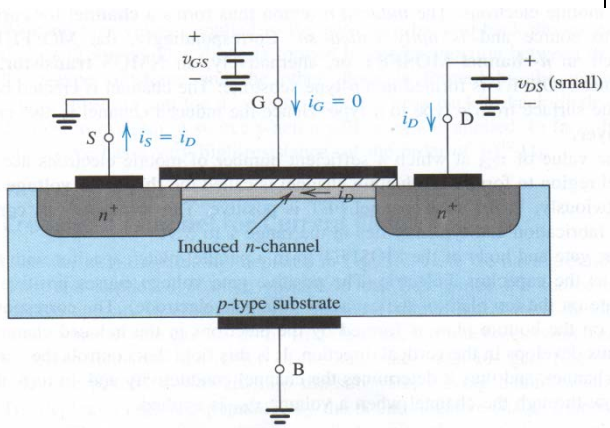
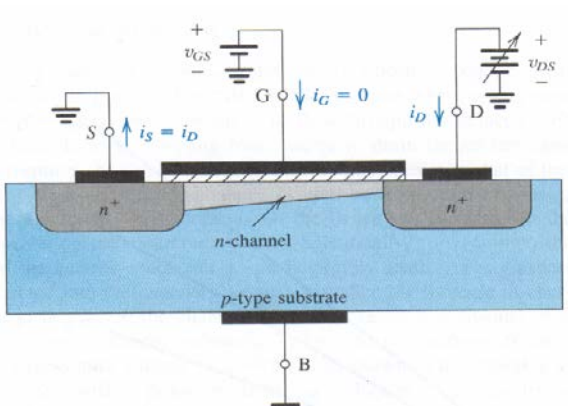
Chapter 3 – Summary

3.1.4 Operation of NMOS Enhancement Transistor



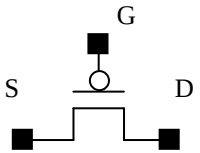
Source (S) ←→ Drain (D)
Source is the terminal with a *lower* voltage
($V_S < V_D$)

3 modes of Operation

Pg 4 **Cut-off Mode**	Pg 4 **Triode/Linear Mode**	Pg 5 **Saturation Mode**
		
OFF : $V_{gs} < V_t$ where $V_t \approx 0.2V_{dd}$ $V_{dd} = 1.2V$ $I_{ds} = 0$	ON : $V_{gs} \geq V_t$ where $V_t \approx 0.2V_{dd}$ Linear : $V_{ds} < (V_{gs} - V_t)$ $I_{ds} = K_n[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$	ON : $V_{gs} \geq V_t$ where $V_t \approx 0.2V_{dd}$ Saturation : $V_{ds} \geq (V_{gs} - V_t)$ $I_{ds} = \{K_n/2\}[W/L](V_{gs} - V_t)^2$

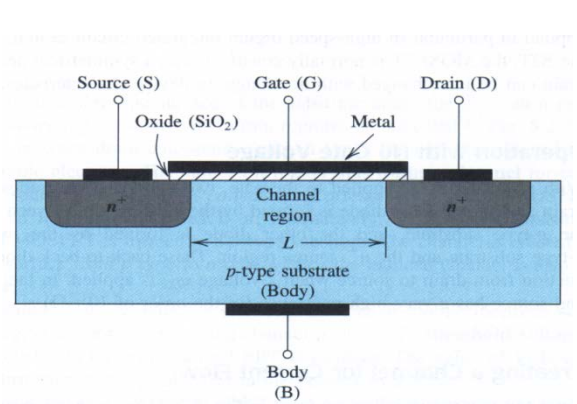
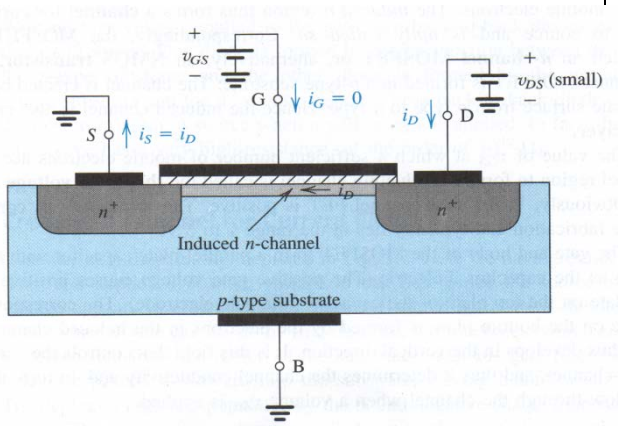
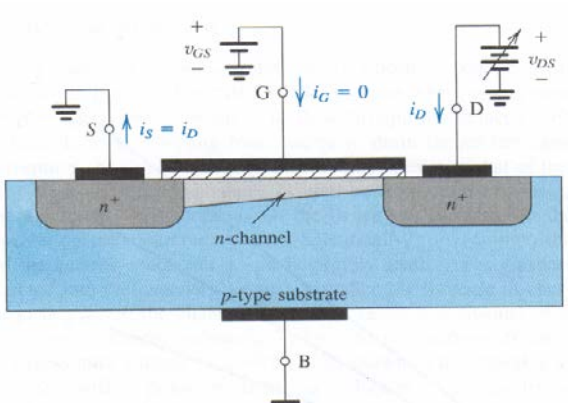
Note : Gate voltage controls ON/OFF of transistor (ON : "1", OFF : "0")
with Source (S) voltage at : $V_S < (V_G - V_T)$, $V_{S(min)} = 0V$

3.1.4 Operation of PMOS Enhancement Transistor



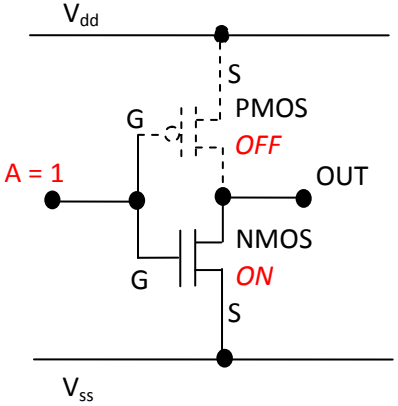
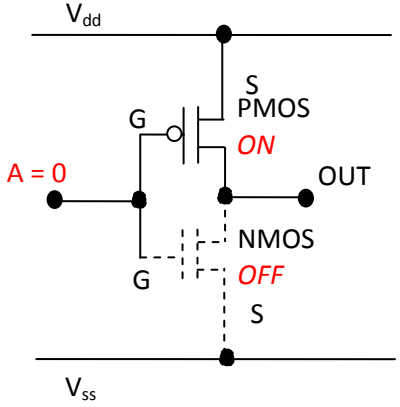
Source (S) ←→ Drain (D)
Source is the terminal with a *higher* voltage
($V_S > V_D$)

3 modes of Operation

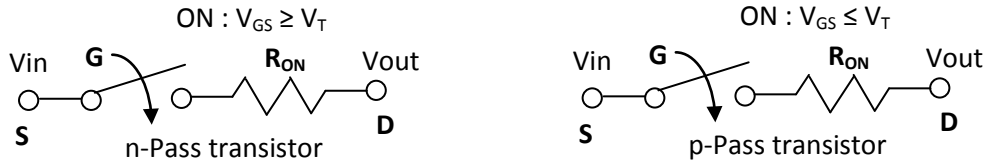
Pg 7 **Cut-off Mode**	Pg 7 **Triode/Linear Mode**	Pg 7 **Saturation Mode**
		
