DIGITAL ELECTRONICS

Digital Electronics CPE

Dr.Mohammad Abdel-Majeed

Assistant Professor

University of Jordan

Outline

- **Course Introduction**
  - **Course Information**
  - **Textbook**
  - **Grading**
  - **Course Outline**

Course Information

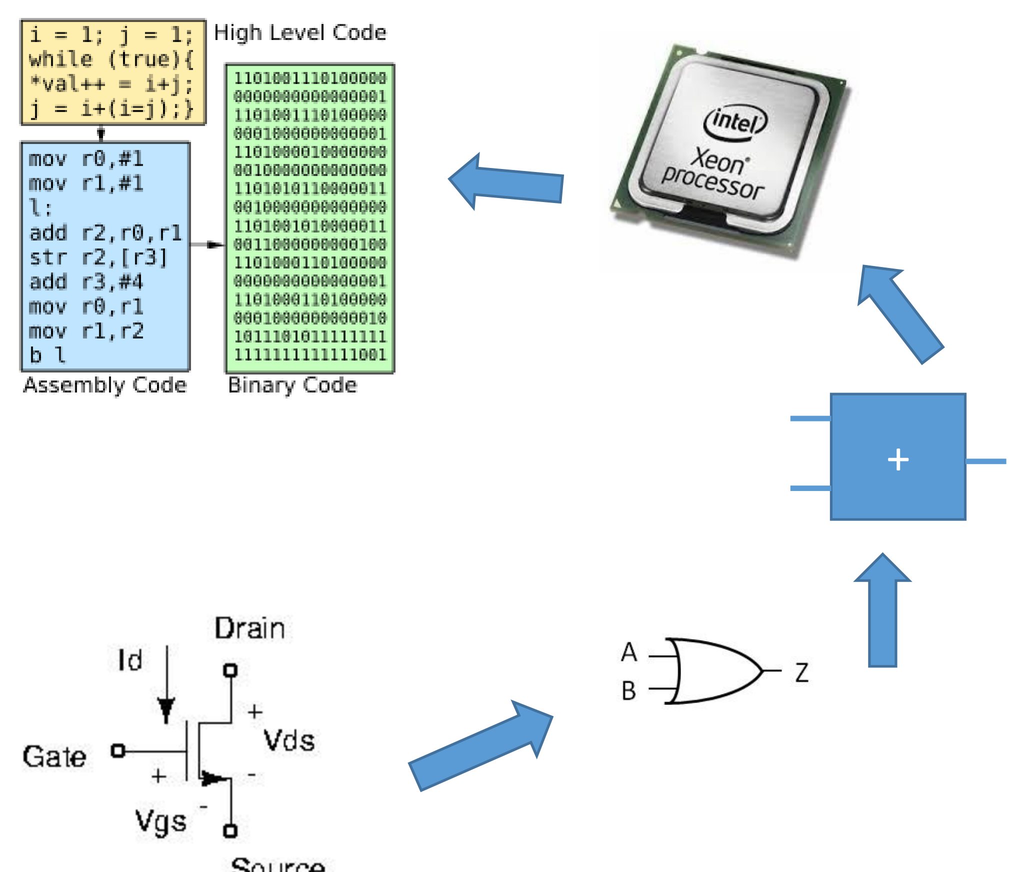
- Instructor: **Mohammad Abdel-Majeed**
- Email: m.abdel-majeed@ju.edu.jo
- Homepage: **www.piazza.com**
- Office: **CPE 421**
- Office Hours:
- Prerequisites:

Textbook

- *CMOS Digital Integrated Circuits Analysis and Design.(Kang and Leblebici)*
- *CMOS VLSI Design, A Circuits And Systems Perspective.(Weste and Harris)*

Grading Information

- **Grading**
  - Midterm Exam 30%
  - Assignments and Quizzes 20%
  - Final Exam 50%
- **Policies**
  - Attendance is required
  - Cheating will not be tolerated
  - No Makeup Exams



C/ C++	Apps
Assembly/ Machine code	Processor/Mem/ I/O
Processor/Mem/ I/O	
Functional units	
Logic Gates	
Transistors	
Voltage/Current	

Objective and Organization

- Provide in-depth understanding of the intricate issues in digital circuit design like:
  - **Performance**
  - **Interconnect**
  - Reliability
  - Low-power design
- Sequence of topics
  - Design Flow
  - MOS transistor (Overview)
  - MOS Inverter
  - MOS Inverter switching characteristics
  - Combinational MOS circuits
  - Sequential MOS Circuits
  - MOS Fabrication, schematic and Layout.(LTSpice and Electric Tools)

Objective and Organization

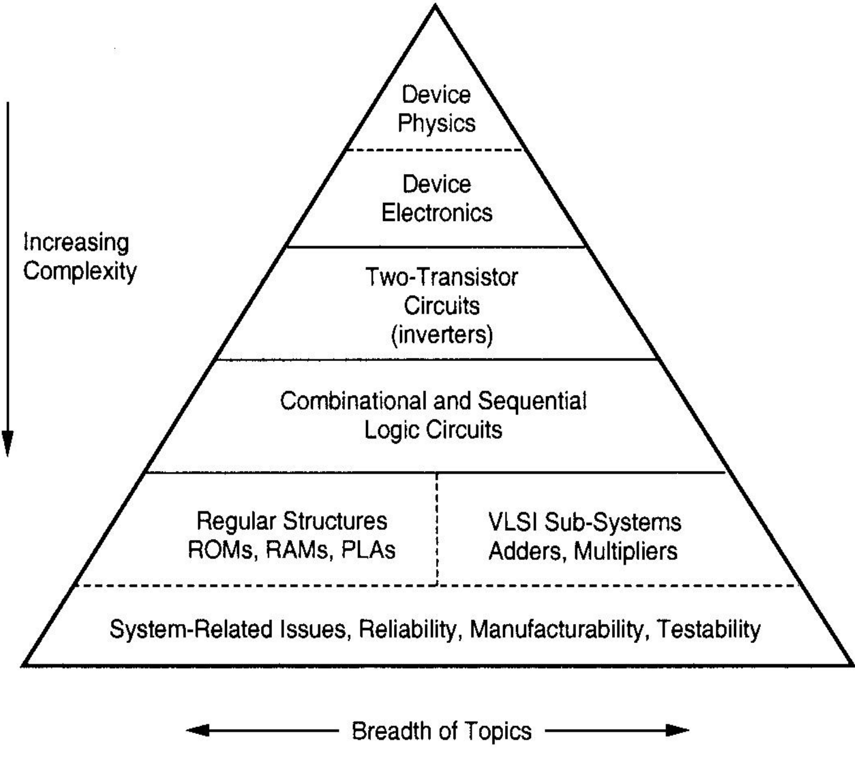


Figure 1.4 The ordering of topics covered in a typical digital integrated circuits course.

Lecture1
Design Flow

Dr.Mohammad Abdel-Majeed

Assistant Professor

University of Jordan

Coping with Complexity

- How to design System-on-Chip?
  - Many millions (even billions!) of transistors
  - Tens to hundreds of engineers
- Structured Design
- Design Partitioning

Structured Design

- _________: Divide and Conquer
  - Recursively system into modules
- _________:
  - Reuse modules wherever possible
  - Ex: Standard cell library
- _________: well-formed interfaces
  - Allows modules to be treated as black boxes
- _________:
  - Physical and temporal

Structural Decomposition of 4-b Adder

Structural Decomposition of 4-b Adder

- Easier to handle

add4

add add add add

carry sum carry sum carry sum carry sum

and or and or and or and or

ci
b[3:0]
a[3:0]
s[3:0]
c4

c
b
a
s
co

a
b
c
sum
s

a
b
c
carry
co

Structural Hierarchy of 16-b Adder

Structural Hierarchy of 16-b Adder

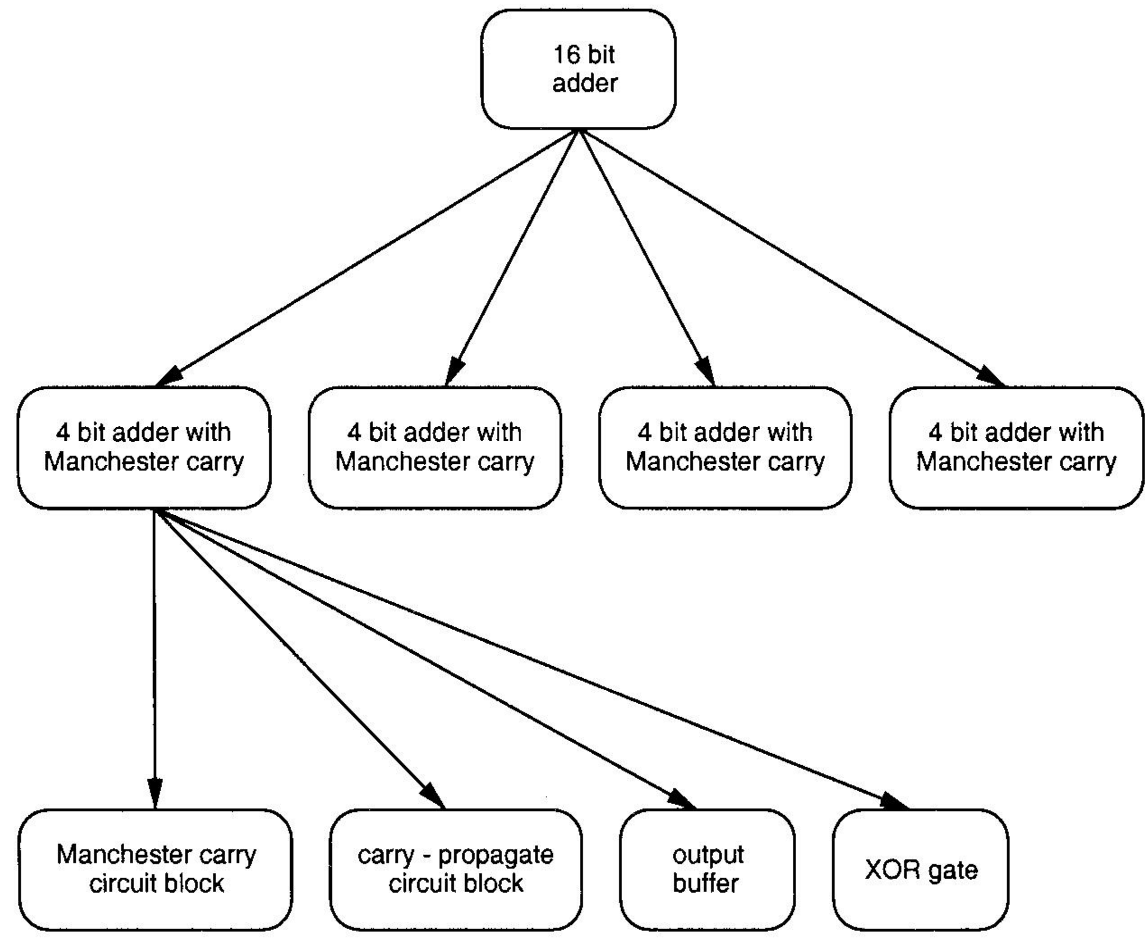
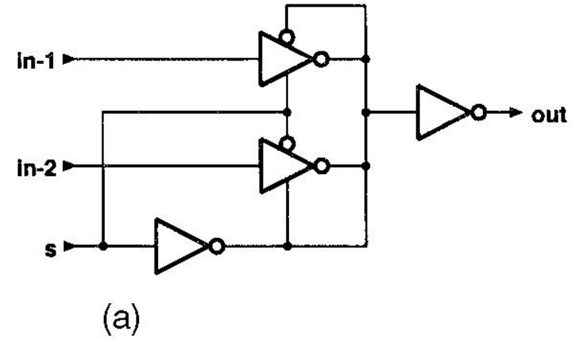


Figure 1.25 Structural hierarchy of the 16-bit adder circuit.

Concepts of Regularity

- Regularity
  - decomposition into similar blocks
  - Example: parallel multiplication array



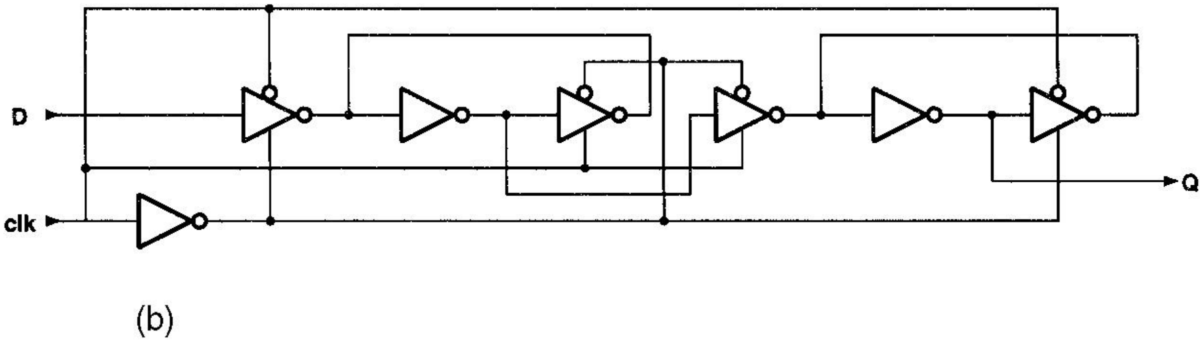


Figure 1.26 Regular design of (a) 2-1 MUX and (b) DFF, using inverters and tri-state buffers as basic building blocks.

