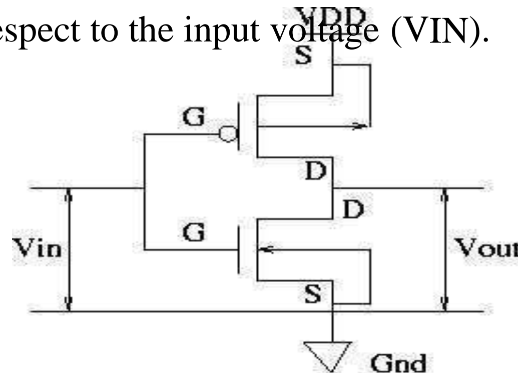
DC characteristics of MOs transistor

- A complementary CMOS inverter consists of a p-type and an n-type device connected in series.
- The DC transfer characteristics of the inverter are a function of the output voltage (Vout) with respect to the input voltage (VIN).



www.Binils.com

Fig 1.4.1MOS Transistor

[Source: Neil H.E. West, CMOS VLSI Design …]

- The MOS device first order Shockley equations describing the transistors in cut-off, linear and saturation modes can be used to generate the transfer characteristics of a CMOS inverter.
- Plotting these equations for both the n- and p-type devices produces the traces below.

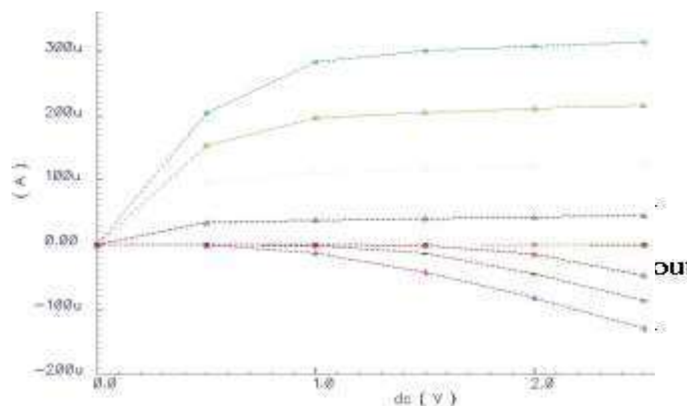


Fig 1.4.2 MOS Transistor IV Characteristics

[Source: Neil H.E. West, CMOS VLSI Design …]

- The DC transfer characteristic curve is determined by plotting the common points of Vgas intersection after taking the absolute value of the p-device IV curves, reflecting

  Them about the x-axis and superimposing them on the n-device IV curves.
- We basically solve for Vin(n-type) = Vin(p-type) and Ids(n-type)=Ids(p-type)
- The desired switching point must be designed to be 50 % of magnitude of the supply voltage i.e. VDD/2.
- Analysis of the superimposed n-type and p-type IV curves results in five regions in which the inverter operates.

- **Region A** occurs when 0 leavein let VT (n-type).
  - The n-device is in cut-off (Ids =0).
  - p-device is in linear region,
  - Ids = 0 therefore -Ids = 0
  - Vds = Vout – VDD, but Vdsp =0 leading to an output of Vout = VDD.
- **Region B** occurs when the condition VTN let VIN le VDD/2 is met.
  - Here p-device is in its non-saturated region Ds new 0.
  - n-device is in saturation current

IDs is obtained by setting Vgas = Vin resulting in the equation:

- In **region B** Ido is governed by voltages Vgas and Vds described by:
- Saturation currents for the two devices are:
- **Region D** is defined by the inequality
- P-device is in saturation while n-device is in its non-saturation region.

www.Binils.com

- Equating the drain currents allows us to solve for Out. (See supplemental notes for algebraic manipulations).
- In **Region E** the input condition satisfies:
- The p-type device is in cut-off: Idip=0
- The n-type device is in linear mode
- Vgap = Vin –VDD and this is a more positive value compared to Vtp.
- Vout = 0

www.Binils.com

Download Binils Android App in Playstore Download Photoplex App

Stick diagrams:

Stick diagrams are used to convey layer information through the use of a color code for example in NMOS design.

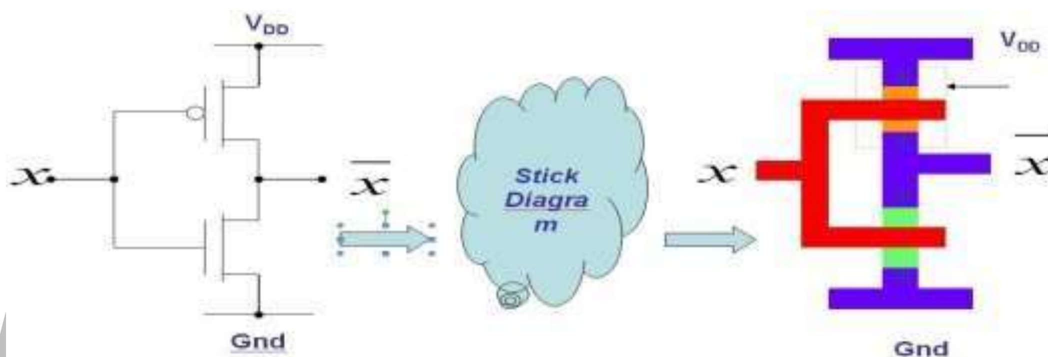
- Green for n- diffusion
- Red for poly silicon
- Blue for metal
- Yellow for implant
- Black for contact areas
- The designer can draw a layout using colored lines to represent the various process layers such as diffusion, metal and poly silicon.

Where poly silicon crosses the diffusion, transistors are created and where metal wires join diffusion or poly silicon, contacts are formed.

A stick diagram is a cartoon of a chip layout. They are not the exact models of layout

www.binils.com

- The stick diagram represents the rectangles with lines which represents wires are component symbols.
- The color cooling has been complemented by monochrome encoding of the lies so the black and white copies of stick diagrams do not lose the layer information.
- The color and monochrome encoding scheme used has been evolved to cover NMOS and CMOS processes.
- To illustrate the stick diagram inverter circuits are presented below in NMOS, and in P well CMOS technology.



www.binils.com

Figure 1.2.1 : Stick diagrams

[Source: Wayne Wolf, ―Modern VLSI Design: System On Chip ….]

- Having conveyed layer information and topology by using stick or symbolic diagrams. These diagrams relatively easily turned into mask layouts.
- The below diagram stressing the ready translation into mask layout form. In order that the mask layout produced during design will be compatible with the fabrication process.