OFF : $V_{gs} > V_t$ where $V_t \approx -0.2V_{dd}$ $V_{dd} = 1.2V$ $I_{ds} = 0$	ON : $V_{gs} \leq V_t$ where $V_t \approx -0.2V_{dd}$ Linear : $V_{ds} > (V_{gs} - V_t)$ $I_{ds} = K_p[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$	ON : $V_{gs} \leq V_t$ where $V_t \approx -0.2V_{dd}$ Saturation : $V_{ds} \leq (V_{gs} - V_t)$ $I_{ds} = \{K_p/2\}[W/L](V_{gs} - V_t)^2$

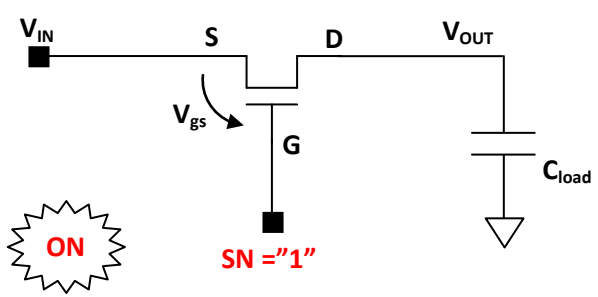
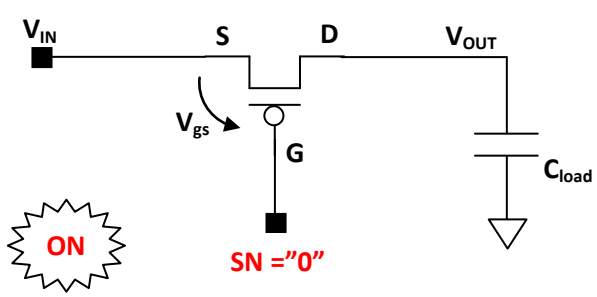
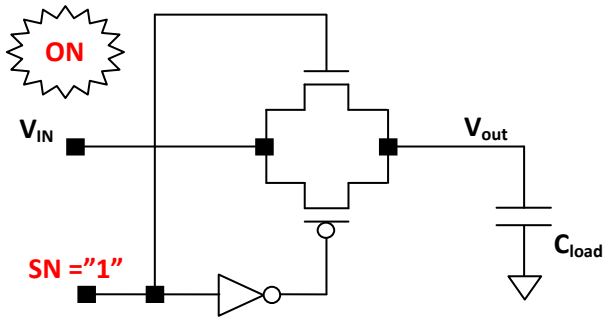
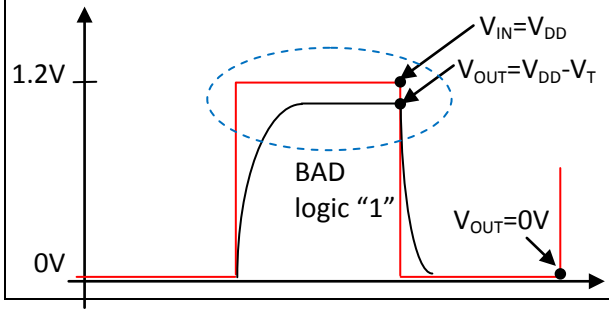
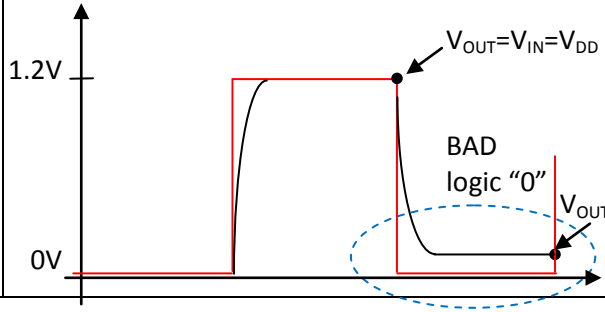
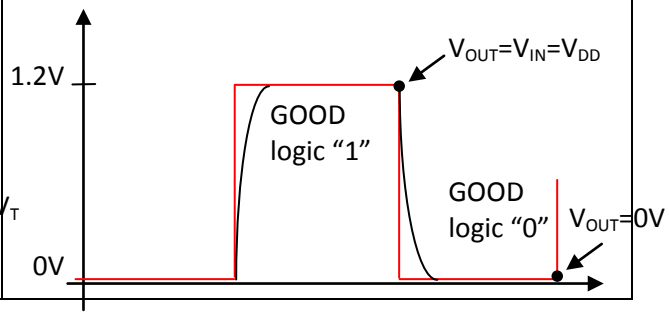
Note : Gate voltage controls ON/OFF of transistor (ON : "0", OFF : "1")
with Source (S) voltage at : $V_S > (V_G - V_T)$, $V_{S(max)} = V_{dd}$

Summary of Operation Modes

	Pg 7 **NMOS**	Pg 7 **PMOS**
OFF	$V_{gs} < V_t$ where $V_t \approx 0.2V_{dd}$ and $V_{dd} = 1.2V$	$V_{gs} > V_t$ where $V_t \approx -0.2V_{dd}$ and $V_{dd} = 1.2V$
ON : Linear Region	$V_{gs} \geq V_t$ $V_{ds} < (V_{gs} - V_t)$ $I_{ds} = K_n[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$	$V_{gs} \leq V_t$ $V_{ds} > (V_{gs} - V_t)$ $I_{ds} = K_p[W/L][(V_{gs} - V_t)V_{ds} - V_{ds}^2/2]$
ON : Saturation Region	$V_{gs} \geq V_t$ $V_{ds} \geq (V_{gs} - V_t)$ $I_{ds} = [K_n/2][W/L](V_{gs} - V_t)^2$ where $K_n = \mu_n \varepsilon_o \varepsilon_r / t_{ox}$ and ε_o = permittivity of free space ε_r = relative permittivity of SiO_2 t_{ox} = thickness of gate oxide	$V_{gs} \leq V_t$ $V_{ds} \leq (V_{gs} - V_t)$ $I_{ds} = [K_p/2][W/L](V_{gs} - V_t)^2$ where $K_p = \mu_p \varepsilon_o \varepsilon_r / t_{ox}$ and ε_o = permittivity of free space ε_r = relative permittivity of SiO_2 t_{ox} = thickness of gate oxide
PMOS and NMOS operate in complementary mode	 Logic "1" at Gate terminal NMOS ➔ ON On condition that : $V_S \leq V_G - V_T$	 Logic "0" at Gate terminal PMOS ➔ ON On condition that : $V_S \geq V_G - V_T$