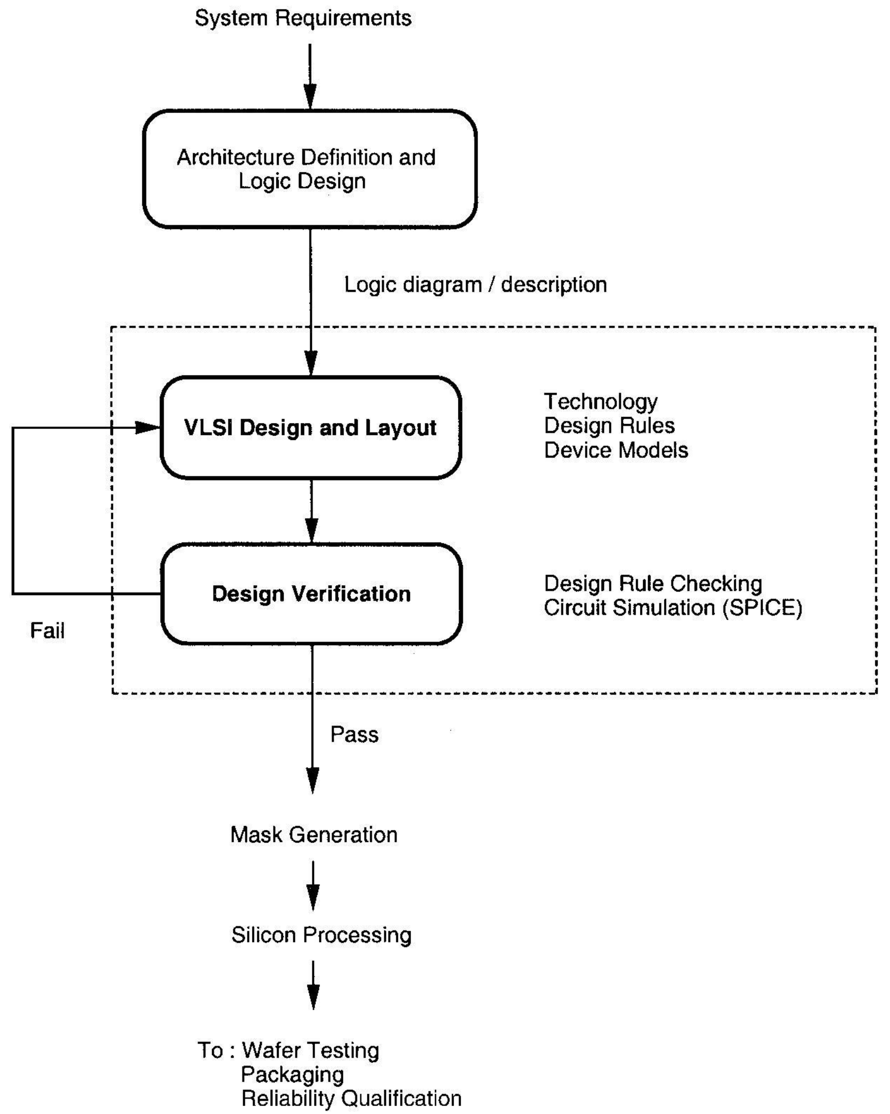
Concepts of Modularity and Locality

- Modularity
  - Functional blocks have well-defined functions and interfaces
  - Each block can be designed independently and combined easily
  - Design process parallelized
- Locality
  - Ensures connections are mostly between neighboring modules
  - Delay minimized by avoiding long interconnect

Design Partitioning

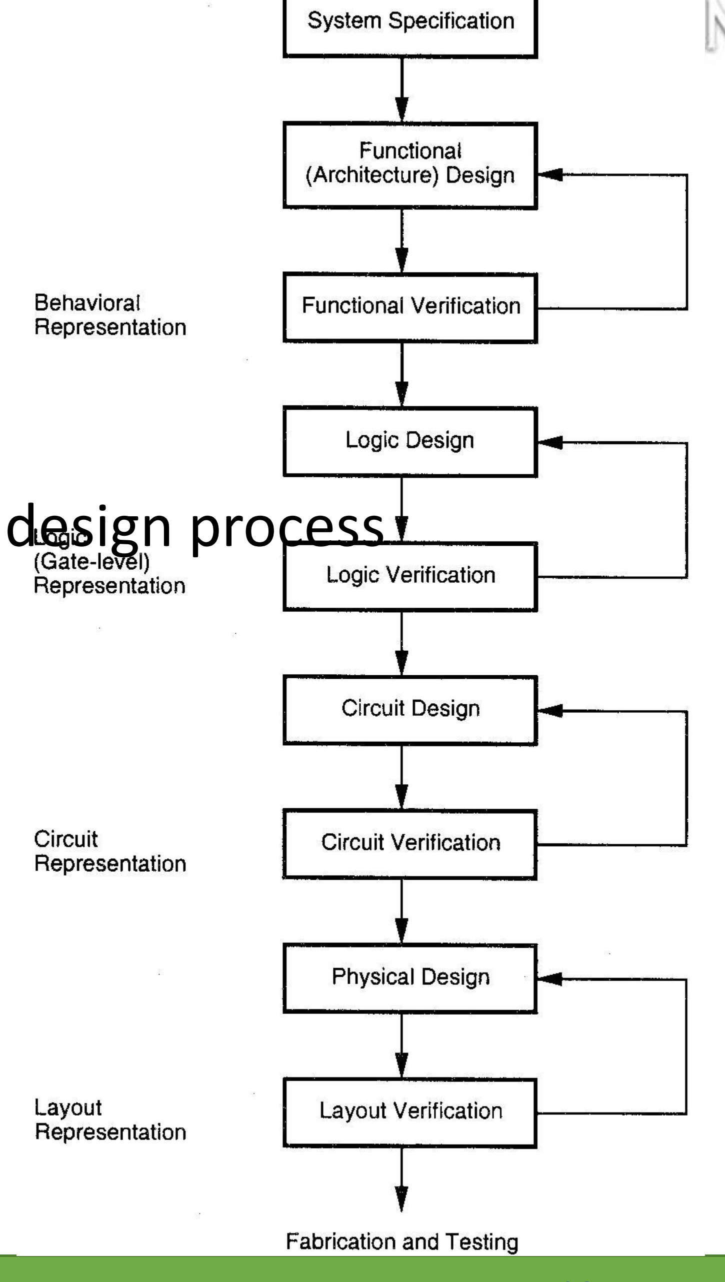
- **Architecture**: User’s perspective, what does it do?
  - Instruction set, registers
  - MIPS, x86, Alpha, PIC, ARM, …
- **Microarchitecture**
  - Single cycle, multcycle, pipelined, superscalar?
- **Logic**: how are functional blocks constructed
  - Ripple carry, carry lookahead, carry select adders
- **Circuit**: how are transistors used
  - Complementary CMOS, pass transistors, domino
- **Physical**: chip layout
  - Datapaths, memories, random logic

Flow of Circuit



More Simplified VLSI Design Flow

- Simplified design flow
  - Verification plays an important role in every step
  - Top-down and bottom-up approaches combined in the design process



Example 1.1 (1)

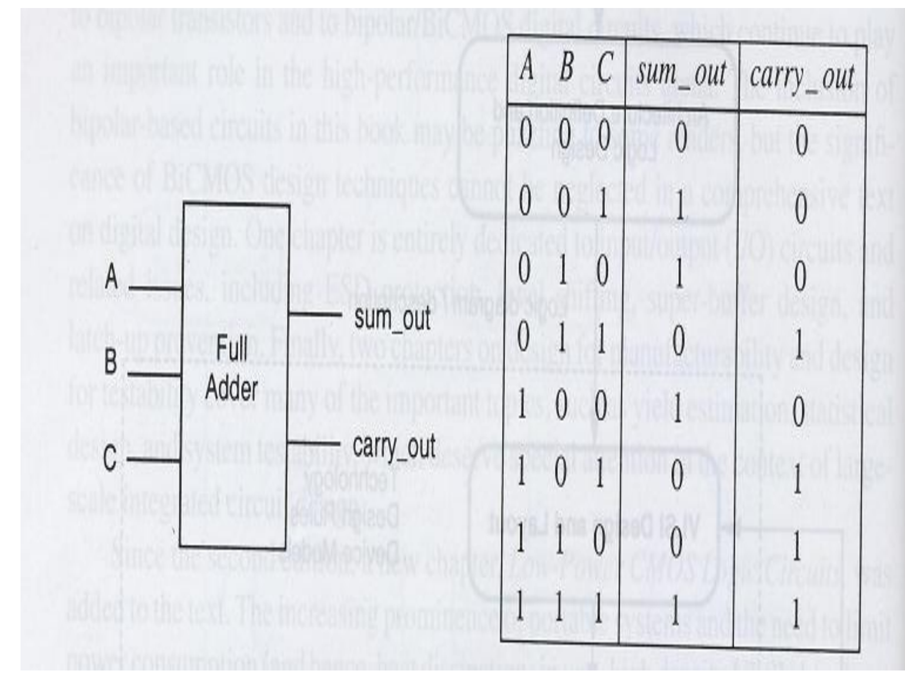
- Problem: Design of 1-bit full-adder circuit using __ nm, twin-well CMOS technology
  - Specifications:
    - Propagation delay of sum and carry_out <_____ __
    - Transition delay of sum and carry_out < ____ __
    - Circuit area < 10 μm^2
    - Dynamic power dissipation (@V_{DD} = 1.1V and f_{max} = 500MHz) < 20 μW

Example 1.1 (2)

- Boolean Description
  - Boolean Functions:
    - A , B = Two inputs
    - C = Carry in
    - sum_out =ABC +AB’C’ + A’B’C + A’C’B
    - carry_out = AB +AC+BC
    - Alternatively, sum_out = ABC+(A+B+C)(carry_out)’

Example 1.1 (2)

- Boolean Description
  - Boolean Functions:
    - A , B = Two inputs
    - C = Carry in
    - sum_out =ABC +AB'C' + A'B'C + A'C'B
    - carry_out = AB +AC+BC
    - Alternatively, sum_out = ABC+(A+B+C)(carry_out)'



A	B	C	sum_out	carry_out
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Example 1.1 (3)

- Logi

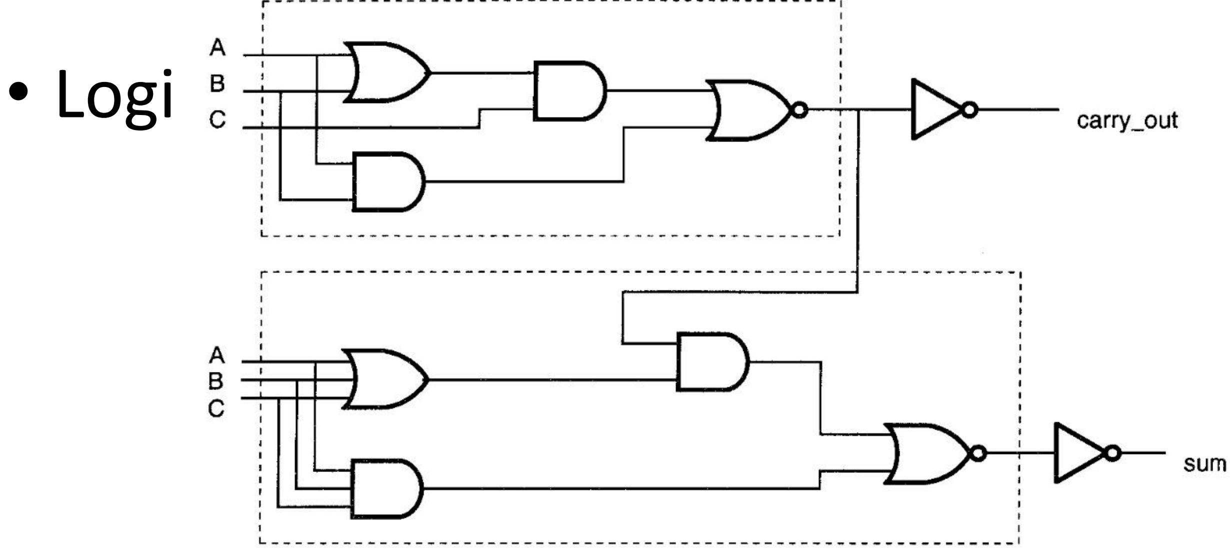


Figure 1.7 Gate-level schematic of the one-bit full-adder circuit.

Example 1.1 (4)

- Transistor-level circuit
  - AND = Series-connected nMOS
  - OR = Parallel- connected nMOS
  - pMOS network = dual of nMOS network

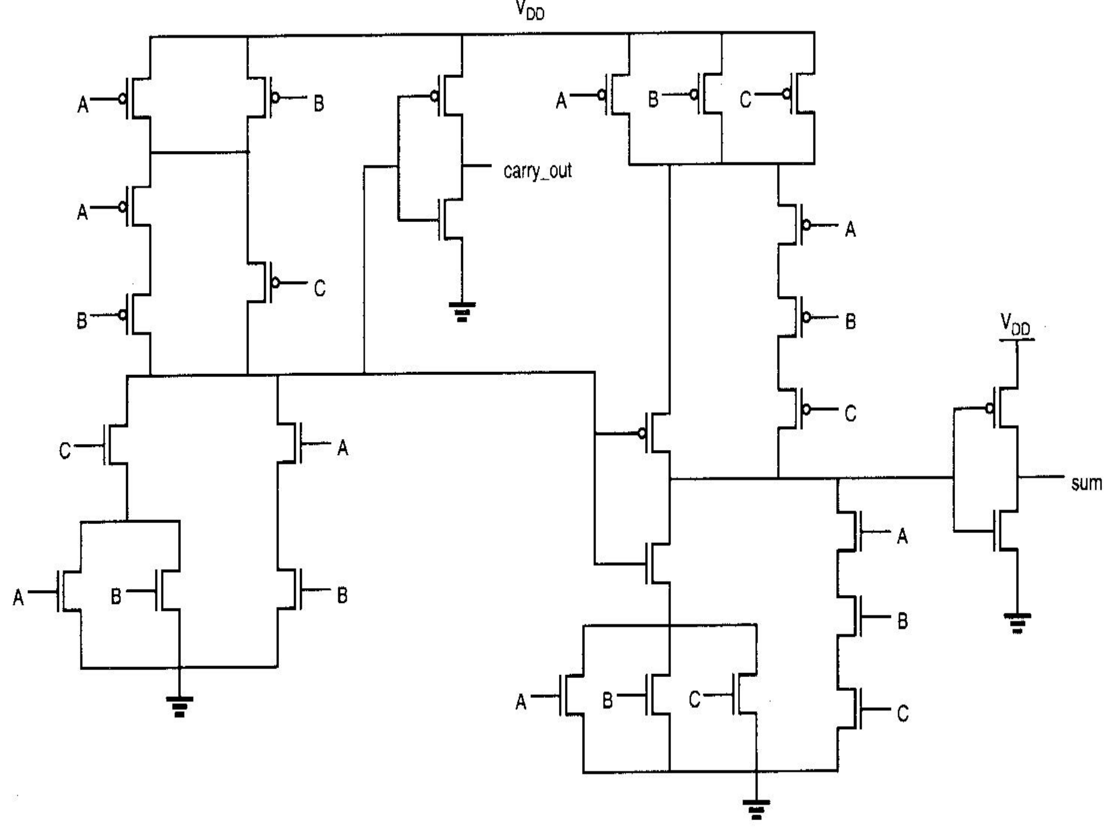


Figure 1.8 Transistor-level schematic of the one-bit full-adder circuit.