  A seer of design rules are set out for layouts.

Stick diagram using NMOS Design:

We consider single metal, single poly silica NMOS technology. The layout of NMOS involves.

- N-diffusion and other thin oxide regions- green
- Poly silicon - red
- Metal -blue
- IM pant -yellow
- Contacts - black or brown

A transistor is formed wherever poly silicon crosses n-diffusion and all diffusion wires are n-type. The various steps involved in the design style are.

Step1: Draw the metal VDD and GND rails in parallel allowing enough space between them for the other circuit element which will be required.

Step 2: Draw the thin ox paths between the rails for inverters and inverter based logic.

Step 3: Draw the pull up structure which comprises a depletion mode transistor interconnected between the output point and VDD.

www.binils.com

Step 4: Draw the pull down structure comprising an enhancement mode structure interconnected between the output point and GNO.

Step 5: Signal paths may be switched by pass transistor, and along signal paths often require metal buses.

Design Rules and layout:

The design rules primarily address two issue

1) The geometrical reproduction of features that can be reproduced by the mask-making and lithographical process.
2) The interactions between different layers. There are several approaches that can be taken in describing the design rules. These include

- Micron design rules:

- Stated at some micron resolution
- Usually given as a list of minimum feature sizes and spacing's for all masks required in a given process.
- Normal style for industry.

- Lambda () based design rules

- These rules popularized by Mead and Conway are based on a single parameter, which Characterized the linear feature- the resolution of the complete wafer implementation process – and permits first order scaling.
- They have been widely used, particularly in the educational context and in the design of multi project chips.

Layout () based design Rules:

The lambda, design rules are bases on mead and Conway work and in general, design rules and layout methodology are based on the concept of which provides a process and feature size. Independent way of making mask dimensions to scale.

- All paths in all layers will be dimensioned in units and sub-Sequent can be allocated an appropriate value compactible with the feature size of the fabrication process.
- Design rules can be conveniently set out in diagrammatic form as shown below.

www.binils.com

Contact cuts:

The contacts between layers are set out as shown below. Here it will be observed that connection can be made between two or, in the case of NMOs design, three layers.

1) Metal to poly silicon or to diffusion

There are three possible approaches for making contacts between poly silicon and diffusion in NMOS circuits. There are

i) Poly silicon to metal then metal to diffusion
ii) Buried contact poly silicon to diffusion
iii) Butting contact.

- The 2 x2 contact cut indicates and area in which the oxide is to be removed down to the underlying poly silicon or diffusion surface.
- When the deposition of the metal layer takes place, the metal is deposited through the contact cut areas on to the underlying areas so that contact is made between

The layers.

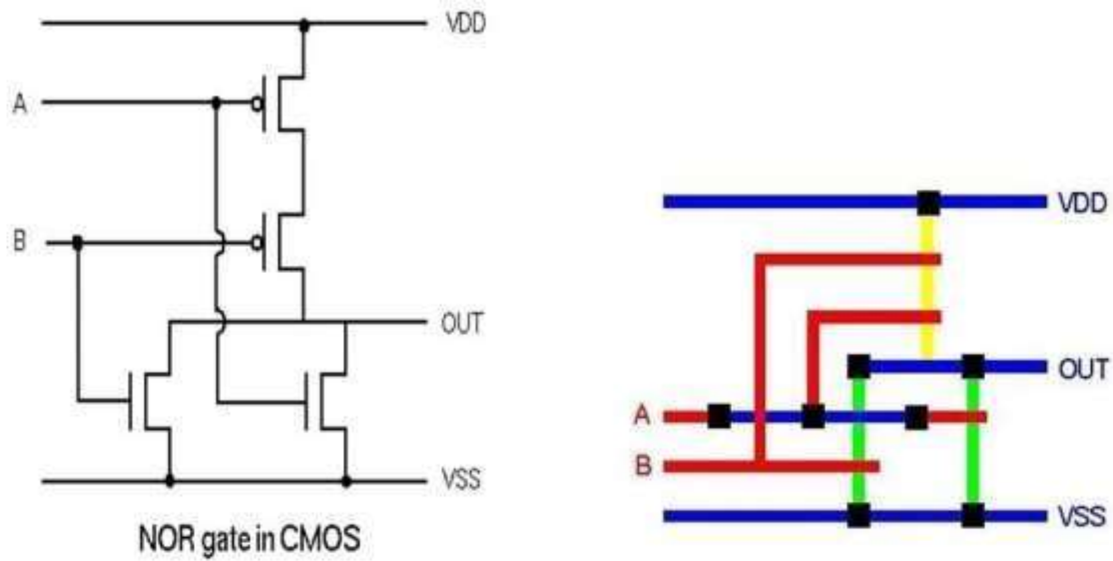


Figure 1.2.2: Stick diagrams

[Source: Neil H.E. Waste, David Money Harris ―CMOS VLSI Design ….]

www.binils.com

The non- ideal IV effect include the following:-

1) Velocity saturation & mobility degradation.
2) Channel length modulation
3) Body effect
4) Sub threshold condition
5) Junction Leakage
6) Tunneling
7) Temperature dependence
8) Geometry Dependence.

Velocity saturation and mobility degradation:-

- Carrier drift velocity and current increase linearly with the lateral field E let = Vds/ L between source and drain.
- At high field strength, drift velocity rot off due to carrier scattering and usually saturates at V sat.

www.binils.com

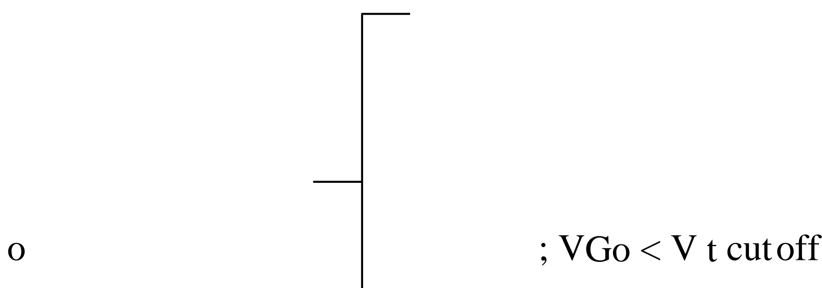
- Without velocity saturation the saturation current is

$$I_{ds} = \square C_0 \times \frac{W}{L} \frac{(V_{s.} - V_{ds})^2}{2}$$

- If the transistor is completely Velour saturated V = Vsat and saturation current become.