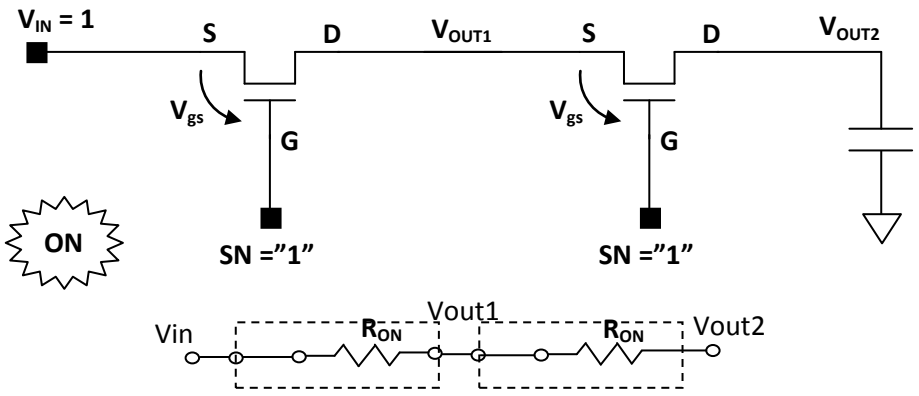
3.2 THE MOS SWITCHES



Pg 8 **NMOS Pass-Transistor**	Pg 9 **PMOS Pass-Transistor**	Pg 9 **CMOS Transmission Gate**
		
Good : When $V_{IN} = 0V$ ("0") , $V_{OUT} = 0V$ Transfer logic "0" without degradation Bad : When $V_{IN} = V_{dd}$ ("1") , $V_{OUT} = V_{dd}-V_t$ (V_{out} cannot reach V_{dd}) Transfer logic "1" with degradation Losses = V_T + Resistive drop	Good : When $V_{IN} = V_{dd}$ ("1") , $V_{OUT} = V_{dd}$ Transfer logic "1" without degradation Bad : When $V_{IN} = 0V$ ("0"), $V_{OUT} = V_t$ (V_{OUT} cannot reach 0V) Transfer logic "0" with degradation Losses = V_T + Resistive drop	Good : Tranfer logic "1" without degradation ($V_{out} = V_{dd}$) Good : Transfer logic "0" without degradation ($V_{out} = 0V$) Because at least one transistor is ON at any time
		

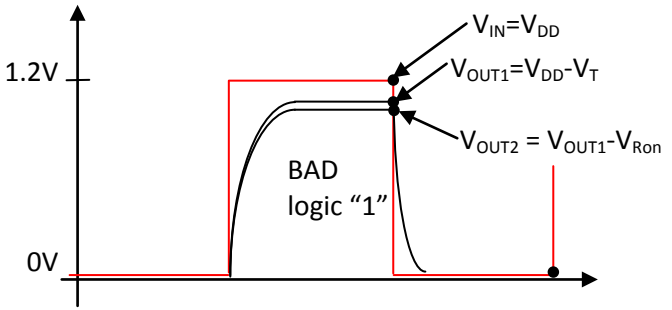
3.2 THE MOS SWITCHES

Lab 2 **NMOS Pass-Transistors In Series**

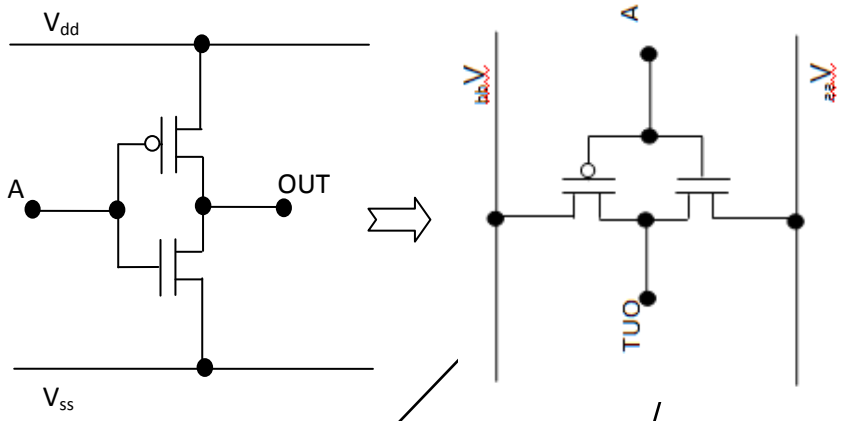


When $V_{IN} = V_{dd}$ ("1")
SW1 : $V_{OUT1} = V_{dd}-V_t$ (V_{out} cannot reach V_{dd})
due to Losses = $V_T + V_{Ron}$

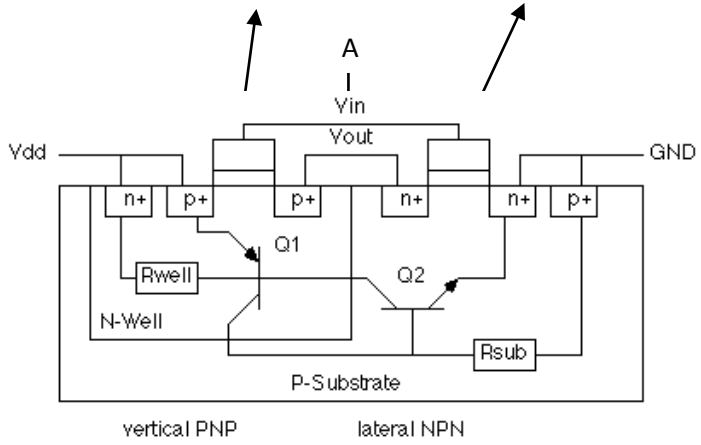
SW2 : $V_{IN} = V_{OUT1} = V_{dd}-V_t$ (degraded "1"), $V_{OUT2} = V_{OUT1}-V_{Ron}$
due to Losses = V_{Ron}



Pg 11 **CMOS Inverter**



PMOS (in n-well) NMOS (in p-substrate)