Example 1.1 (6)

- Initial sizes
  - nMOS, (*W/L*) = 90nm/50nm
  - pMOS, *(W/L)* = 90nm/50nm
  - May need to be changed depending on performance

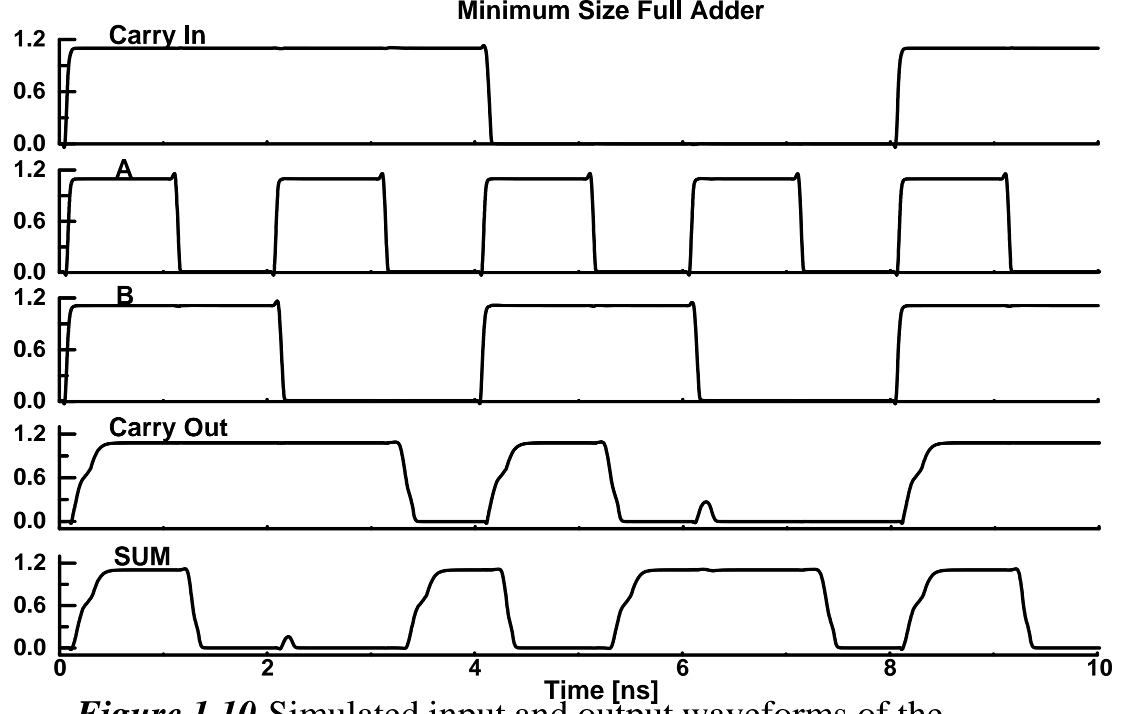


Figure 1.10. Simulated input and output waveforms of the full-adder circuit.

Example 1.1 (7)

- Timing constraint violation
    - sum_out and carry_out violate timing constraints
    - Worst-case delay 250 ps (> 220 ps)
    - Modification necessary

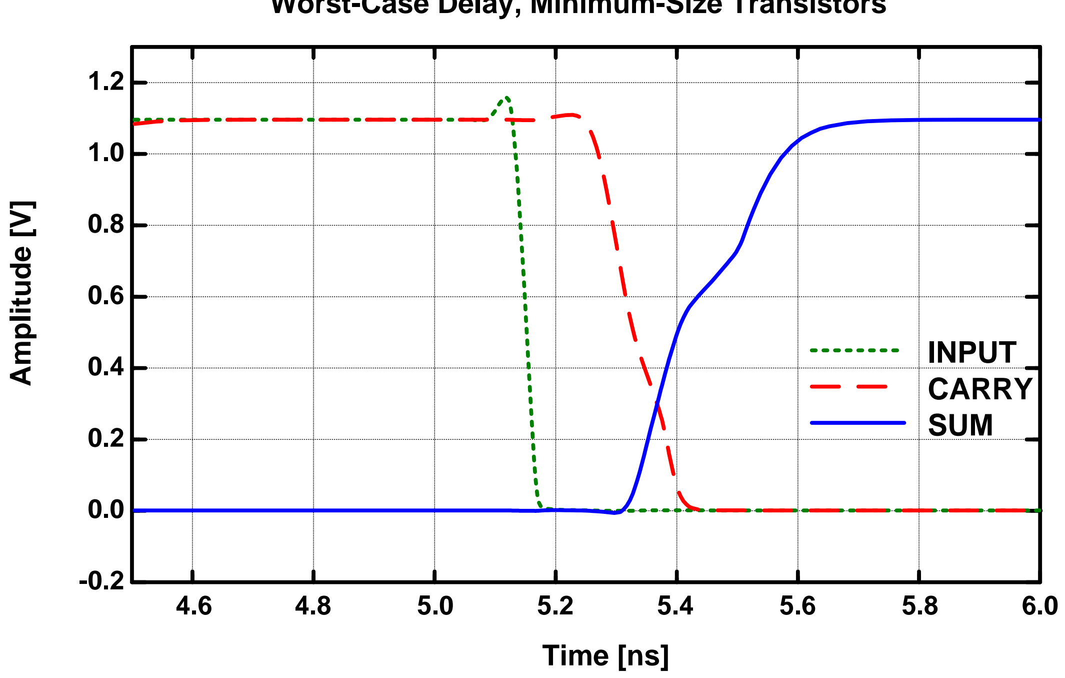


Figure 1.11. Simulated output waveforms of the full adder circuit with minimum transistor dimensions, showing the signal propagation delay during one of the worst-case transitions.

Example 1.1 (8)

- Resizing transistors to improve design
  - is an iterative process
  - To meet timing specifications (W/L) of (n/p)MOS is increased

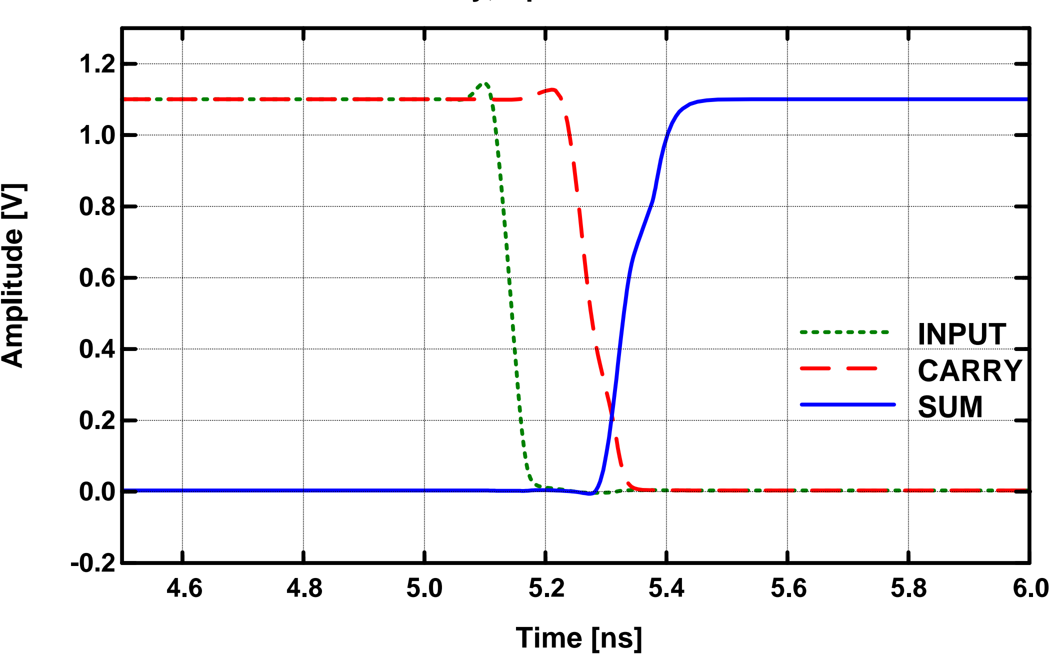


Figure 1.13. Simulated output waveforms of the full-adder circuit with optimized transistor dimensions, showing the signal propagation delay during the same worst-case transition.

Example 1.1 (8)

- Layout Design
  - Design rule checker (DRC) tool used to check violation of design rules
  - Parasitic capacitances and resistances extracted
- Design Verification
  - Extracted parasitics used to create SPICE input file
  - Simulation is run
- Simulation Results
  - Not all specifications met

Example 1.1 (9)

- New and compact layout for 1-bit full adder (optimized)
- Now, all the design specifications are satisfied
  - Propagation and transition (rise/fall) delay within 220 ps
  - Dynamic power dissipation = 4.9 µW (<20µW)
  - Area = (2.04 µm x 3.01 µm) = 6.14 μm^2 (<10 μm^2)

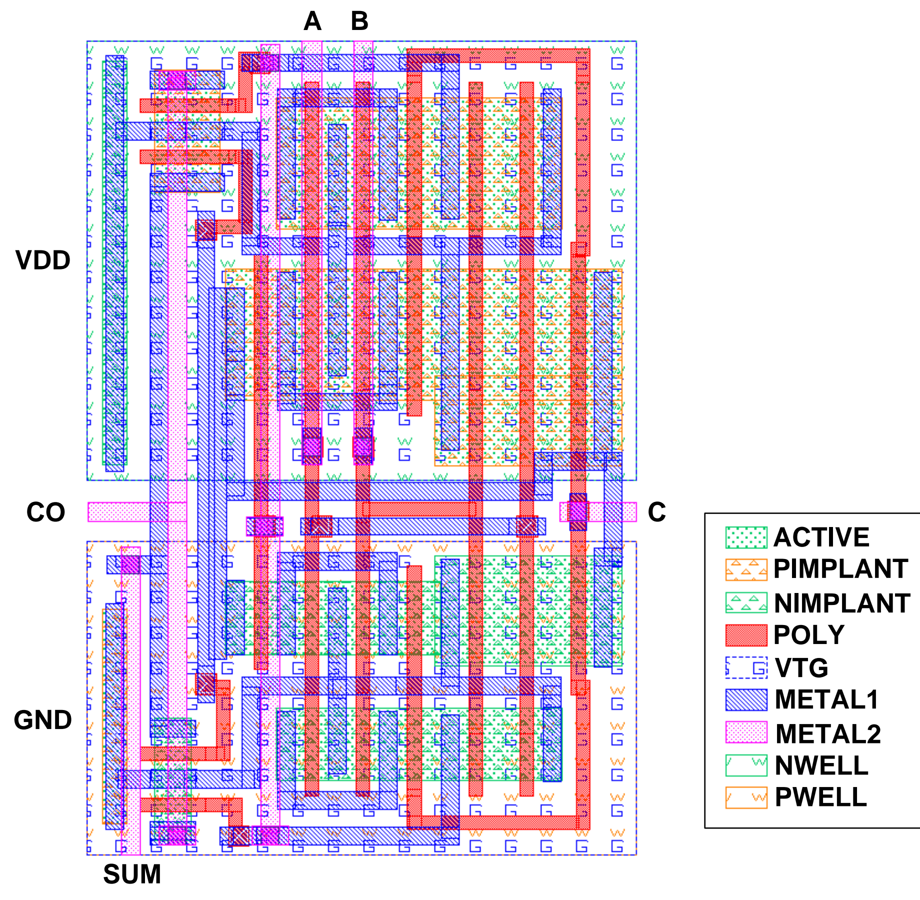


Figure 1.12. Layout of the full-adder circuit, with optimized transistor dimensions.