- Ids = C_0 x W (Vgs - V t) V sat without velocity saturation

Drain current is quodratically dependent on voltage



o ; VGo < V t cutoff

Ids = I Sat = V ds ; V ds < V sat linear

I Sat ; V ds > V sat saturation.

Where

IV dsat = Pave (Vs. - V t) ∝/2.

- As channel length becomes shorter, the lateral field increases and transistors become more velocity saturated, and the supply voltage is held constant.

Channel Length Modulation:-

- Ideally Ids is independent of VDs in saturation.
- The reverse biased p-n junction between the drain and body forms a depletion region with a width LD that increases with VdB.
- The depletion region effectively shortens the channel length to Luff =L-Ld.
- Imagine that the source voltage is close to the body voltage. Increasing VDs decreases the effective channel length. Shorter length results in higher current. Thus Ids increases with VDs in saturation as shown below.

In saturation region

Where □ = Channel length modulation factor

$$I_{ds} = \frac{□}{2} (V_s - V_t)^2 (1 + □ V_{ds})$$

www.binils.com

Body Effect:

Transistor has four terminals named gate, source, drain and body. The potential difference between the source and body Vsub affects the threshold voltage.

VT = Vto +ȣ ($\sqrt{ps + V_{sb}}$ - $\sqrt{\varphi s}$)

Where

VTo = Threshold Voltage when the source is at the body potential

□ s = Surface Potential at threshold = 2vT in

ND Ni

VSub = Potential difference between the source and body.

Sub threshold condition:

- Ideally current flows from source to drain when Vgs > V t. In real transistor, current does not abruptly cut off below threshold, but rather drops off exponentially as

Ids = I do e Vgo - V t [1- e V ds] Vt.

This is also called as leakage and often this results in underwired current when a transistor is normally OFF. Ido is the current at threshold and is dependent on process and device geometry

Applications:-

- This is used in very low power analog circuit
- This is used in dynamic circuits and OR AM

Advantage:

1) Leakage increases exponentially as VT decreases or as temperature rises.

Disadvantages:

1) It becomes worse by drain induced barrier lowering in which a positive VDs effectively reduces VT This effect is especially pronounced in short channel transistors.

Junction Leakage:

www.binils.com

- The P-n junction between diffusion and the substrate or well form diodes are shown below.
- The substrate and well are tied to GND or VDD to ensure that these diodes remain reverse biased.

The reverse biased diodes still conduct a small amount of current I do. ID = Is [e VD -1] VT

Where

ID = diode current

Is = diode reverse- biased saturation current that depends ondoping levels and on the area and perimeter of the diffusion region.

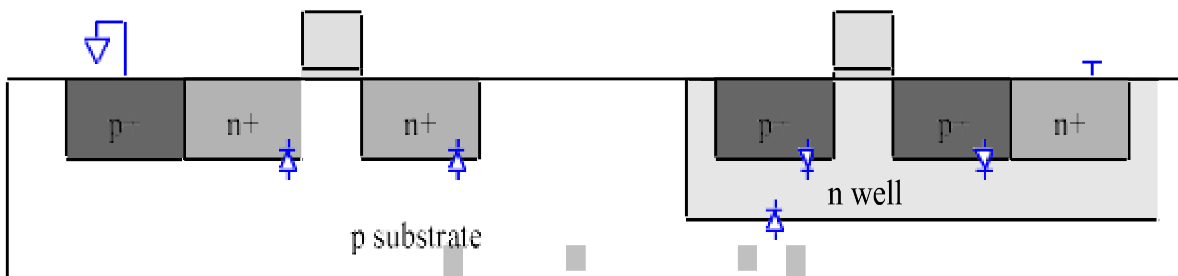


Figure 1.3.1: Junction Leakage

[Source: Sung-Mo kanga, Yusuf leblebici, Charlwood Kim —CMOS Digital Integrated Circuits: Analysis & Design….]

Tunneling:

Based on quantum mechanics, we see that the is a finite probability that carriers will tunnel through the gate oxide. This results in gate leakage current flowing into the gate. The probability of tunneling drops off exponentially with oxide thickness.

- Large tune ling currents impact not only dynamic nodes but also quiescent power consumption and thus may limit oxide thickness tor.
- Tunneling can purposely be used to create electrically erasable memory devices. Different dielectrics may have different tunneling properties.

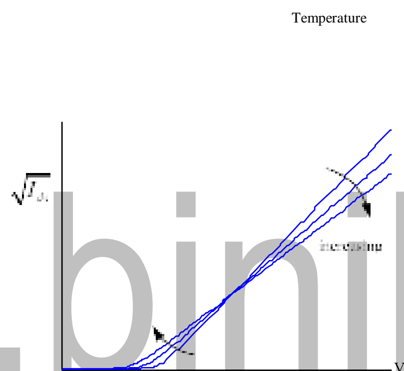
Temperature Dependence:

Temperature influences the characteristics of transistors. Carrier mobility decreases with temperature.

μ (T) = μ (TR) (T) $^{-k}$ μ T

www.binils.com

- Junction leakage increases with temperature because. Is is strongly temperature dependent. The combined temperature effect is shown below.
  Where on current decreases and OFF current increases with temperature.
  The figure below shows how the on current Isat decreases with temperature. Circuit performance is worst at high temperature, called negative temperature coefficient.
- Circuit performance can be improved by cooling. Natural convection, fans with heat sink, water cooling thin film refrigerators, and liquid nitrogen can be used as cooling.



gs

www.binils.com

Figure 1.3.2: Temperature Dependence

[Source: Sung-Mo kanga, Yusuf leblebici, Charlwood Kim —CMOS Digital Integrated Circuits: Analysis & Design….]

Advantages of Operating at low temperature:

1) Velocity saturation occurs at higher fields providing more current.
2) For high mobility, power is saved.
3) Wider depletion region results in less junction capacitance.

Geometry Dependence:

- The layout designer draws transistors with width and length W drawn and L drawn. The actual gate dimensions may differ by factors XW and XL.
- The source and drain trends to diffuse later under the gate by LDi producing a shorter effective

Between source and drain.

Luff = L drawn + XL -2LP

Weft = W drawn + XW – 2wD

Long transistors experience less channel length modulation. In a process blow 0.25 m the effective length of the transistor depends on the orientation of the transistor.