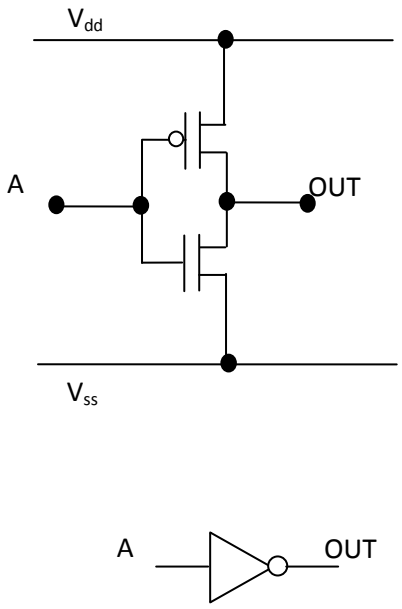
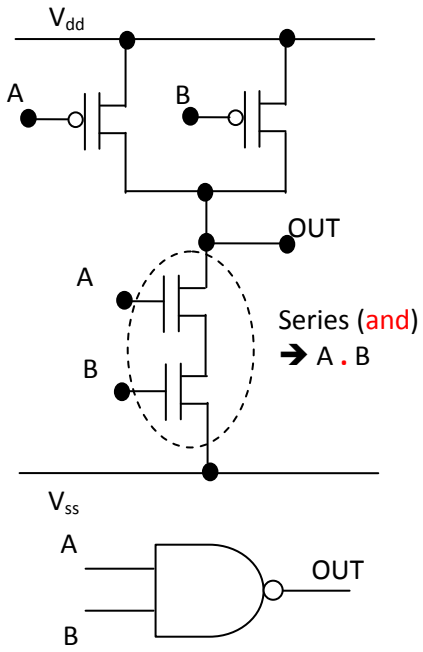
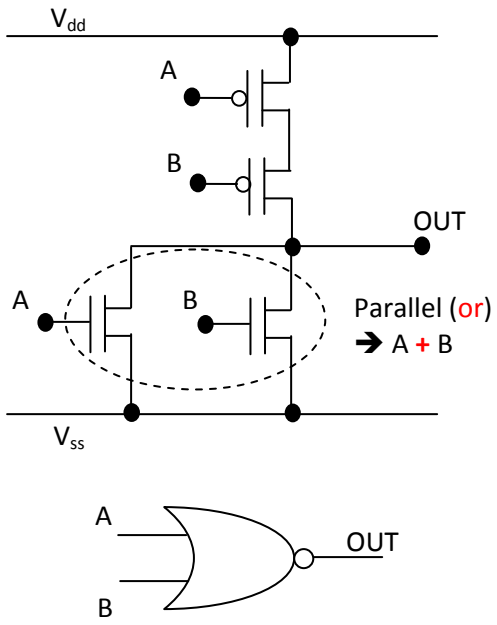
well-tie
- to conr
the we
to V_{dd}

rate-tie
connect
substrate
V_{ss} (Gnd)

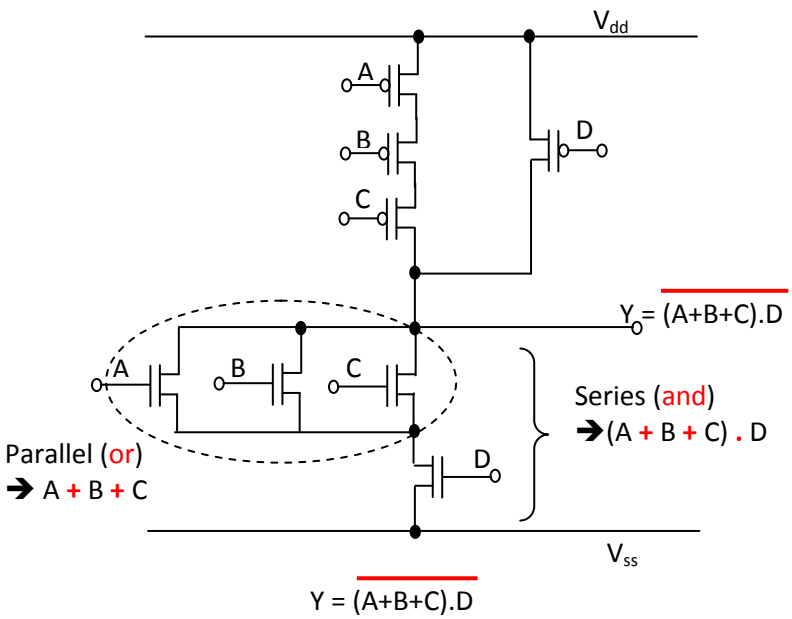
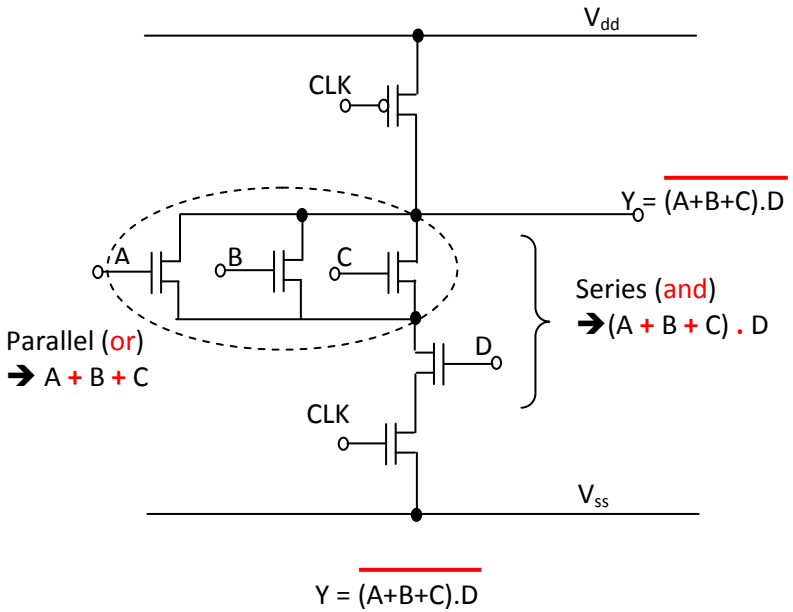
Cross-sectional view of CMOS Inverter

3.3 CMOS CIRCUITS

(I) STATIC *STANDARD* GATE

Pg 10 **CMOS Inverter**	Pg 22 **CMOS Nand Gate**	Pg 24 **CMOS Nor Gate**
		
$OUT = \overline{A}$	$OUT = \overline{A \cdot B}$	$OUT = \overline{A + B}$
Boolean equations are derived from the NMOS transistors	Boolean equations are derived from the NMOS transistors	Boolean equations are derived from the NMOS transistors

3.3 CMOS CIRCUITS

(II) ***STATIC COMPLEX* GATE** Pg 25	***(III)*** ***DYNAMIC COMPLEX* GATE** Pg 26
	
1. Each input signal (eg A, B or C) drives TWO transistors 2. Output is charged / discharged upon logic combinations outcome	1. Each input signal (eg A, B or C) drives only ONE transistor ➔ *smaller input capacitance*, hence faster response (T=RC) 2. Uses the idle cycle of the clock to *pre-charge* the output line (Y) ➔ save time to charge output to "1" when logic requires it to be at "1" Therefore, dynamic circuits are faster than static circuits.
Does not require clock circuitry. Suited for small circuitries.	Requires clock circuitry (high overhead). Not suitable for small circuits.