8-bit Binary Adder (1)

- Obtained by cascading 8 full adders – called "carry ripple adder"
  - Speed limited by the delay of carry bits

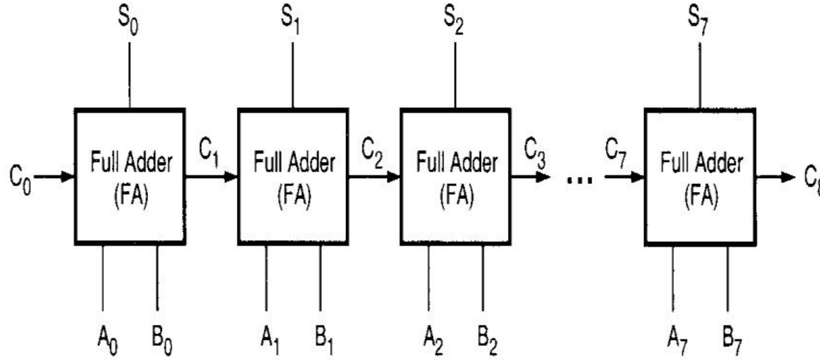
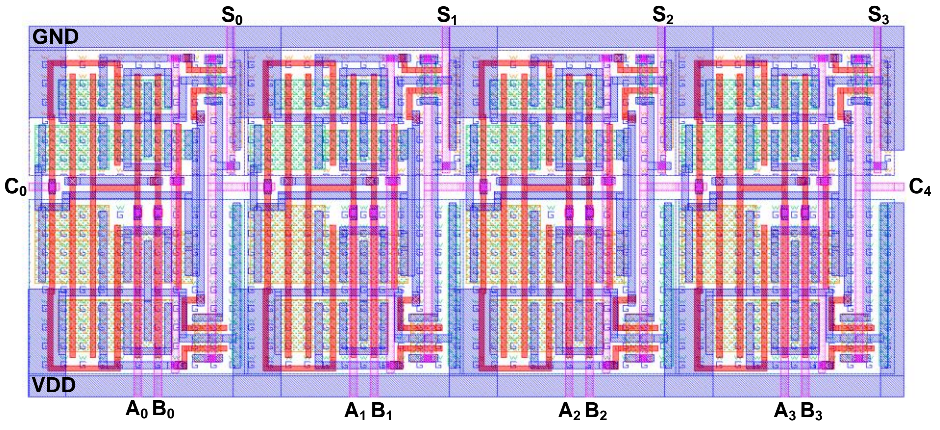
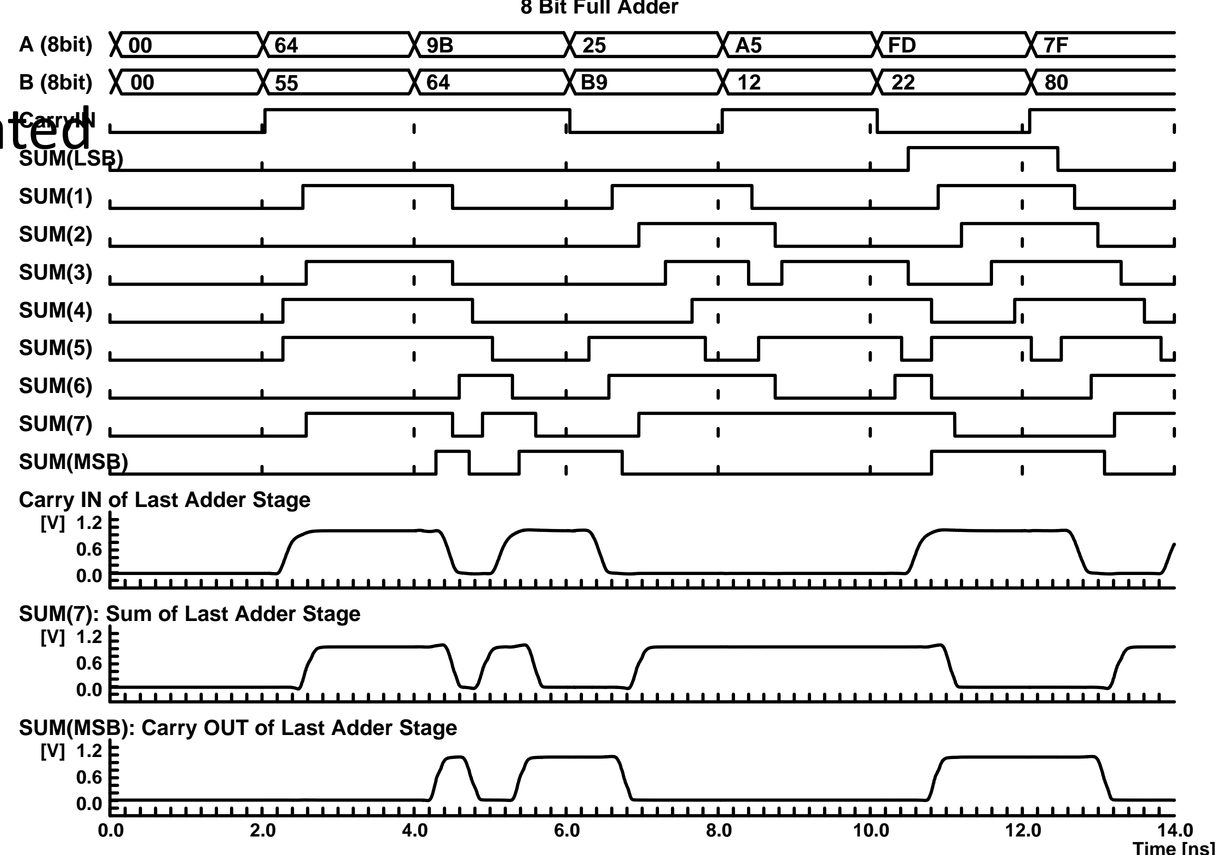


Figure 1.15 Block diagram of a carry ripple adder chain consisting of full adders.

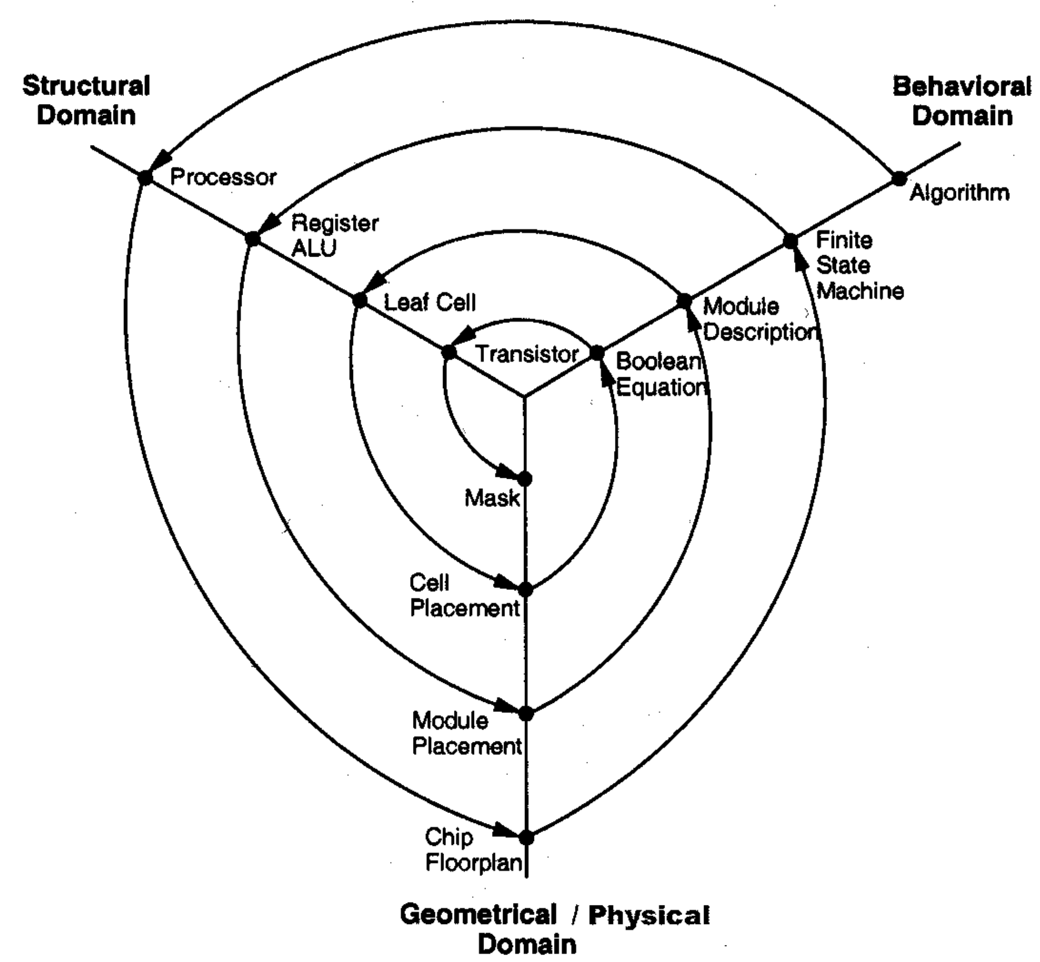


8-bit Binary Adder (2)

- Simulation results
  - Sum bit of last adder stage is generated last
  - Overall delay as long as 0.7 ns



Y-Chart



VLSI Design Styles

- Field Programmable Gate Array (FPGA)
  - Consists of
    - I/O buffers
    - Array of configurable logic blocks (CLBs)
    - Programmable interconnect structure
  - Contains thousands of logic gates
  - Routing between CLBs and I/O blocks done by setting the configurable switch matrices
- Proper choice of design style is essential to delivering the product in time with low cost
  - Full-custom
  - Semi-custom

Field Programmable Gate Array (1)

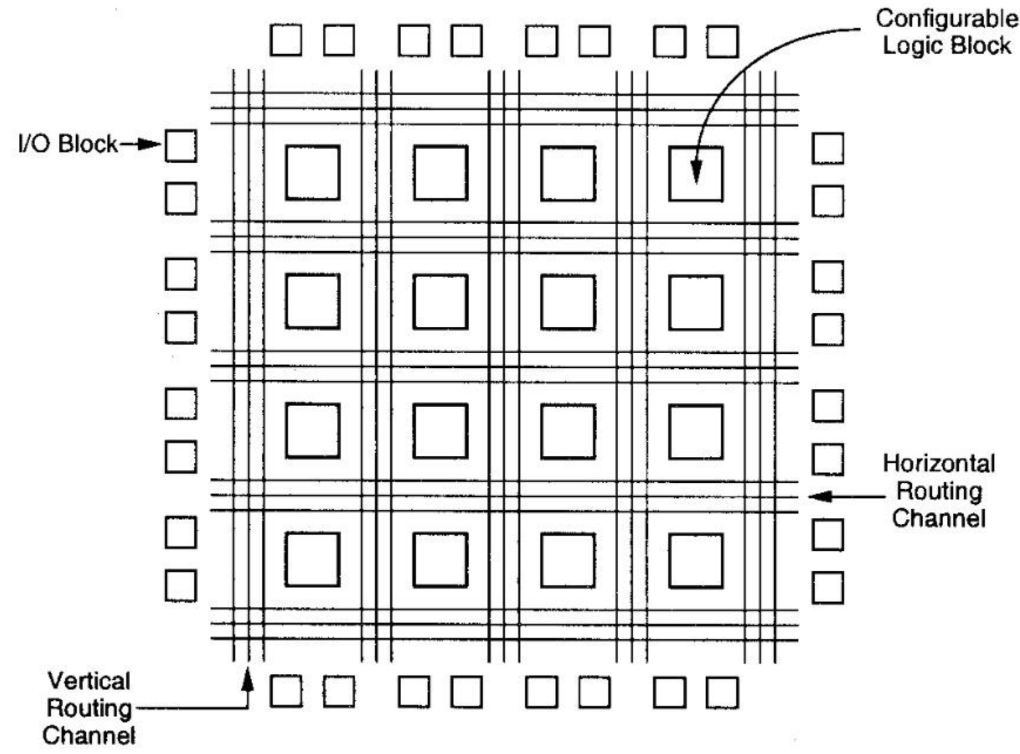


Figure 1.27 General architecture of Xilinx FPGAs.

https://www.youtube.com/watch?v=gUsHwi4M4xE

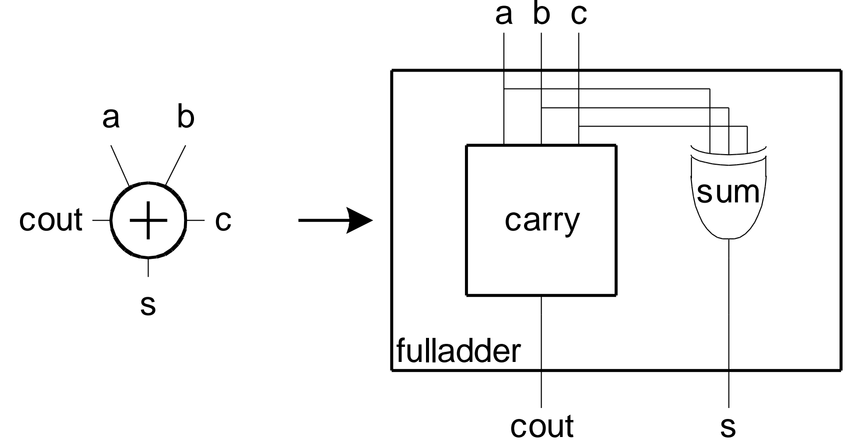
Verilog Example

```
module fulladder(input  a, b, c,
                 output s, cout);

  sum          s1(a, b, c, s);
  carry        c1(a, b, c, cout);
endmodule

module carry(input  a, b, c,
             output cout)

  assign cout = (a&b) | (a&c) | (b&c);
endmodule
```



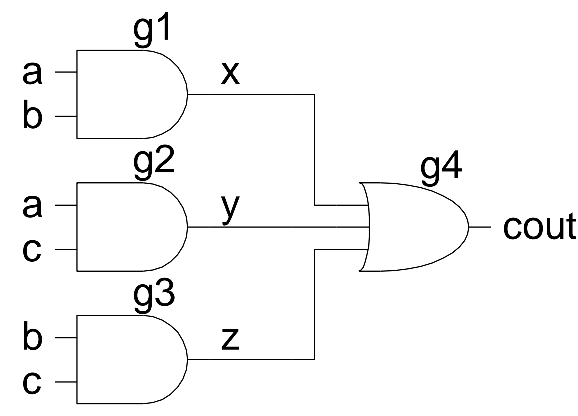
Gate-level Netlist(Synthesis)

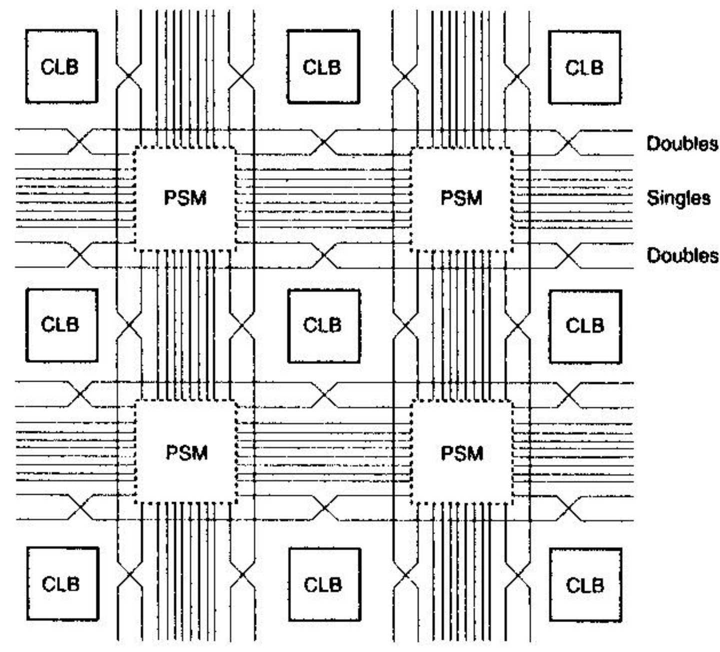
```
module carry(input  a, b, c,
                    output cout)

    wire   x, y, z;

    and g1(x, a, b);
    and g2(y, a, c);
    and g3(z, b, c);
    or  g4(cout, x, y, z);
endmodule
```