Ideal I-V Characteristics of a MOs and MOs Device

MOS transistors have three regions of operation:

_ Cutoff or sub threshold region

_ Linear region

_ Saturation region

The current and voltage (I-V) for a no transistor in each of these regions. The model assumes that the channel length is long enough that the lateral electric field (the field between source and drain) is relatively low, which is no longer the case in nanometer devices. This model is variously known as the *long-channel*, *ideal*, *first-order,* or *Shockley* model. Subsequent sections will refine the model to reflect high fields, leakage, and other no idealities. The long-channel model assumes that the current through an OFF transistor is 0. When a transistor turns ON (*Vs.* > *VT*), the gate attracts carriers (electrons) to form a channel. The electrons drift from source to drain at a rate proportional to the electric field between these regions. Thus, we can compute currents if we know the amount of charge in the channel and the rate at which it moves.

We know that the charge on each plate of a capacitor is *Q* = *CV*. Thus, the charge in the channel *C*hannel is

www.binils.com

QChannel=Cg (Vgo-Vt)

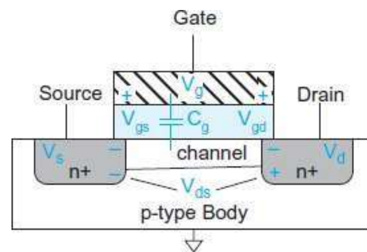


Figure 1.3.3: Ideal I-V Characteristics

[Source: Sung-Mo kanga, Yusuf leblebici, Charlwood Kim —CMOS Digital Integrated Circuits: Analysis & Design….]

Where *Cg* is the capacitance of the gate to the channel and *Vic* - *VT* is the amount of voltage attracting charge to the channel beyond the minimum required to invert from p ton. The gate voltage is referenced to the channel, which is not grounded. If the source is at *Vs.* and the drain is at *Veda,* the average is *Vic* = (*Vs.* + *Veda*)/2 = *Vs.* + *Vs.* /2. Therefore, the mean difference between the gate and channel potentials *Vic* is *Vg* – *Vic* = *vs.* – *Vs.* as shown in Figure.

www.binils.com

We can model the gate as a parallel plate capacitor with capacitance proportional to area over thickness. If the gate has length *L* and width *W* and the oxide thickness is *t*ax, as shown in below Figure, the capacitance is

Cg=ox (WL/tox) =Cowl

where ox is the permittivity of free space, 8.85 × 10–14 F/cm, and the permittivity of SiO2 is *o*x = 3.9 times as great. Often, the ox/*t*ax term is called *C*ox, the capacitance per unit area of the gate oxide.

Each carrier in the channel is accelerated to an average velocity, *v*, proportional to the lateral electric field, i.e., the field between source and drain. The constant of proportionality μ is called the *mobility*.

$v = \mu E$

The time required for carriers to cross the channel is the channel length divided by the carrier velocity: L/v. Therefore, the current between source and drain is the total amount of charge in the channel divided by the time required to cross

$$I_{ds} = \frac{Q_{channel}}{L/v}$$
$$= \mu C_{ox} \frac{W}{L}\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$
$$= \beta\left(V_{GT} - V_{ds}/2\right)V_{ds}$$

$$\beta = \mu C_{ox}\frac{W}{L};\ V_{GT} = V_{gs} - V_t$$

The term *Vs.* – *VT* arises so often that it is convenient to abbreviate it as *VGT*.

K' = *co*d

If *Vs.* > *V*isit *VGT*, the channel is no longer inverted in the vicinity of the drain; we say it is pinched off. Beyond this point, called the *drain saturation voltage*, increasing the drain voltage has no further effect on current. Substituting *ds* = *V*isit at this point of maximum current in above eon, we find an expression for the saturation current that is independent of *Vs.*

Ids= (β/2) V2GT

This expression is valid for *Vs.* > *VT* and *Vs.* > *V*isit. Thus, long-channel MOS transistors are said to exhibit *square-law behavior* in saturation. Two key figures of merit for a transistor are *I*on and *I* off. *I*on (also called *I*D sat) is the ON current, *Ids*, when *Vs.* = *Vs.* = *VDD*. *I* off is the OFF current when *vs.* = 0 and *Vs.* = *VDD*. According to the long-channel model, *I* off = 0 and

Ion= (β/2) (Vdid-Vt)

www.binils.com

Below fig shows the I-V characteristics for the transistor. According to the first-order model, the current is zero for gate voltages below *VT* For higher gate voltages, current increases linearly with *vs.* for small *ds.* As *ds* reaches the saturation point *V*isit = *VGT*, current rolls off and eventually becomes independent of *ds* when the transistor is saturated. We will later see that the Shockley model overestimates current at high voltage because it does not account for mobility degradation and velocity saturation caused by the high electric fields.