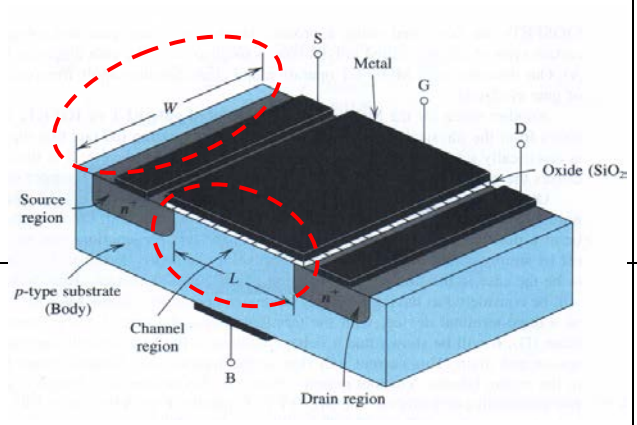
3.3 R & C (Transistor Path) Pg 13

<u>Resistance</u>

$R = R_s[L/W]$

L ↑ ➔ R ↑
W ↑ ➔ R ↓

L (// to current flow)
W (⊓ to current flow)



<u>Capacitance</u>

C = relative area * relative capacitance * □C_g

$= \frac{actual\ area}{reference\ area}$ * relative capacitance * □C_g

<u>Sheet Resistance (R_s)</u>

NMOS : R_{sn}

PMOS : $R_{sp} = 2.5R_{sn}$

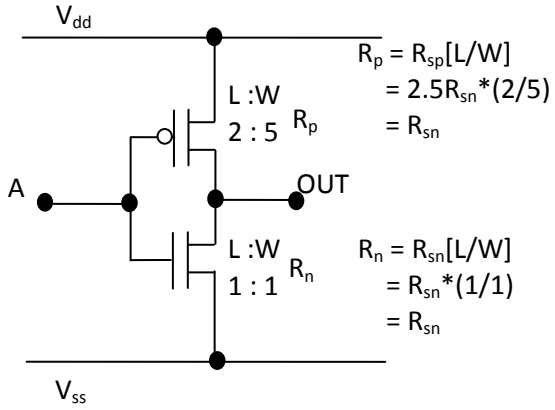
<u>Reference Area</u>

= Area of minimum-sized transistor

= 1*1 or 2λ*2λ

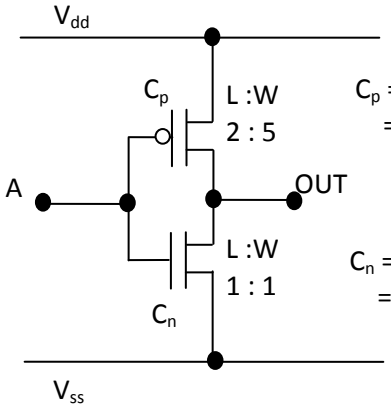
<u>Relative Capacitance</u>

Gate : 1



$R_p = R_{sp}[L/W]$
$= 2.5R_{sn}*(2/5)$
$= R_{sn}$

$R_n = R_{sn}[L/W]$
$= R_{sn}*(1/1)$
$= R_{sn}$

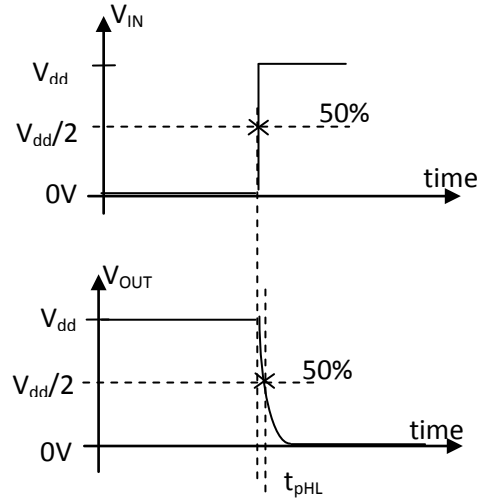


$C_p = (2*5) / (1*1) * 1 *$ □C_g
$= 10*$ □C_g

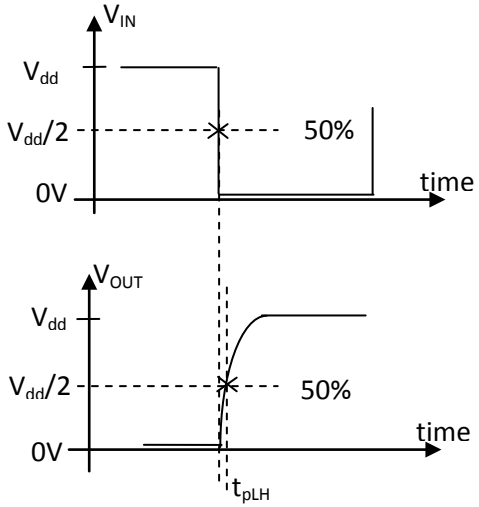
$C_n = (1*1) / (1*1) * 1 *$ □C_g
$= 1*$ □C_g

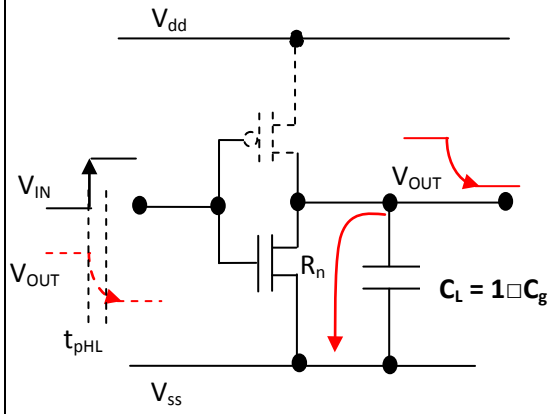
3.3.3.2 Propagation Delay (t = RC) Pg 17

tpHL :
Time taken for output to discharge from HIGH to LOW (wrt input)



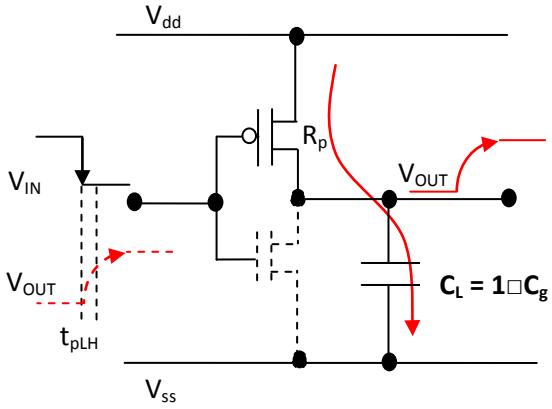
tpLH :
Time taken for output to charge from LOW to HIGH (wrt input)





For V_{IN} = 1,
NMOS ➔ ON, PMOS ➔ OFF
V_{OUT} will discharge (to LOW)

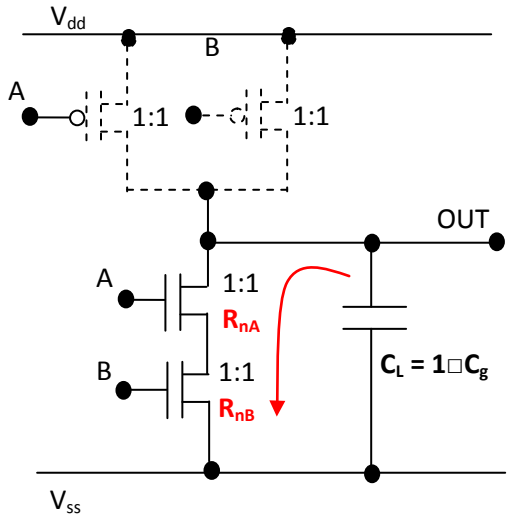
Time taken = tpHL = R*C
= R_n * C_L
= R_{sn}(L/W) * C_L



For V_{IN} = 0,
PMOS ➔ ON, NMOS ➔ OFF
V_{OUT} will charge (to HIGH)