Place and route





Circuit Design

- How should logic be implemented?
  - NANDs and NORs vs. ANDs and ORs?
  - Fan-in and fan-out?
  - How wide should transistors be?
- These choices affect speed, area, power
- Logic synthesis makes these choices for you
  - Good enough for many applications
  - Hand-crafted circuits are still better

Standard-Cell Based Design (1)

- One of the most prevalent full custom design styles
  - Commonly used logic cells are optimized and developed
  - Several versions are stored in a standard library cell

- Each cell is characterized by
  - Delay time vs. load capacitance
  - Circuit simulation model
  - Timing simulation model
  - Fault simulation model
  - Cell data for place-and-route
  - Mask data

Standard-Cell Based Design (2

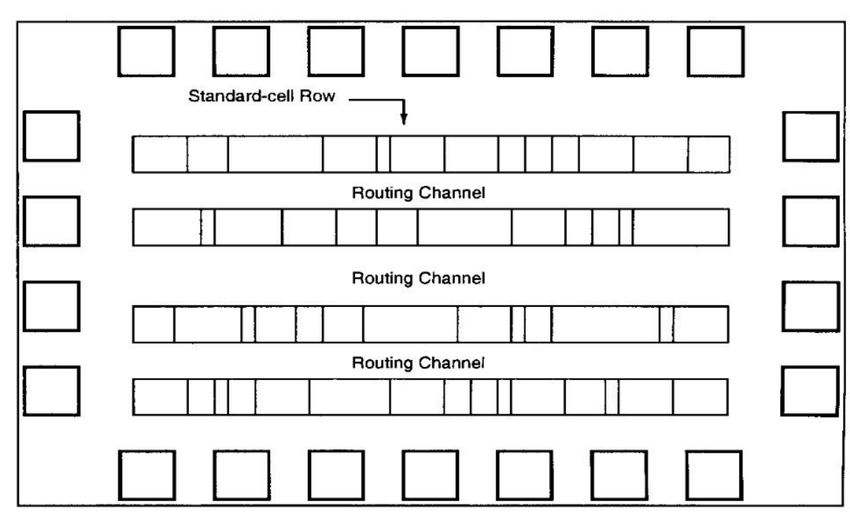


Figure 1.37 A simplified floorplan of standard-cells based design.

- Each cell layout is designed with fixed height
  - Cells can be placed side-by-side
  - Routing of intercell connection is easy

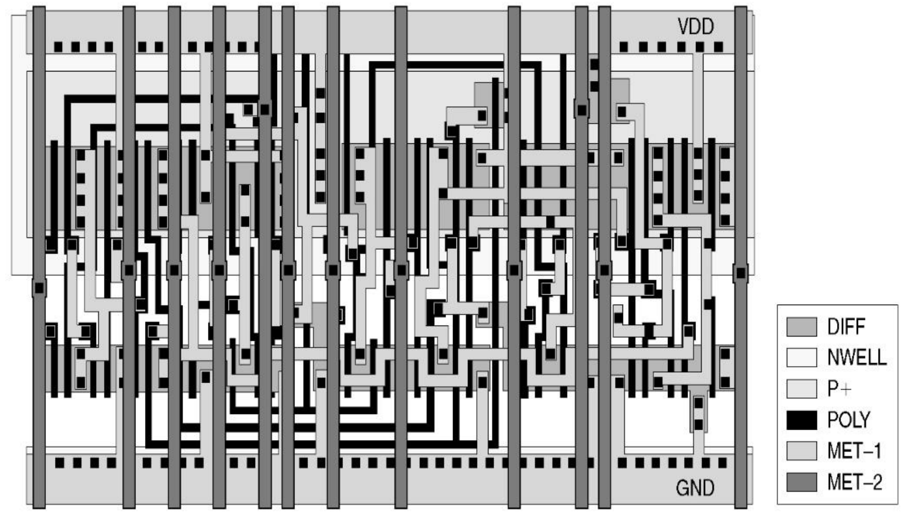


Figure 1.36 A standard cell layout example.

- Floorplan for a standard-cell based design contains
  - I/O frame, cell rows
  - Channels between rows
    - channels may be reduced or removed if over-the-cell routing is done

Standard-Cell Based Design (3)

- Common bus may be incorporated if cells must share same input and/or output signals

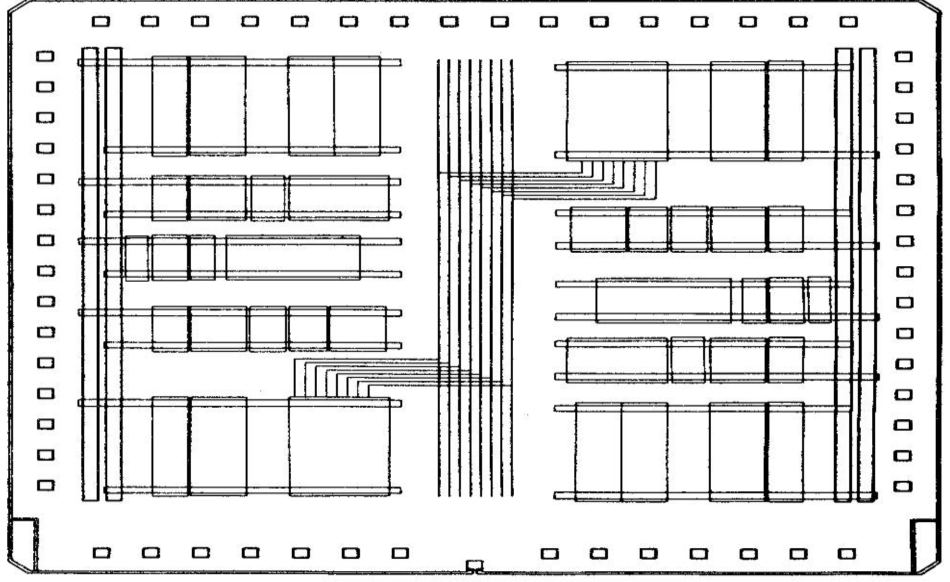


Figure 1.38 Simplified floorplan of a standard-cells based design, consisting of two separate blocks and a common signal bus.

Structured ASIC (______________________________)

FPGA	Vs.	Standard Cell ASIC
Easy to Design		Difficult to Design
Short Development Time		Long Development Time
Low NRE Costs		High NRE Costs
Design Size Limited		Support Large Designs
Design Complexity Limited		Support Complex Designs
Performance Limited		High Performance
High Power Consumption		Low Power Consumption
High Per-Unit Cost		Low Per-Unit Cost (at high volume)

Structured ASIC's Combine the Best of Both Worlds

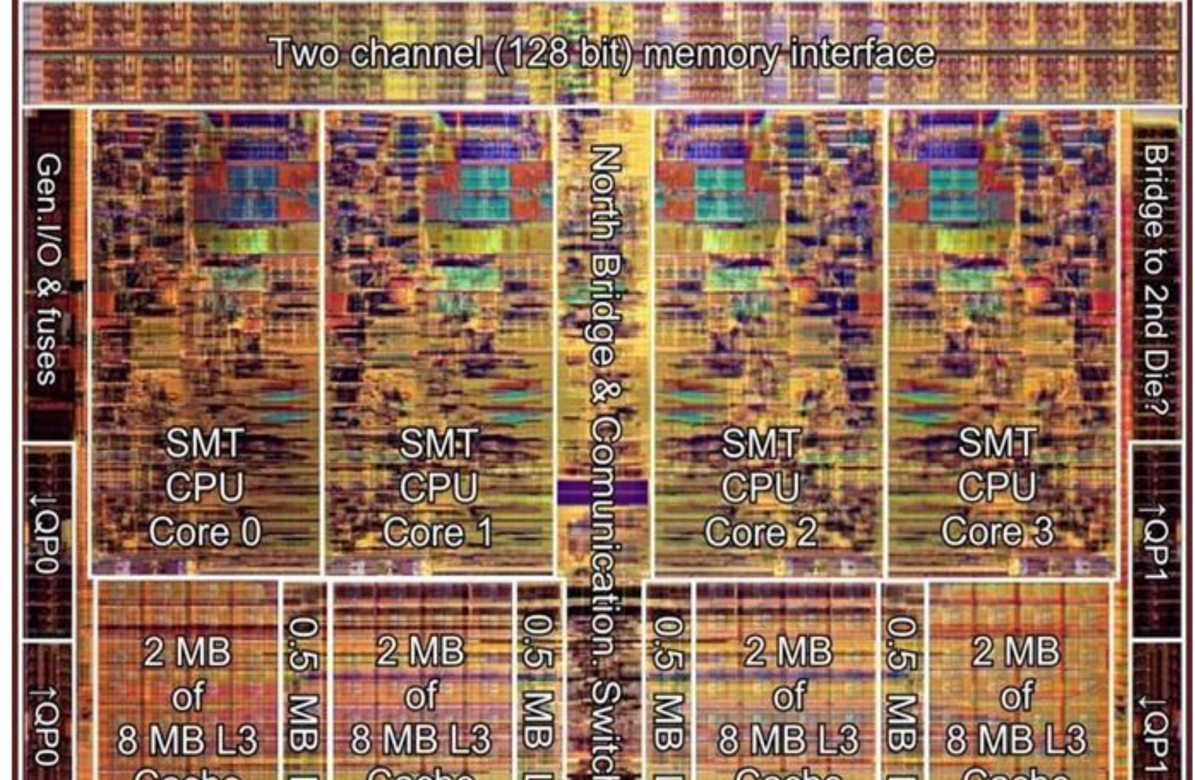
- Generally speaking
  - 100:33:1 ratio between the number of gates in a given area for _______, ________, _______
  - ___:___:___ ratio for performance (based on clock frequency)
  - ___:___:___ ratio for power

Full Custom Design (1)

- Design is done from scratch
  - Geometry, orientation and placement of every transistor done by designer
  - Development cost and time very high
    - "Design Reuse" becoming popular to reduce cost and time
  - Example of a true full custom design – design of memory cell (static or dynamic)

Full Custom Design (2)

- Full custom design rarely used due to high labor cost
  - Rather combination of different design styles are used to develop a chip



Design Quality

- Important metrics for measuring the quality of design
  - Testability
  - Yield and manufacturability
  - Reliability
  - Technology updateability

Testability

- Fabricated chips should be fully testable which requires
  - Generation of good test vectors
  - Availability of reliable test fixture at speed
  - Design of testable chip

Yield and Manufacturability

- Yield may be defined in two ways
  - (1) No. of good tested chips divided by the total no. of tested chips
  - (2) No. of good tested chips divided by the total no. of chip sites available at the start of wafer processing – strictest definition
- Chip yield can be further divided into
  - Functional yield – obtained by testing the functionality of the chip at a speed lower than required
    - Weeds out problems of short, open and leakage
    - Can detect logic and circuit design faults
  - Parametric yield – performed at the required speed on chips that passed functional test
    - Delay testing done in this phase

Reliability

- Reliability depends on design and process conditions
  - Major causes of chip reliability problem are
    - Electrostatic discharge (ESD) and electrical overstress (EOS) and electromigration
    - Latch-up in CMOS I/O internal circuits
    - Hot carrier induced aging
    - Oxide breakdown and single event upset
    - Power and ground bouncing
    - On-chip noise and crosstalk
- Measures taken to ensure reliability
  - Metal wire widened to avoid over-etching
  - Rise time of signals applied to nMOS gate reduced to avoid aging

Technology Updateability

- Process technology advancing at a high pace
  - Design styles should be chosen such that chips are technology updateable
    - “Silicon Compilation” – where physical layout is done automatically – is used

References

Backup

MIPS Architecture

- Example: subset of MIPS processor architecture
  - Drawn from Patterson & Hennessy
- MIPS is a 32-bit architecture with 32 registers
  - Consider 8-bit subset using 8-bit datapath
  - Only implement 8 registers ($0 - $7)
  - $0 hardwired to 00000000
  - 8-bit program counter
- You’ll build this processor in the labs
  - Illustrate the key concepts in VLSI design

Instruction Set

Table 1.7 MIPS instruction set (subset supported)

Instruction	Function		Encoding	op	funct
`add $1, $2, $3`	addition:	$1 → $2 + $3	R	000000	100000
`sub $1, $2, $3`	subtraction:	$1 → $2 – $3	R	000000	100010
`and $1, $2, $3`	bitwise and:	$1 → $2 and $3	R	000000	100100
`or $1, $2, $3`	bitwise or:	$1 → $2 or $3	R	000000	100101
`slt $1, $2, $3`	set less than:	$1 → 1 if $2 < $3 $1 → 0 otherwise	R	000000	101010
`addi $1, $2,`	add immediate:	$1→ $2 + imm	I	001000	n/a
`beq $1, $2, imm`	branch if equal:	PC → PC + imm[a]	I	000100	n/a
`j destination`	jump:	PC_destination[a]	J	000010	n/a
`lb $1, imm($2)`	load byte:	$1 → mem[$2 + imm]	I	100000	n/a
`sb $1, imm($2)`	store byte:	mem[$2 + imm] → $1	I	110000	n/a