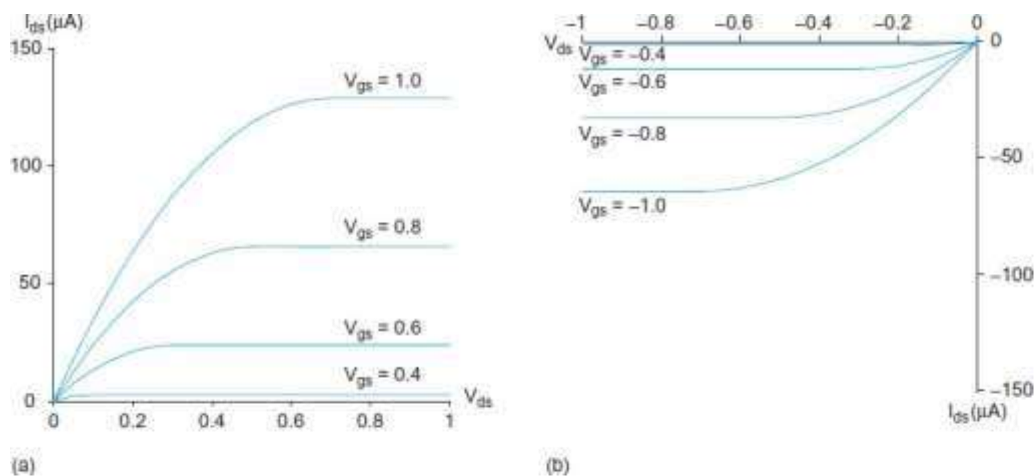


Figure 1.3.4: I-V Characteristics

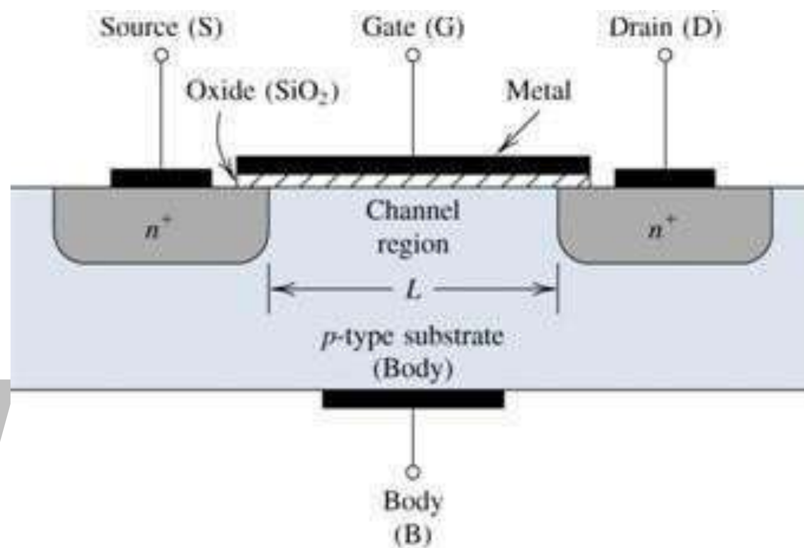
[Source: Sung-Mo kanga, Yusuf leblebici, Charlwood Kim —CMOS Digital Integrated Circuits: Analysis & Design….]

MOS TRANSISTOR

nMOS Transistor

NMos transistors are built on a p-type substrate of moderate doping. Source and drain are formed by diffusing heavily doped n-type impurities (n+)adjacent to the gate. A layer of silicon dioxide (SiO_2) or glass is place over the substrate in between the source and drain. Over SiO_2, a layer of polycrystalline silicon or polysilicon is formed, from which the gate terminal is taken.

The following figure shows the structure and symbol of nMOS transistor.R



www.binils.com

Fig: 1.1.1 no transistor.

[Source: Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

Threshold Voltage (V_t)

It can be defined as the voltage applied between the gate and the source of a MOS device (V_{gs}) below which the drain-to-source current (I_{ds}) "effectively" drops to zero. V_t depends on the following:

- Gate conduction material
- Gate insulation material
- Gate insulator thickness
- Channel doping
- Impurities at the silicon-insulator interface
- Voltage between the source and the substrate, V_{sb}.

Modes of operation of MOS Transistor:

The following are the three modes of operation of nMOS transistor:

1. Accumulation mode
2. Depletion mode
3. Inversion mode

a. Accumulation Mode

In this mode a negative voltage is applied to the gate. So there is negative charge on the gate. The mobile positively charged holes are attracted to the region beneath the gate.

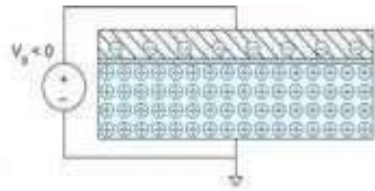


Fig: 1.1.2 Accumulation Mode.

[Source: Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

b. Depletion Mode:

www.binils.com

In this mode a low positive voltage is applied to the gate. This results in some positive charge on the gate. The holes in the body are repelled from the region directly beneath the gate.

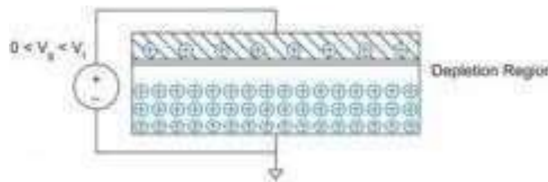


Fig: 1.1.3 Depletion Mode.

[Source: Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

c. Inversion Mode:

In this mode, a higher positive potential exceeding a critical threshold voltage is applied. This attracts more positive charge to the gate. The holes are repelled further and a small number of free electrons in the body are attracted to the region beneath

the gate. This conductive layer of electrons in the p-type body is called the inversion layer.

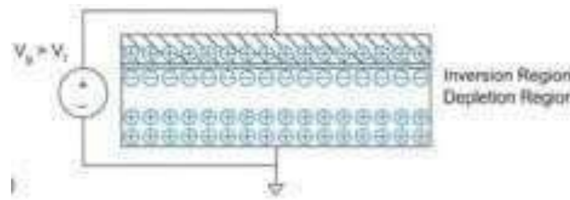


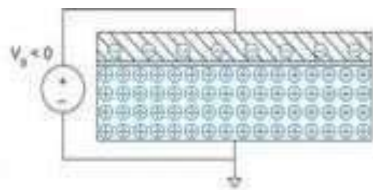
Fig: 1.1.4 Inversion Mode.

[Source: Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

Behavior of nMOS with different voltages:

The Behavior of nMOS with different voltages can be classified into the following three cases and is illustrated in below figure.

i. Cut-off region



ii. Linear region

iii. Saturation region

www.binils.com

a. Cut-off region:-

In this region $V_{gs} < V_t$ .The source and drain have free electrons. The body has free holes but no free electrons. The junction between the body and the source or drain is reverse biased. So no current will flow. This mode of operation is called cut-off.

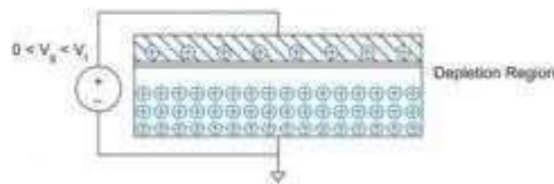


Fig: 1.1.5 Cut off region.

[Source: R.Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

Linear region:-

In this region $V_{gs} > V_T$ .Now an inversion region of electrons called the channel connects the source and drain. This creates a conductive path between source and drain. The number of carriers and the conductivity increases with the gate voltage. The potential difference between drain and source is $V_{ds} = V_{gs} - V_{gd}$. If $V_{ds}=0$, there is no electric field tending to push current from drain to source.

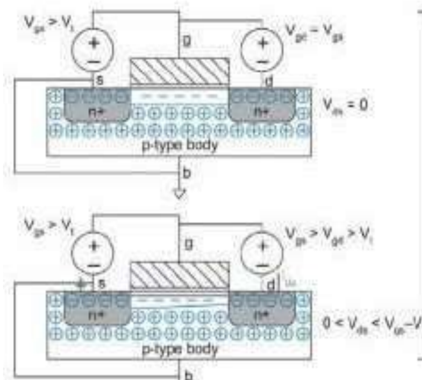


Fig: 1.1.6 linear region.

www.binils.com

[Source: Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

b. Saturation region:-

In this region V_{ds} becomes sufficiently larger than $V_{gd} < V_T$, the channel is no longer inverted near the drain and becomes pinched off .Above this drain voltage, the I_{ds} is controlled only by the gate voltage. This mode is called saturation mode.