Time taken = tpLH = R*C
= R_p * C_L
= R_{sp}(L/W) * C_L
= 2.5R_{sn}(L/W)*C_L

3.3.6 Propagation Delay (NAND Gate)

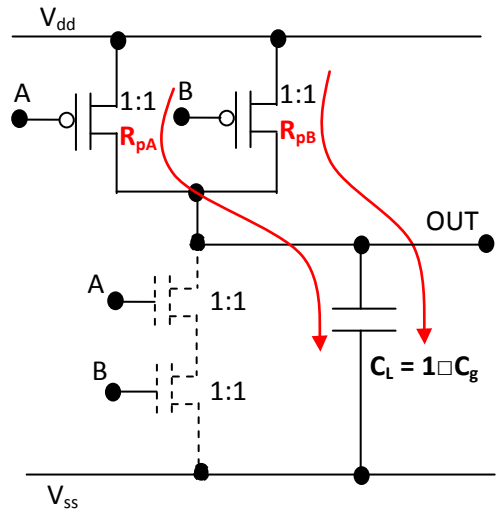


For A = 1, B = 1
NMOS ➔ ON, PMOS ➔ OFF
OUT will discharge (to LOW)

Time taken = tpHL
$= (R_{nA}+R_{nB}) * C_L$

$tpHL = 2R_{sn}(1/1) * C_L$
$= 2R_{sn} * 1\square C_g$
$= 2\ \tau$

Note : $\tau = R_{sn}*\square C_g$

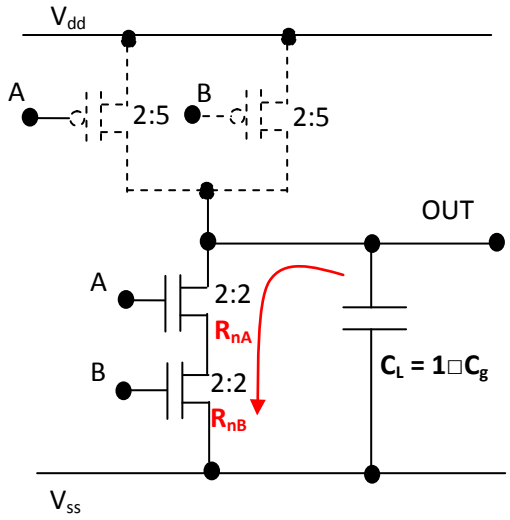


For A = 0, B = 0
PMOS ➔ ON, NMOS ➔ OFF
OUT will charge (to HIGH)

Time taken = tpLH
$= (R_{pA} // R_{pB}) * C_L$

$= \frac{1}{2}R_{sp}(1/1)* C_L$
$= \frac{1}{2}\ (2.5R_{sn}) * 1\square C_g$
$= 1.25\ \tau$

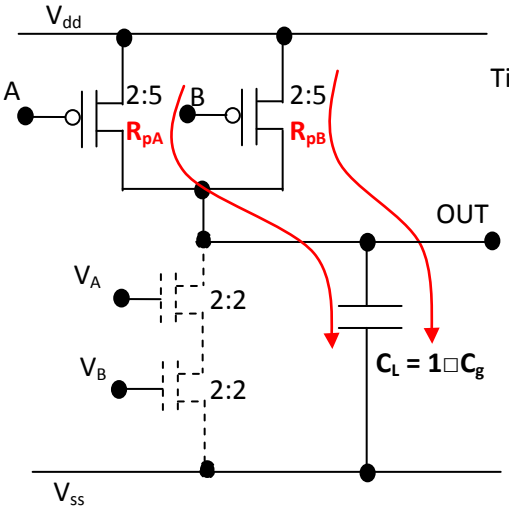
Note : $\tau = R_{sn}*\square C_g$



Time taken = tpHL
$= (R_{nA}+R_{nB}) * C_L$

$tpHL = 2R_n * C_L$
$= 2R_{sn}(2/2) * C_L$
$= 2R_{sn} * 1\square C_g$
$= 2\ \tau$

Note : $\tau = R_{sn}*\square C_g$



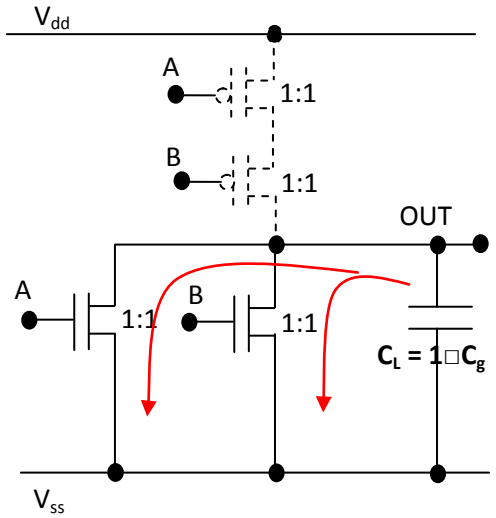
Time taken = tpLH
$= (R_{pA} // R_{pB}) * C_L$

$= \frac{1}{2}R_p * C_L$
$= \frac{1}{2}R_{sp}(2/5)* C_L$
$= \frac{1}{2}(2.5R_{sn})(2/5)*1\square C_g$
$= 0.5\ \tau$

Note : $\tau = R_{sn}*\square C_g$

Timing performance is affected by :
1. Size of transistors and
2. Number of transistor turned ON

3.3.6 Propagation Delay (NOR Gate) Pg 23 /24

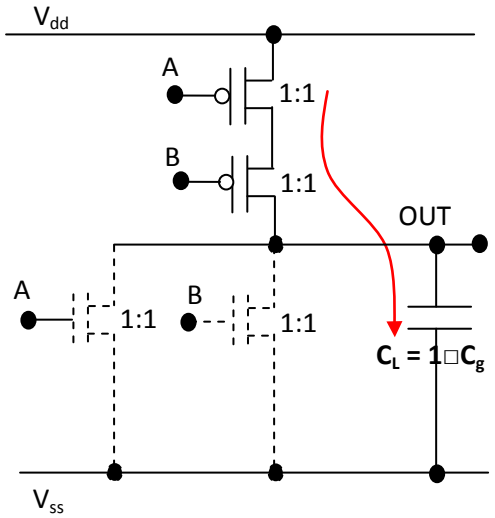


For A = 1, B = 1
NMOS ➔ ON, PMOS ➔ OFF
OUT will discharge (to LOW)

Time taken = tpHL
$= (R_{nA}//R_{nB}) * C_L$

$= ½R_n * C_L$
$= ½R_{sn}(1/1) * C_L$
$= ½R_{sn} * 1\square C_g$
$= ½ \tau$

Note : $\tau = R_{sn}*\square C_g$

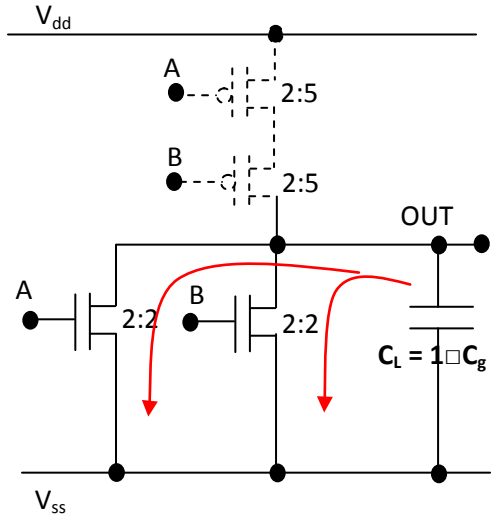


For A = 0, B = 0
PMOS ➔ ON, NMOS ➔ OFF
OUT will charge (to HIGH)