Instruction Encoding

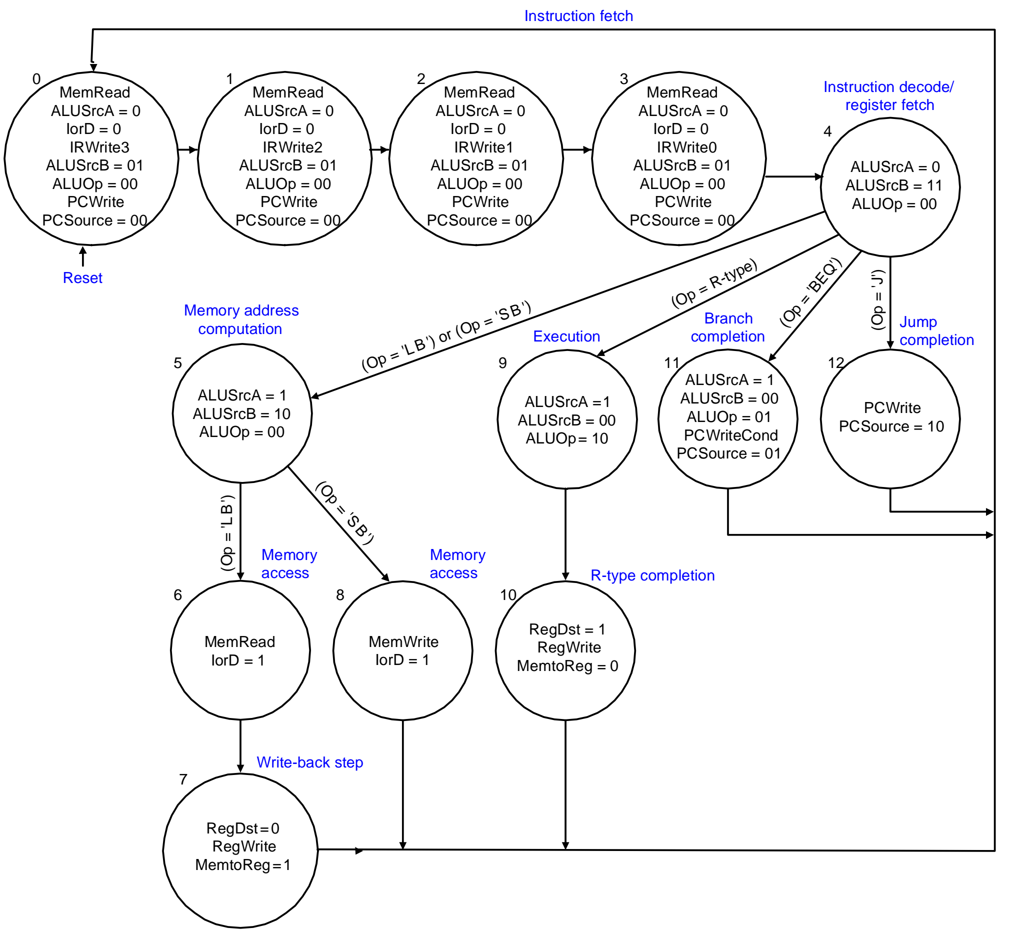
- 32-bit instruction encoding
  - Requires four cycles to fetch on 8-bit datapath

format	example	encoding					
		6	5	5	5	5	6
R	add $rd, $ra, $rb	0	ra	rb	rd	0	funct
		6	5	5	16		
I	beq $ra, $rb, imm	op	ra	rb	imm		
		6	26				
J	j dest	op	dest				

MIPS Microarchitecture

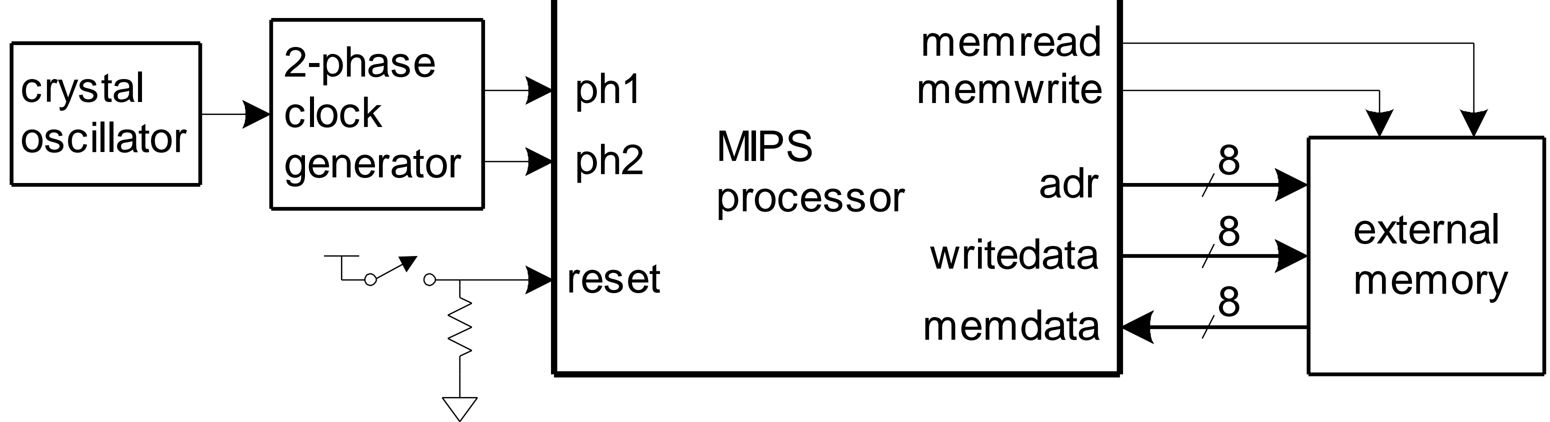
- Multicycle μarchitecture ([Paterson04], [Harris07])

Multicycle Controller

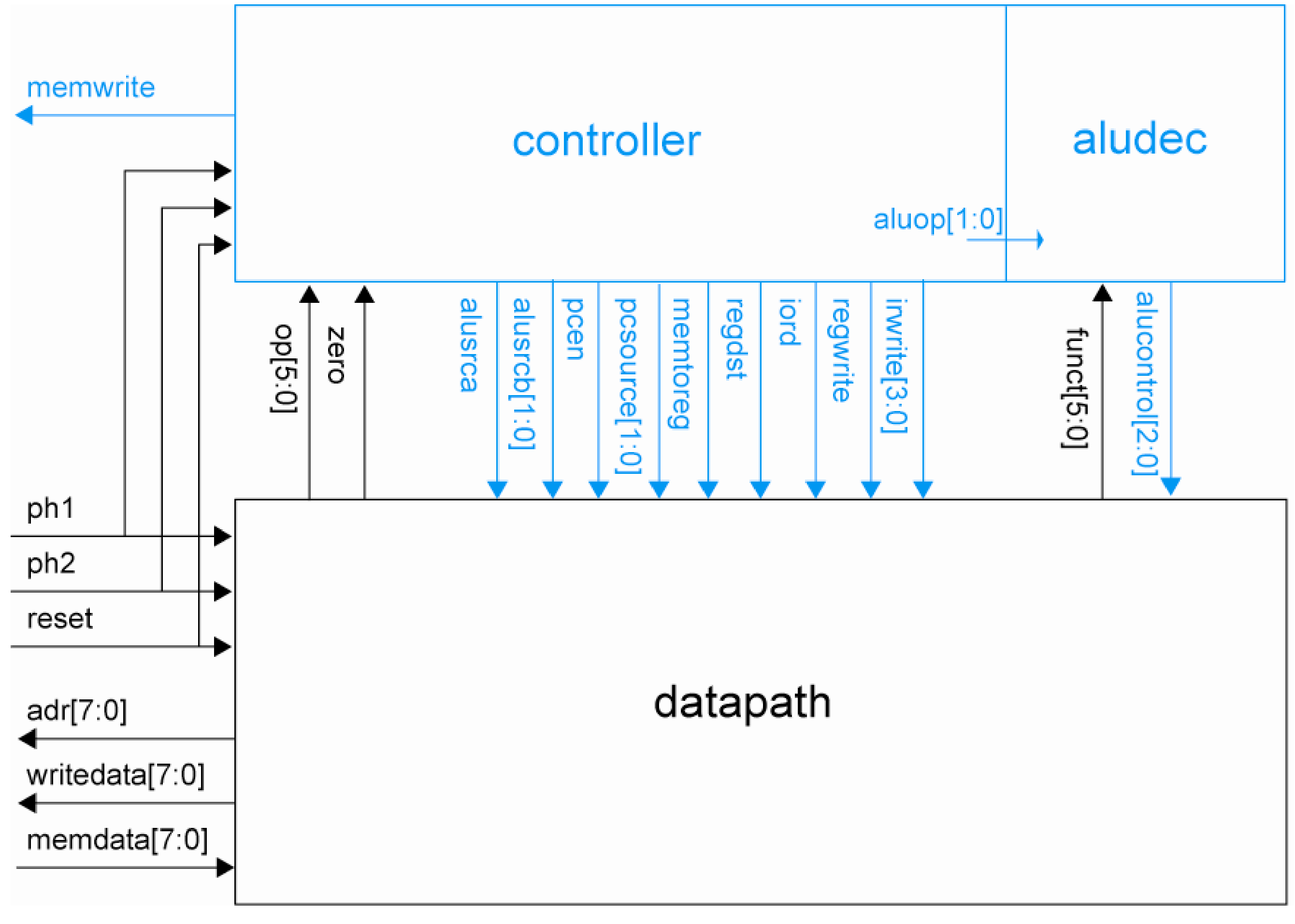


Logic Design

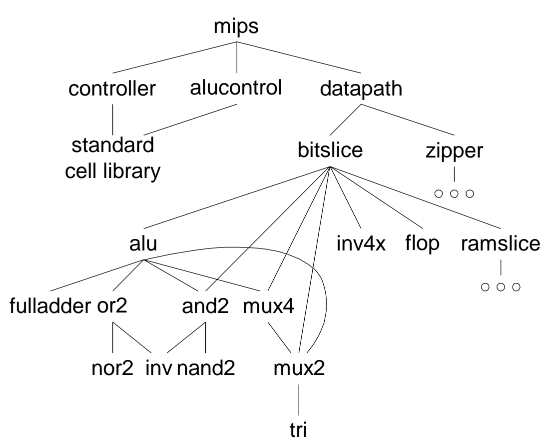
- Start at top level
  - Hierarchically decompose MIPS into units
- Top-level interface



Block Diagram



Hierarchical Design



CMOS Transistor

Dr.Mohammad Abdel-Majeed

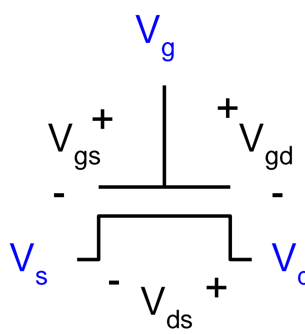
Assistant Professor

University of Jordan

How Does a Transistor Work?

- http://www.youtube.com/watch?v=IcrBqCFLHIY

Terminal Voltages



- Mode of operation depends on V_g, V_d, V_s
  - $V_{gs} = V_g – V_s$
  - $V_{gd} = V_g – V_d$
  - $V_{ds} = V_d – V_s = V_{gs} - V_{gd}$
- Source and drain are symmetric diffusion terminals
  - By convention, source is terminal at lower voltage
  - Hence $V_{ds} \geq 0$
- nMOS body is grounded. First assume source is 0 too.
- Three regions of operation
  - *Cutoff*
  - *Linear*
  - *Saturation*

NMOS cross section

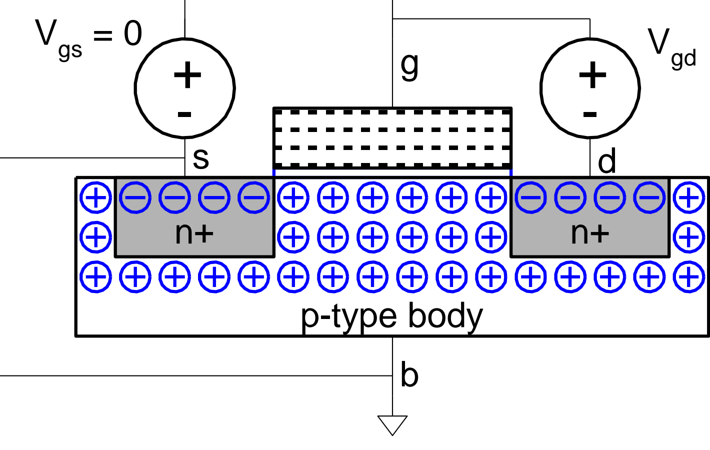
Cutoff

Linear

Saturation

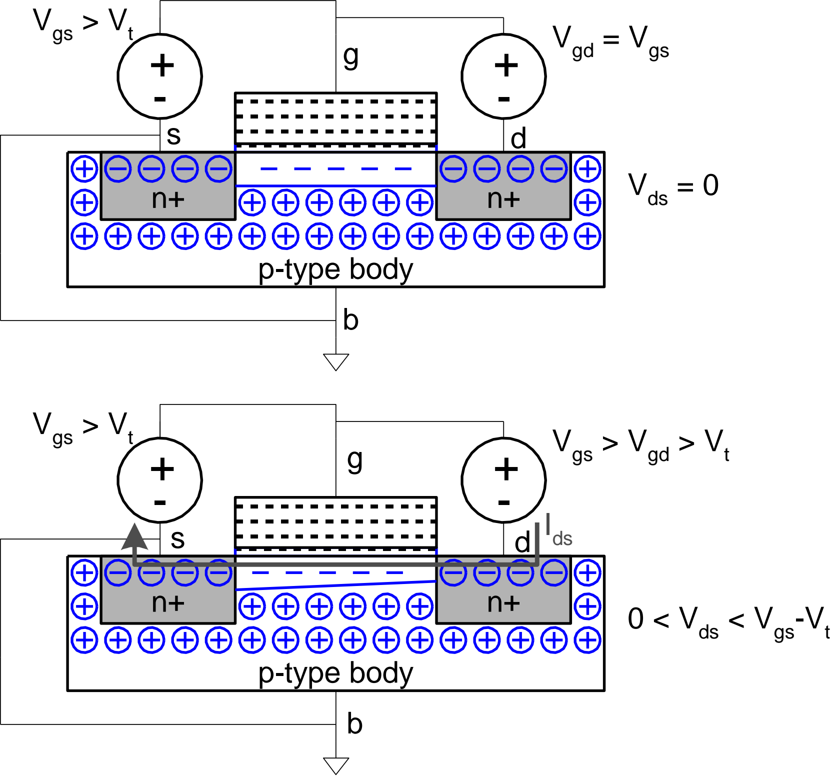
nMOS Cutoff

- No channel
- $I_{ds} \approx 0$



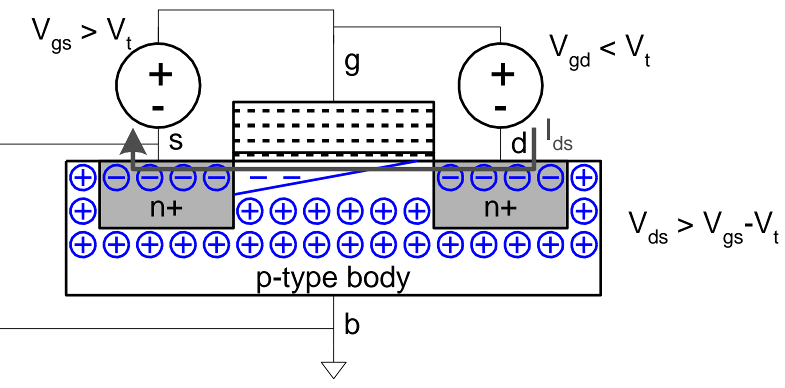
nMOS Linear

- Channel forms
- Current flows from d to s
  - e^- from s to d
- I_{ds} increases with V_{ds}
- Similar to linear resistor



nMOS Saturation

- Channel pinches off
- I_{ds} independent of V_{ds}
- We say current *saturates*
- Similar to current source