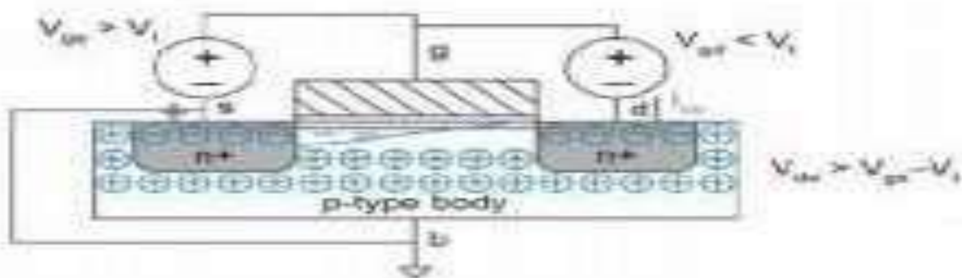
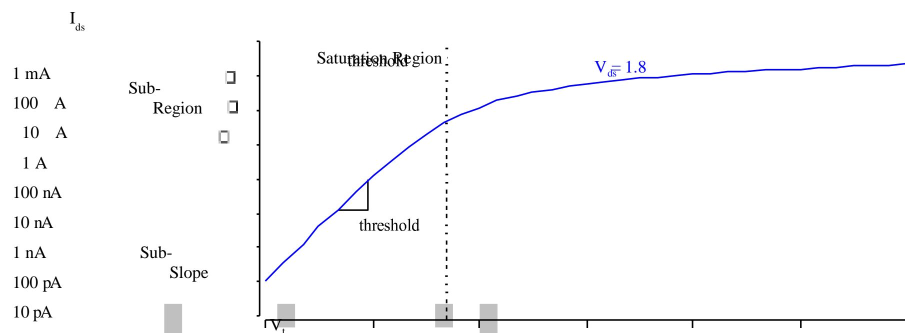


Fig: 1.1.7 Saturation region.

[Source: R.Jacob Baker, CMOS Circuit Design, Layout and Simulation...]

Non- Ideal I-V Effects:

The Ids value of an ideal I- v model neglects many effects that are important to modern devices.



www.binils.com

Fig: 1.1.8 Non ideal IV Model.

[Source: Neil H.E. Weste, David Money Harris —CMOS VLSI Design...]

Simulated I-V Characteristics

- While compared to the ideal devices, the saturation current increases less than a quarter with increasing V_{gs}. This is caused by two effects.
  1) Velocity Saturation
  2) Mobility degradation.
- At high lateral field strengths Ds. Carrier velocity ceases to increase L linearly with field strength. This is called velocity saturation and this results in lesser I_{ds} than expected at high V_{Ds}.
- Current between source and drain is the total amount of charge in the channel divide the time required to cross it.

$$I_{ds} = \frac{Q\ Channel}{V\ t} = \frac{Q\ Channel}{(\square/\square)} = \frac{Q\ Channel\ *}{L}$$

By Sub the values we get

$$I_{ds} = \square C \square \square \frac{W}{L} (V_{gas} - V_t - V_{ds)}) V_{ds}$$

$$Ids = \square (V_{gs} - V_t - \frac{V_{ds)}}{2}) V_{ds}$$

Where $\square = \square C_0 \times W/L$

In linear region $V_{as} > V_t$ and V_{ds} is relatively small.

Saturation Region:-

- In saturation region, if $V_{ds} > V_{sat,}$ the channel is pinched off. i.e.; $V_{Ds} = V_{as} - V_t$
- Beyond this point it is often called the drain saturation voltage. Sub

$V_{ds} = V_{as} - V_t$ in the Ids values for linear region we get.

$$I_{ds} = \frac{\square}{2} (V_{as} - V_{t)})^2 \text{ for } V_{ds} > V_{sat.}$$

In saturation, I_{sat} is

$V_{as} = V_{ds} + V_{DD}$

www.binils.com

$I_{sat} = \quad (V_{DD} - V_t)^2$

Summarizing the three regions we get.

$$I_{ds} = \begin{cases} 0 & ; V_{gs} < V_t ; \text{cut off} \\ \square\,(V_{gs} - V_t - \frac{V_{ds}}{2})\,V_{ds} & ; V_{ds} < V_{dsat} ; \text{Linear} \\ \frac{\square}{2}(V_{gs} - V_t)^2 & ; V_{ds} > V_{dsat} ; \end{cases}$$

www.binils.com

www.binils.com for Anna University | polytechnic and schools

- At high vertical field strengths V_{gs} / tor the carrier scatters more often. This is called mobility degradation and this leads to less current than expected at high V_{gs}
- The threshold voltage itself is influenced by the voltage difference between the source and body called the body effect.

Download Binils Android App in Playstore Download Photoplex App

www.binils.com

The propagation delay. Delay Estimations: -

In most designs, there exist many logic paths called critical paths. These paths are recognized by a timing analyzer or circuit simulator. Critical paths are affected by the following four levels.

i. Architectural level
ii. Logic level
iii. Circuit level
iv. Layout level

Propagation delay time (Tpd) or max time is the maximum time from the input crossing 50%to the output crossing 50%.The delay can be estimated by the following ways,

i. RC delay models
ii. Linear delay models
iii. Logic efforts
iv. Parasitic delay

www.binils.com

1. **RC delay models:**

The delay of logic gate is computed as the product of RC\where R is the effective driver resistance and C the load capacitance. Logic gates use minimum –length devices for least delay, area and power consumption. The delay of a logic gate depends on the transistor width in the gate and the capacitance of the load.

Effective Resistance and Capacitance:

An NMOS transistor with width of one unit has effective Esistance R An PMOS transistor with width of one unit has effective resistance 2R Capacitance consists of gate capacitance cg and source/diffusion capacitance c diff .in most processes cg is equal to c diff , c g and cdiff are proportional to transistor width.

Diffusion capacitance layout effects:

Download Binils Android App in Playstore Download Photoplex App

To reduce the diffusion capacitance in the layout, diffusion nodes are shared. Un contacted nodes have less capacitance. Diffusion capacitance depends on the layout.

2. **Elmore delay model:**

Elmore delay model estimates the delay of an RC ladder .this is equal to the sum over each node in the ladder of the resistance between the node and supply multiplied by capacitance on the node.