Time taken = tpLH
$= (R_{pA} + R_{pB}) * C_L$

$= 2R_p * C_L$
$= 2R_{sp}(1/1) * C_L$
$= 2\ (2.5R_{sn}) * 1\square C_g$
$= 5 \tau$

Note : $\tau = R_{sn}*\square C_g$

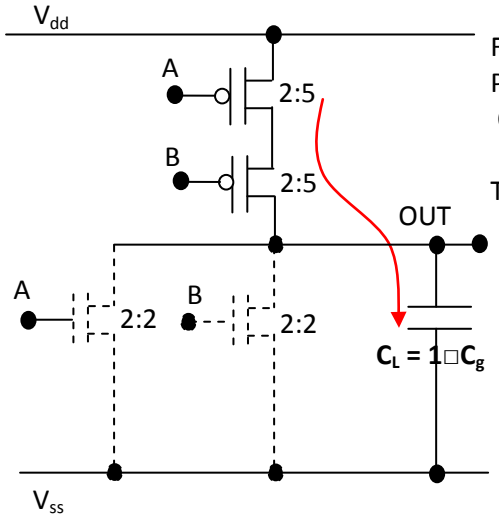


For A = 1, B = 1
NMOS ➔ ON, PMOS ➔ OFF
OUT will discharge (to LOW)

Time taken = tpHL
$= (R_{nA}//R_{nB}) * C_L$

$= ½R_n * C_L$
$= ½R_{sn}(2/2) * C_L$
$= ½R_{sn} * 1\square C_g$
$= ½ \tau$

Note : $\tau = R_{sn}*\square C_g$



For A = 0, B = 0
PMOS ➔ ON, NMOS ➔ OFF
OUT will charge (to HIGH)

Time taken = tpLH
$= (R_{pA} + R_{pB}) * C_L$

$= 2R_p * C_L$
$= 2R_{sp}(2/5) * C_L$
$= 2(2.5R_{sn})(2/5) * 1\square C_g$
$= 2 \tau$

Note : $\tau = R_{sn}*\square C_g$

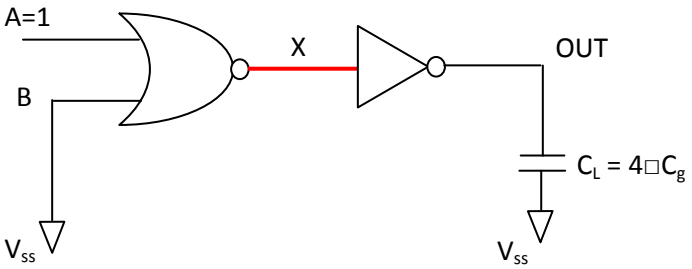
Timing performance is affected by :
1. Size of transistors and
2. Number of transistor turned ON

3.3.3.2 Propagation Delay Pg 42

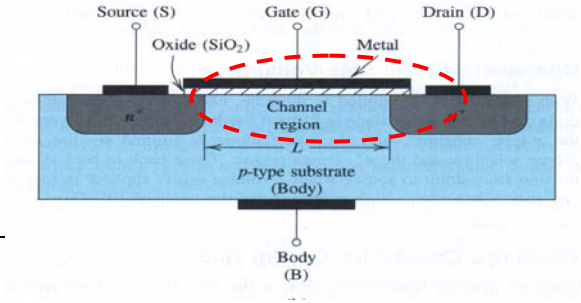
Given

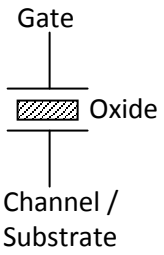
PMOS ➔ L:W = 2:5
NMOS ➔ L:W = 2:2

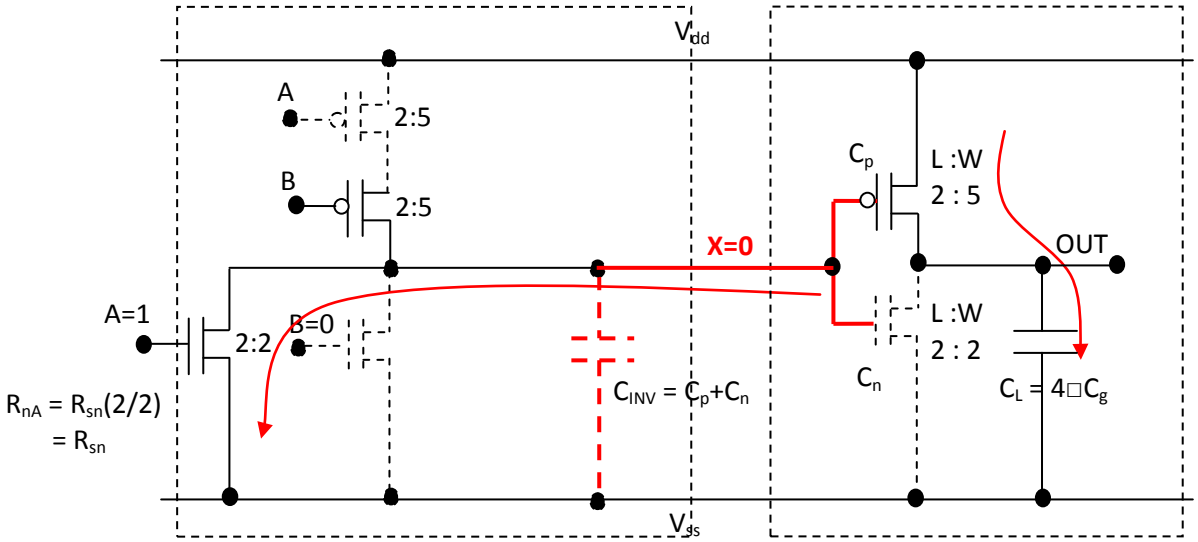
Calculate the time taken for signal at A to reach OUT when A=1.