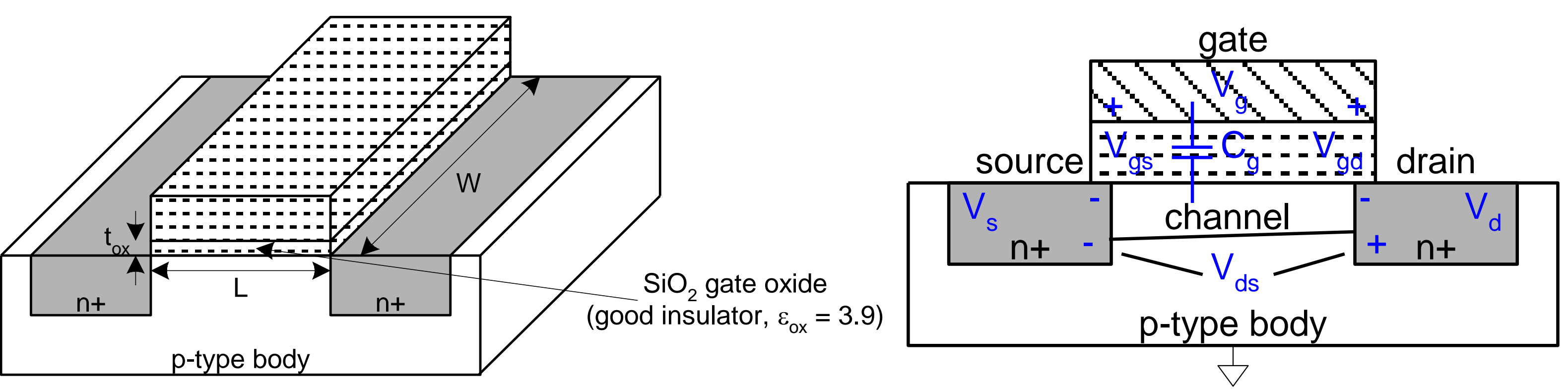


I-V Characteristics

- In Linear region, I_{ds} depends on
  - How much charge is in the channel?
  - How fast is the charge moving?

Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversions
  - Gate – oxide – channel
- $Q_{channel}$ =
- $C = C_g$ =
- V =



Carrier velocity

- Charge is carried by e-
- Electrons are propelled by the lateral electric field between source and drain
  - E =
- Carrier velocity *v* proportional to lateral E-field
  - *v* =
- Time for carrier to cross channel:
  - *t* =

nMOS Linear I-V

- Now we know
  - How much charge $Q_{channel}$ is in the channel
  - How much time *t* each carrier takes to cross

$$
\begin{aligned}
I_{ds} &= \frac{Q_{\text{channel}}}{t} \\
&= \mu C_{\text{ox}} \frac{W}{L} \left(V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds} \\
&= \beta \left(V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds}
\end{aligned}
$$

$$\beta = \mu C_{\text{ox}} \frac{W}{L}$$

nMOS Saturation I-V

- If $V_{gd} < V_t$, channel pinches off near drain
  When $V_{ds} > V_{dsat} = V_{gs} - V_t$
- Now drain voltage no longer increases current

$$I_{ds} = \beta \left(V_{gs} - V_t - \frac{V_{dsat}}{2} \right) V_{dsat}$$

$$= \frac{\beta}{2} \left(V_{gs} - V_t \right)^2$$

nMOS I-V Summary

- *Shockley* 1st order transistor models

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t & \text{cutoff} \\ \beta\left(V_{gs} - V_t - \frac{V_{ds}}{2}\right)V_{ds} & V_{ds} < V_{dsat} & \text{linear} \\ \frac{\beta}{2}\left(V_{gs} - V_t\right)^2 & V_{ds} > V_{dsat} & \text{saturation} \end{cases}$$

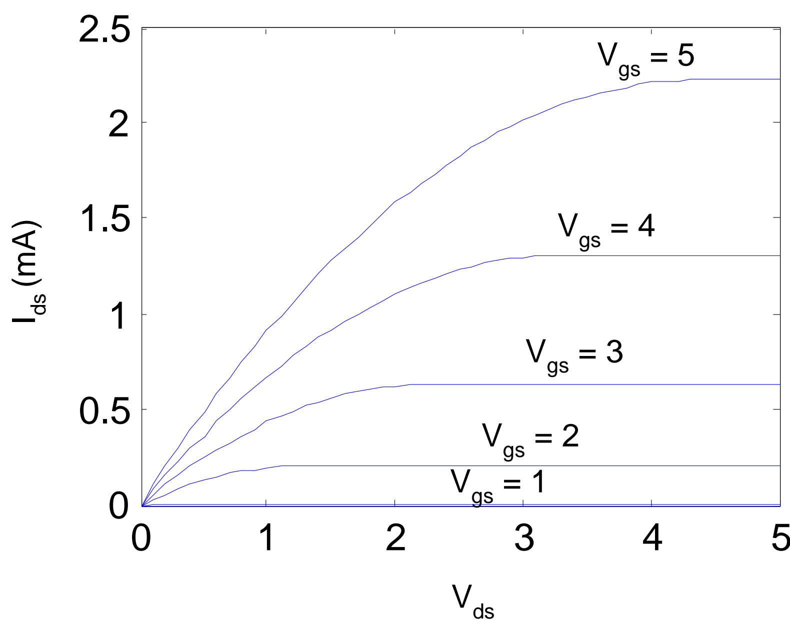
Example

- Plot I_{ds} vs. V_{ds} Given that
  - t_{ox} = 100 Å
  - μ = 350 $cm^2/V*s$
  - V_t = 0.7 V
  - $\xi_0=8.85*10^{-12}F/m$
  - $\xi_{0x}=3.9$
  - **V_{gs} = 0, 1, 2, 3, 4, 5**
  - **Use W/L = 4/2 λ**

THE UNIVERSITY OF
JORDAN

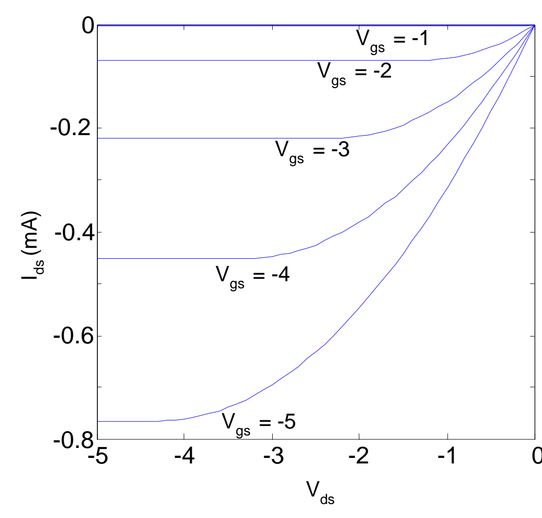
Example

$$\beta = \mu C_{ox} \frac{W}{L} = (350)\left(\frac{3.9 \times 8.85 \cdot 10^{-14}}{100 \cdot 10^{-8}}\right)\left(\frac{W}{L}\right) = 120 \frac{W}{L} \mu\text{A/V}^2$$



pMOS I-V

- All dopings and voltages are inverted for pMOS
  - Source is the more positive terminal
- Mobility μ_p is determined by holes
  - Typically 2-3x lower than that of electrons μ_n
  - 120 $cm^2/V \bullet s$ in AMI 0.6 μm process
- Thus pMOS must be wider to provide same current
  - In this class, assume $\mu_n / \mu_p = 2$



Problem

- NMOS Transistor
- V_{th}=.7 V_s=1V V_G=2.5 V_d=4V B=23uA/V^2
- Find W/L

The Threshold Voltage (1)(Optional)

- Physical components of the threshold voltage of a MOS structure
  - The work function difference between the gate and the channel.
  - The gate component to change the surface potential.
  - The gate voltage component to offset the depletion region charge.
  - The voltage component to offset the fixed charges in the gate oxide and in the silicon-oxide interface.

The Threshold Voltage (2) (Optional)

- The work function difference Φ_{GC} between the gate and the channel determines the built-in potential of the MOS system.

  - For metal gate

$$\Phi_{GC} = \phi_F(substrate) - \phi_M$$

  - For polysilicon gate

$$\Phi_{GC} = \phi_F(substrate) - \phi_F(gate)$$

The Threshold Voltage (3) (Optional)

- Because of the fixed acceptor ions located in the depletion region near the surface, depletion charge exists.
  - Depletion region charge

$$Q_{B0} = -\sqrt{2q \cdot N_A \cdot \varepsilon_{Si} \cdot |-2\phi_F|}$$

  - Consider the voltage bias of the body.

$$Q_B = -\sqrt{2q \cdot N_A \cdot \varepsilon_{Si} \cdot |-2\phi_F + V_{SB}|}$$

- The component that offsets the depletion region charge is equal to:

$$-Q_B / C_{OX}$$

$$C_{OX} = \frac{\varepsilon_{ox}}{t_{ox}}$$

The Threshold Voltage (4) (Optional)

- There always exists a fixed positive charge density Q_{ox} at the interface between the gate oxide and the silicon substrate.
- The gate voltage component that is necessary to offset this positive charge at the interface is .

  - For zero substrate bias

$$V_{T0} = \Phi_{GC} - 2\phi_F(substrate) - \frac{Q_{B0}}{C_{ox}}$$

The Threshold Voltage (optional)

- We can use the (3.23) for both n-channel device and p-channel device.
- However, some of the terms and coefficient in (3.23) have different polarities for the n-channel case and for the p-channel case.
  - The substrate Fermi potential ϕ_F is negative in nMOS, positive in pMOS.
  - The depletion region charge density Q_{B0} and Q_B are negative in nMOS, positive pMOS.
  - The substrate bias coefficient γ is positive in nMOS, negative in pMOS.
  - The substrate bias voltage V_{SB} is positive in nMOS, negative in pMOS.

Example 3.2 (1) (optional)

- Calculate the threshold voltage V_{T0} (@ V_{SB}=0).
  - N_A = 4 x 10^{18} cm^{-3}
  - N_D = 2 x 10^{20} cm^{-3}
  - t_{ox} = 26.3 Å
  - $N_{i=}1.45*10^{18}$

- Sol.
  - Calculate the Fermi potentials

$$\phi_F(substrate) = \frac{kT}{q}\ln\left(\frac{n_i}{N_A}\right) = 0.026\text{V}\cdot\ln\left(\frac{1.45\cdot 10^{10}}{4\times 10^{18}}\right) = -0.51\text{V}$$

  - Calculate the work function difference

$$\Phi_{GC} = \phi_F(substrate) - \phi_F(gate) = -0.51\text{V} - 0.55\text{V} = -1.06\text{V}$$

Example 3.2 (2) (optional)

- Sol.(Cont'd)
  - The depletion region charge density at V_{SB} = 0

$$Q_{B0} = -\sqrt{2 \cdot q \cdot N_A \cdot \varepsilon_{Si} \cdot |-2\phi_F(substrate)|}$$
$$= -\sqrt{2 \cdot 1.6 \cdot 10^{-19} \cdot (4 \times 10^{18}) \cdot 11.7 \cdot 8.85 \cdot 10^{-14} \cdot |-2 \cdot 0.51|}$$
$$= -1.16 \cdot 10^{-6}\,\mathrm{C/cm^2}$$