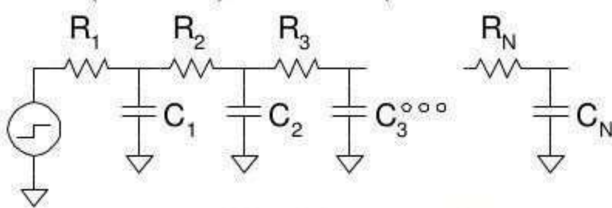


Fig 1.5.1: RC ladder for Elmore Delay Model

[Source: Neil H.E. Weste, David Money Harris —CMOS VLSI Design..]

www.binils.com

3. **Linear delay model:**

The propagation delay of a gate is d,

$$d= f + p$$

F= effort delay or state effort, which depends on the complexity and fan-out of the gate.

P=parasitic delay

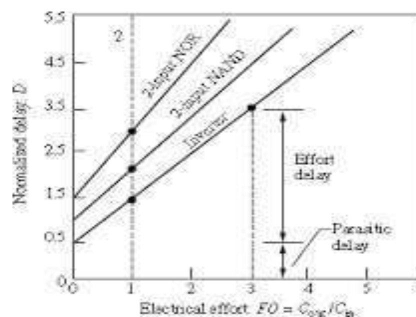


Fig 1.5.2: Normalized delay vs. fan-out

[Source: Neil H.E. Weste, David Money Harris —CMOS VLSI Design..]

Download Binils Android App in Playstore Download Photoplex App

Logical effort:

Logical effort is defined as the ratio of the input capacitance of the gate to the input capacitance of an inverter that delivers the same output current.

Parasitic delay:

Parasitic delay is defined as the delay of the gate when it drives zero load. This can be estimated with RC delay models. The inverter has 3 units of diffusion capacitance on the output.

Gate type	Number of input				
	1	2	3	4	n
INVERTER	1				
AND		2	3	4	N
NOR		2	3	4	n
TRISTATE,MULTIPLEX	2	4	6	8	2

www.binils.com

Logical effort and transistor sizing:

Logical effort provides a simple method to choose the best topology and number of stages of logic for a function. This quickly estimates the minimum possible delay for the given topology and to choose gate sizes that achieve this delay.

Delay in multistage logic networks:

Logical effort is independent of size and electrical effort is dependent on size.

1. path logical effort

2. path electrical effort

Download Binils Android App in Playstore Download Photoplex App

3. path effort
4. branching effort
5. path branching effort
6. path delay
7. minimum possible delay

Choosing the best number of stages:

Inverters can be added at the end of a path without changing its function. Extra inverters and parasitic delay, but do not change the path logical effort.

www.binils.com

Device Modeling SPICE provides a wide variety of MOS transistor models with various trade-offs between complexity and accuracy. Level 1 and Level 3 models were historically important, but they are no longer adequate to accurately model very small modern transistors. BSIM models are more accurate and are presently the most widely used. Some companies use their own proprietary models. This section briefly describes the main features of each of these models. It also describes how to model diffusion capacitance and how to run simulations in various process corners. The model descriptions are intended only as an overview of the capabilities and limitations of the models; refer to a SPICE manual for a much more detailed description if one is necessary Level 1 Model the SPICE Level 1, or Shih man-Hodges Model [Shichman68] is closely related to the Shockley model described in EQ (2.10), enhanced with channel length

Download Binils Android App in Playstore Download Photoplex App

www.binils.com for Anna University | polytechnic and schools

www.binils.com

Download Binils Android App in Playstore Download Photoplex App

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t & \text{cutoff} \\ \text{KP}\frac{W_{\text{eff}}}{L_{\text{eff}}}(1+\text{LAMBDA}\times V_{ds})\left(V_{gs}-V_t-\frac{V_{ds}}{2}\right)V_{ds} & V_{ds} < V_{gs}-V_t & \text{linear} \\ \frac{\text{KP}}{2}\frac{W_{\text{eff}}}{L_{\text{eff}}}(1+\text{LAMBDA}\times V_{ds})\left(V_{gs}-V_t\right)^2 & V_{ds} > V_{gs}-V_t & \text{saturation} \end{cases}$$

The parameters from the SPICE model are given in ALL CAPS. Notice that is written instead as KP(*W*eff /*L*eff), where KP is a model parameter playing the role of *k*. *W*eft and *L*uff are the effective width and length). The LAMBDA term (LAMBDA = 1/*VA*) models channel length modulation The threshold voltage is modulated by the source-to- body voltage *Vs*. through the body effect.

The gate capacitance is calculated from the oxide thickness TOX. The default gate capacitance model in HSPICE is adequate for finding the transient response of digital circuits. More elaborate models exist that capture nonreciprocal effects that are important for analog design. Level 1 models are useful for teaching because they are easy to correlate with hand analysis, but are too simplistic for modern design.

www.binils.com

Level 2 and 3 Models

The SPICE Level 2 and 3 models add effects of velocity saturation, mobility degradation, sub threes hold conduction, and drain-induced barrier lowering. The Level 2 model is based on the Grove-Frohman equations, while the Level 3 model is based on empirical equations that provide similar accuracy, faster simulation times, and better convergence. However, these models still do not provide good fits to the measured I-V characteristics of modern transistors.

BSIM Models

The Berkeley Short-Channel IGFET1 Model (BSIM) is a very elaborate model that is now widely used in circuit simulation. The models are derived from the underlying device physics but use an enormous number of parameters to fit

the behavior of modern transistors. BSIM versions 1, 2, 3v3, and 4 are implemented as SPICE levels 13, 39, 49 and 54 respectively.

Features of the model include:

- Continuous and differentiable I-V characteristics across sub threes hold, linear, and saturation regions for good convergence
- Sensitivity of parameters such as *Vt* to transistor length and width
- Detailed threshold voltage model including body effect and drain-induced barrier lowering
- Velocity saturation, mobility degradation, and other short-channel effects
- Multiple gate capacitance models
- Diffusion capacitance and resistance models
- Gate leakage models

As the BSIM models are so complicated, it is impractical to derive closed-form equations for propagation delay, switching threshold, noise margins, etc., from the underlying equations. However, it is not difficult to find these properties through circuit simulation. Device characterisation will show simple simulations to plot the device characteristics over the regions of operation that are interesting to most digital designers and to extract effective capacitance and resistance averaged across the switching transition. The simple RC model continues to give the designer important insight about the characteristics of logic gates.

www.binils.com

Diffusion Capacitance Models

The p–n junction between the source or drain diffusion and the body forms a diode. We depends on the area and perimeter of the diffusion. HSPICE provides a number of methods to specify this geometry, controlled by the ACM (Area Calculation Method) parameter, which is part of the transistor model. have seen that the diffusion capacitance determines the parasitic delay of a gate and The diffusion capacitance model is common across most device models including

Levels 1–3 and BSIM. By default, HSPICE models use ACM = 0. In this method, the designer must specify the area and perimeter of the source and drain of each transistor.