Gate Capacitance







C = relative area * relative capacitance $*\square C_g$

Gate Capacitance of PMOS (Inverter)

$C_p = (2*5) / (1*1) * 1 *\square C_g$
$= 10*\square C_g$

Gate Capacitance of NMOS (Inverter)

$C_n = (2*2) / (1*1) * 1 *\square C_g$
$= 4*\square C_g$

NOR :
Time taken = tpHL
$= R_{nA} * C_{INV}$
$= R_{sn}(2/2) *(C_p+C_n)$
$= R_{sn} * 14\square C_g$
$=14\ \tau$

INVERTER :
Time taken = tpLH
$= R_p * C_L$
$= R_{sp}(2/5) * C_L$
$= (2.5R_{sn})(2/5) * 4\square C_g$
$= 4\ \tau$

From A to OUT when A=1 :
Total time taken = $t_{NOR} + t_{INV}$
$= 18\tau$

3.3 R & C (Wiring Path) Pg 13 / Pg 15

<u>Resistance</u>

$R = R_s[L/W]$

L = dimension parallel to current
W = dimension perpendicular to length

<u>Capacitance</u>

C = relative area * relative capacitance * $\square C_g$

$= \frac{actual\ area}{reference\ area}$ * relative capacitance * $\square C_g$

<u>Sheet Resistance (R_s)</u>

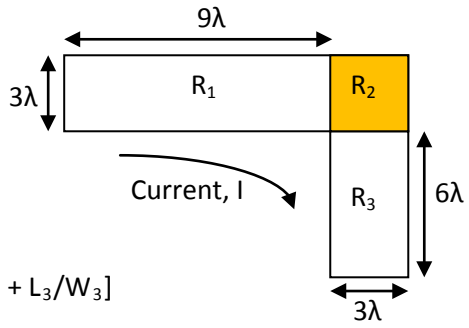
Given (Table 3-14)

<u>Reference Area</u>

= Area of minimum-sized transistor
= 1*1 or 2λ*2λ

<u>Relative Capacitance (Rel Cap)</u>

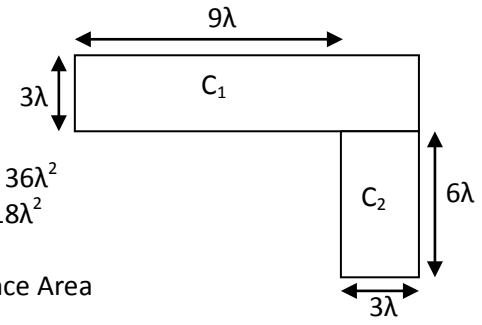
Given (Table 3-16)



Given : $R_{s(poly)} = 7.5\ \Omega/\square$

$R = R_1 + ½(R_2) + R_3$
$= R_{S(poly)}\ [L_1/W_1 + ½(L_2/W_2) + L_3/W_3]$

$R = 7.5 * [9/3 + ½(3/3) + 6/3]$
$= 41.25\ \Omega$



Given : $Rel\ Cap_{(poly)} = 1$

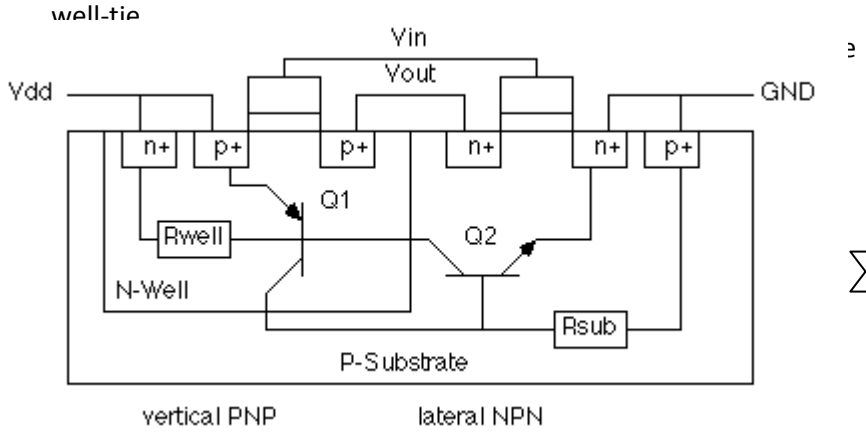
Actual area of $C_1 = L_1 * W_1 = 12λ * 3λ = 36λ^2$
Actual area of $C_2 = L_2 * W_2 = 6λ * 3λ = 18λ^2$

Rel area of C_1 = Actual Area / Reference Area
$= 36λ^2 / 4λ^2 = 9$
Rel area of $C_2 = 18λ^2 / 4λ^2 = 4.5$

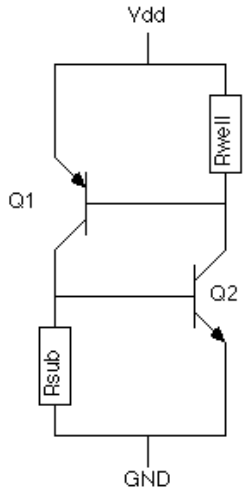
C_1 = Rel Area * Rel Cap * $\square$Cg = 9 * 1 * $\square$Cg = $9\square C_g$
C_2 = Rel Area * Rel Cap * $\square$Cg = 4.5*1* $\square$Cg = $4.5\square C_g$

$C = C_1 + C_2 =$ **$13.5\square C_g$**

3.5.4 CMOS Latchup Pg 31



Parasitic Structure which can cause latchup



Latch-up Equivalent Circuit

A byproduct of the Bulk CMOS structure is a pair of parasitic bipolar transistors.

Latch-up :

(i) is the inadvertent creation of a low-impedance path between V_{dd} and GND of a CMOS circuit when the parasitic structure is activated.

(ii) will result in circuit malfunction or destruction of the device if a large runaway (positive feedback → $\beta_1 \times \beta_2 > 1$) current is established.

It can be minimized through fabrication and design approaches.

Latch up in CMOS circuits can be reduced by :-

(i) increasing the distance between N-well and N+ source/drain diffusion of the NMOS device ⇒ reduction of β_2.

(ii) using guard rings around the transistors to provide low resistance path to collect the carriers before they interfere with the operation, thus reduces parasitic resistance.

(iii) placing a well tie and a substrate tie close to the supplies to reduce R_{sub} and R_{well} ⇒ higher accidental current is required to cause latch-up.

(iv) increase substrate doping level ⇒ reduction in R_{sub}

CMOS Characteristics Summary Pg 21

(i) Power Consumption

In a steady state,

- only half of the circuit is conducting thus no closed path between V_{dd} and V_{ss}
- since no current flows can flow, hence no static power dissipation

In transition state,

- there is a window where the p- and n-transistors are turned on concurrently, so a current can flow between V_{dd} and V_{ss}
- this short-circuit current contributes to the dynamic power dissipation in the CMOS circuits

Dynamic power dissipation is less significant compared to the power consumed during charging or discharging of load capacitance, C_L.

(ii) Ratio-less Device

V_{OUT} of the CMOS circuits is able to give a full voltage swing from V_{dd} to V_{ss} regardless of the relative dimensions of the PMOS and NMOS transistors. Thus CMOS devices are sometimes known as ratio-less devices.

(iii) Timing Performance

Ratios introduced in CMOS gates are for the purpose of improving the timing aspects of the gates. Timing performance is affected by the size of the transistors and the number of transistor turned ON, to achieve the desired output state.

(iv) Threshold Voltage

This voltage is not a constant but is affected by

(i) body effect (occurs when $V_{sb} \neq 0$)
- depletion region increases
- additional charges formed in depletion region resulting in decrease in channel charges
- must increase V_{gs} to maintain the channel conductivity

(ii) hot carrier effect

trapping of high energy (hot) carriers in the gate oxide

(v) Latch Up

An undesirable creation of a low-impedance path between V_{dd} and GND of a CMOS circuit when the parasitic bipolar structure is activated. It will result in circuit malfunction or destruction of the device if a large runaway current is established.

It can be minimized through fabrication and design approaches.