The SPICE models also should contain parameters CJ, CJSW, PB, PHP, MJ, and MJSW. Assuming the diffusion is reverse-biased and the area and perimeter are specified, the diffusion capacitance between source and body is computed as described in

$$C_{sb} = \text{AS} \times \text{CJ} \times \left(1 + \frac{V_{sb}}{\text{PB}}\right)^{-MJ} + \text{PS} \times \text{CJSW} \times \left(1 + \frac{V_{sb}}{\text{PHP}}\right)^{-MJSW}$$

The drain equations are analogous, with S replaced by D in the model parameters. The BSIM3 models offer a similar area calculation model (ACM = 10) that takes into account the different sidewall capacitance on the edge adjacent to the gate. Note that the PHP parameter is renamed to PBSW to be more consistent.

www.binils.com

$$C_{sb} = \text{AS} \times \text{CJ} \times \left(1 + \frac{V_{sb}}{\text{PB}}\right)^{-MJ} + (\text{PS} - W) \times \text{CJSW} \times \left(1 + \frac{V_{sb}}{\text{PBSW}}\right)^{-MJSW} + W \times \text{CJSWG} \times \left(1 + \frac{V_{sb}}{\text{PBSWG}}\right)^{-MJSWG}$$

If the area and perimeter are not specified, they default to 0 in ACM = 0 or 10, grossly underestimating the parasitic delay of the gate. HSPICE also supports ACM = 1, 2, 3, and 12 that provide nonzero default values when the area and perimeter are not specified. Check your models and read the HSPICE documentation carefully. The diffusion area and perimeter are also used to compute the junction leakage current. However, this current is generally negligible compared to sub thres hold leakage in modern devices.

Design Corners

Engineers often simulate circuits in multiple design corners to verify operation across variations in device characteristics and environment. HSPICE includes the .lib statement that makes changing libraries easy. The deck first sets SUP to the nominal supply voltage of 1.0 V. It then invokes .lib to read in the library specifying the TT conditions. In the stimulus, the. alter statement is used to repeat the simulation with changes. In this case, the design corner is changed. Altogether, three simulations are performed and three sets of waveforms are generated for the three design corners.

www.binils.com

Download Binils Android App in Playstore Download Photoplex App

Scaling:

- As the transistors become smaller, they switch faster, dissipate less power and are cheaper to manufacture. Despite the ever increase is challenges process advances have actually accelerated in the past decade.
- Such scaling is unprecedented in the history of technology. However scaling also excessive noise and reliability issues and introduces new problems.
- Designers need to be able to predict the effect of this feature size scaling on chip performance to plan future products, ensures existing products will scale gracefully to future processes for cost reduction and anticipate looming design challenges.

Transistor Scaling:

The characteristics of an MOS device can be maintained and the basic operational characters. Can be preserved if the critical parameters of a device are scaled by a dimensionless factor. These parameters include.

º All dimensions (x, y, z directions) º Device voltages

º Doping concentration densities.

Another approach is **lateral Scaling,** in which only the gate length is scaled. This is commonly called a gate shrink because it can done easily to an existing mask database for a design.

For **constant field scaling**, all devices dimension including channel length L, width W and oxide thickness tor are reduced by a factor of 1/s. The supply voltage VDD and the threshold voltages are also reduced by1/s.

- The substrate doping WA is increased by.
- Because both distance and voltage are scaled equally, the electric field remain constant.
- A gate shrink scales only the channel leng leaving other dimensions, voltages and doping levels unchanged.
- This offers a quadratic improvement in gate delay according to the first order

www.binils.com

Model.

- The gate delay improvement is closer to linear because velocity saturation keeps the current and effective resistance approximately constant.
- The constant voltage scaling increased the electric fields in the devices. By the 1 m generation velocity saturation was server enough that decreasing feature size no longer improved device current.

Inter connect Scaling:

- Two common approaches to interconnect scaling are to either scale all dimensions or keep the wire height constant.
- Wire length decreases for some types of wires, but may increase for others? Local are scaled wires are those that decrease in length during scaling.
- Example: A wire across 64 bits ALU is local because it becomes shorter as the ALU is migrated to finer process. A wire across a particular micro processer is scaled because when the microprocessor is shrunk to the new process the wire will also shrink.
- UN repeated interconnect delay is remaining about constant for local interconnect and increasing for global interconnect. This presents a problem because transistor are getting faster, so the ratio if interconnect to gate delay interconnect with scaling.
- In modern process with aspect ratios 1-5-22 fringing capacitance accounts for the majority of the total capacitance.
- Scaling spacing but not height interconnect the fringing capacitance enough that the extra thickness scarcely improves delay.
- Observe that when wire thickness is called the capacitance per unit length remains constant. Hence, a reasonable initial estimate of the capacitance of a

www.binils.com

Minimum-pitch were is about 0.2fF/ m, independent of the process.

- Wire capacitance is roughly 1/10-1/6 of gate capacitance per unit length.

Impacts on Design:

- One of the limitations of first order scaling is that it gives the wrong impression of being able to scale proportionally to zero dimensions and zero voltage.

Improved performance and cost:

- The most positive impact of scaling is that performance and cost are steadily improving. System architects need to understand the scaling of CMOS technologies and predict the capabilities of the process several years into the future, when a chip will be completed.

Interconnect:

Scaling transistors are steadily improving in delay but scaled wires are holding constant or getting worse.

- The wire problem motivated a number of papers predicting the demise of conventional wires.
- The plot is misleading in two ways.
- First the gate delay is shown for a single unloaded transistor rather than a realistically loaded gate. Second, the wire delay shown for fixed but as technology scales, most local wires connecting gates within a unit also become shorter.

Power:

In classical constant field scaling, power density remains constant and overall chip power increases only slowly with die size.

- Power density has sky rocketed because clock frequencies have increased much faster to classical scaling would predict and VDD is somewhat higher than

www.binils.com

Constant field scaling would demand.

- Dynamic power consumption will not continue to increase at such rates because it will become uneconomical to cool the chips.
- The static power consumption caused by sub threshold leakage was historically negligible but becomes important for threshold voltage below about 0.3 to 04v.

www.binils.com