  - The gate oxide capacitance per unit area

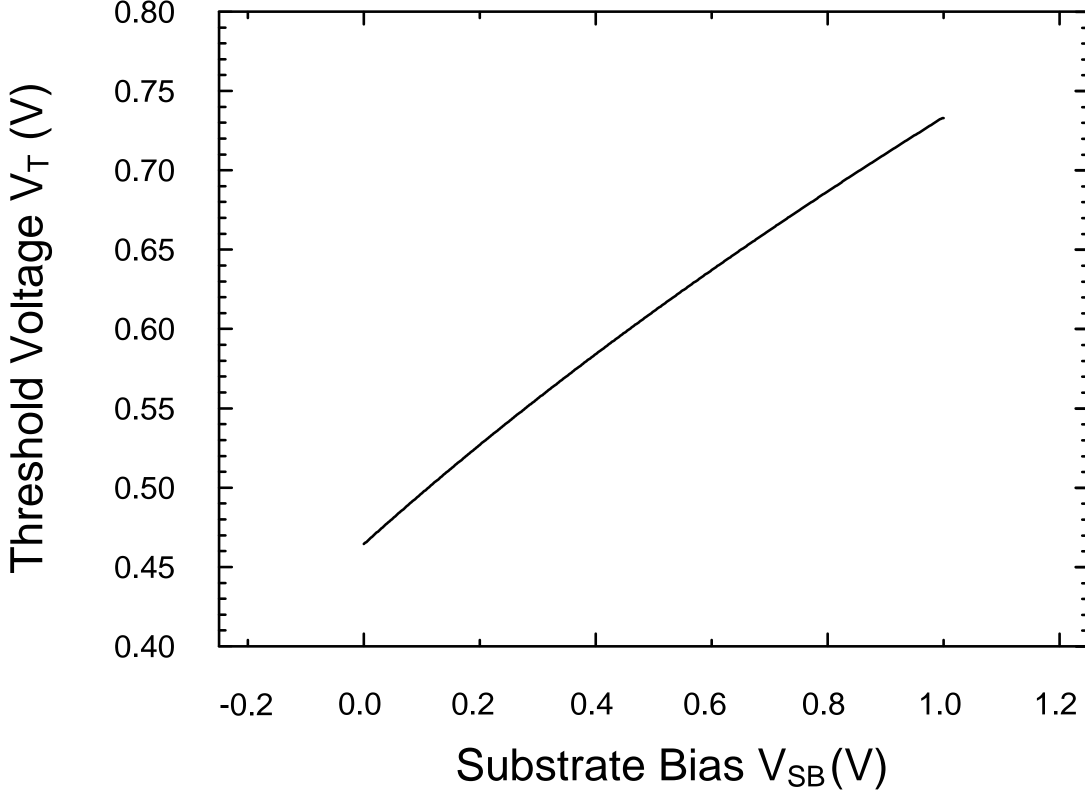
$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} = \frac{3.97 \cdot 8.85 \cdot 10^{-14}\,\mathrm{F/cm}}{1.6 \cdot 10^{-7}\,\mathrm{cm}} = 2.2 \cdot 10^{-6}\,\mathrm{F/cm^2}$$

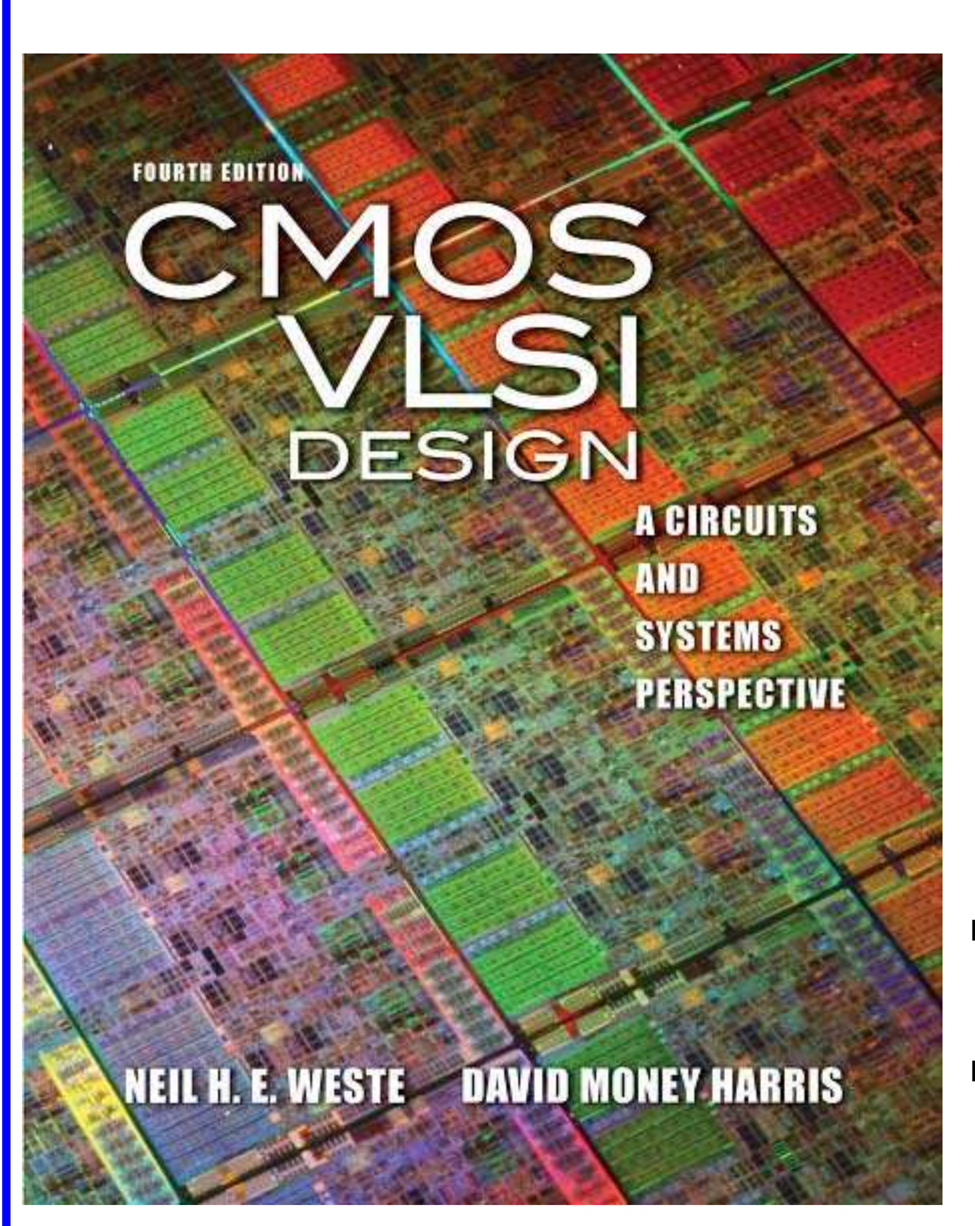
  - Combine all components

$$V_{T0} = \Phi_{GC} - 2\phi_F(substrate) - \frac{Q_{B0}}{C_{ox}}$$
$$= -1.06 - (-1.02) - (-0.53)$$

Body effect

$$V_T = V_{T0} + \gamma\left(\sqrt{\left|-2\phi_F + V_{SB}\right|} - \sqrt{\left|2\phi_F\right|}\right)$$


0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
Threshold Voltage V_T (V)
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Substrate Bias V_{SB} (V)



Lecture 4: Nonideal Transistor Theory

Outline

- Nonideal Transistor Behavior
  - High Field Effects
    - Mobility Degradation
    - Velocity Saturation
  - Channel Length Modulation
  - Threshold Voltage Effects
    - Body Effect
    - Drain-Induced Barrier Lowering
    - Short Channel Effect
  - Leakage
    - Subthreshold Leakage
    - Gate Leakage
    - Junction Leakage
- Process and Environmental Variations

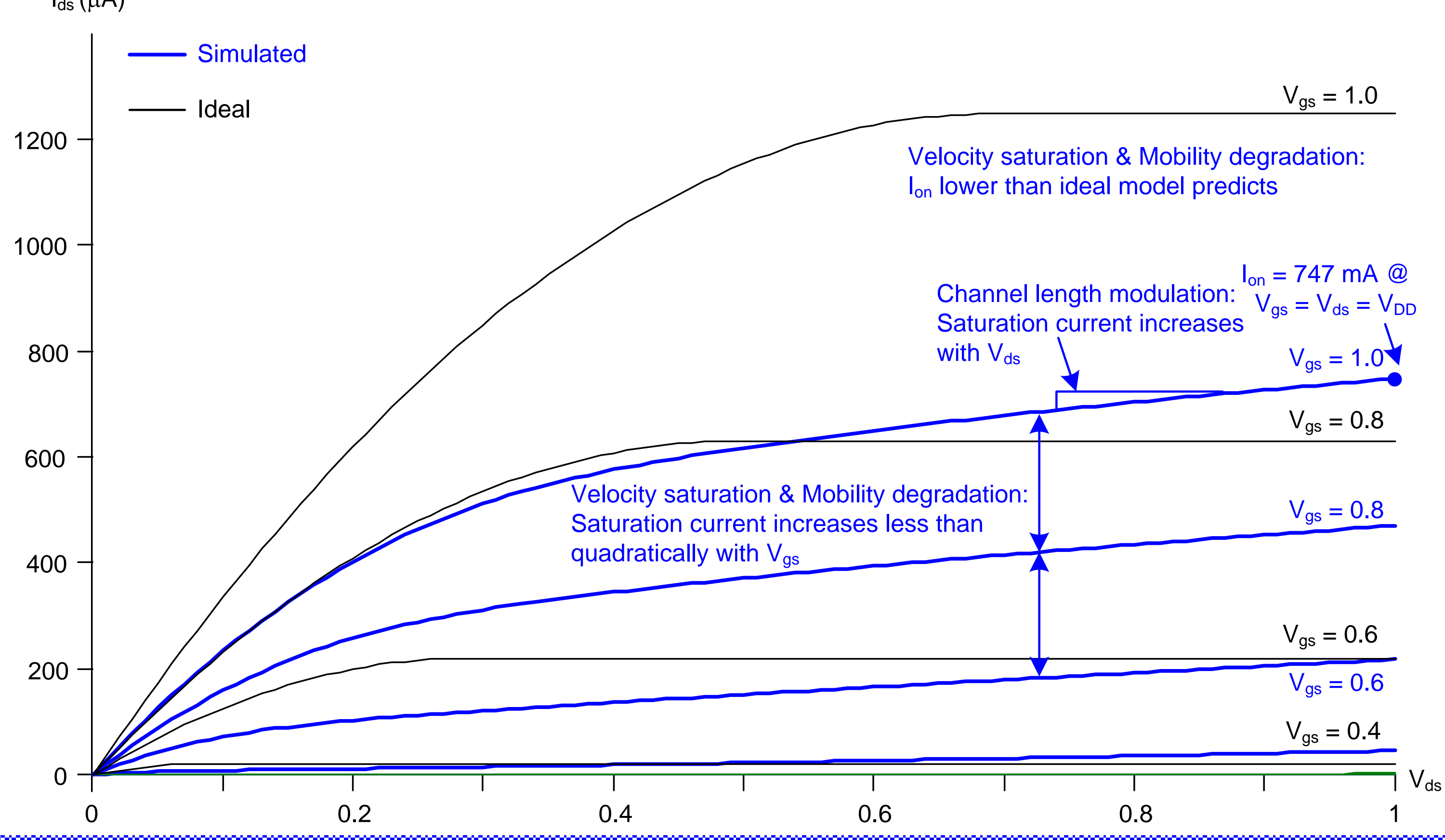
Ideal Transistor I-V

- *Shockley* long-channel transistor models

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t & \text{cutoff} \\ \beta\left(V_{gs} - V_t - \frac{V_{ds}}{2}\right)V_{ds} & V_{ds} < V_{dsat} & \text{linear} \\ \frac{\beta}{2}\left(V_{gs} - V_t\right)^2 & V_{ds} > V_{dsat} & \text{saturation} \end{cases}$$

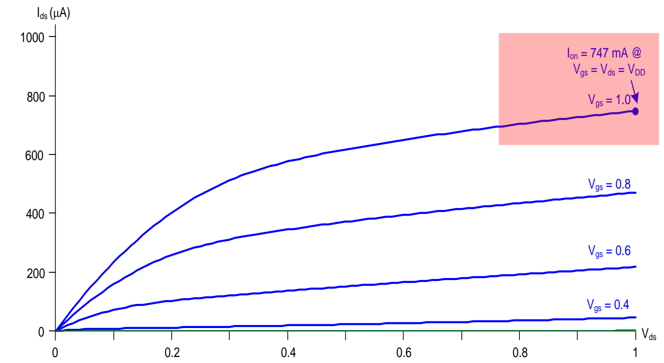
Ideal vs. Simulated nMOS I-V Plot

- 65 nm IBM process, $V_{DD} = 1.0$ V

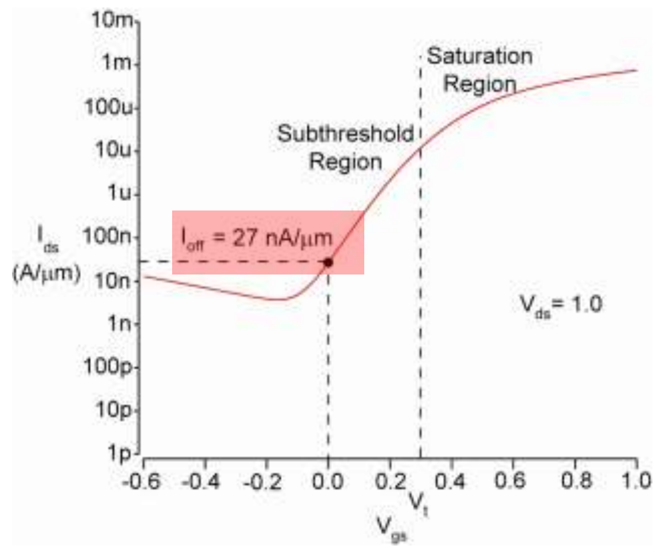


ON and OFF Current

- $I_{on} = I_{ds}$ @ $V_{gs} = V_{ds} = V_{DD}$
  - Saturation



- $I_{off} = I_{ds}$ @ $V_{gs} = 0$, $V_{ds} = V_{DD}$
  - Cutoff



Electric Fields Effects

- Vertical electric field: E_{vert} = ________
  - Attracts carriers into channel
  - Long channel: $Q_{channel} \propto E_{vert}$
- Lateral electric field: E_{lat} = ________
  - Accelerates carriers from drain to source
  - Long channel: $v = \mu E_{lat}$

Coffee Cart Analogy

- Tired student runs from VLSI lab to coffee cart
- Freshmen are pouring out of the physics lecture hall
- V_{ds} is how long you have been up
  - Your velocity = fatigue × mobility
- V_{gs} is a wind blowing you against the glass (SiO_2) wall
- At high V_{gs}, you are buffeted against the wall
  - *Mobility degradation*
- At high V_{ds}, you scatter off freshmen, fall down, get up
  - *Velocity saturation*
    - Don’t confuse this with the saturation region

Mobility Degradation

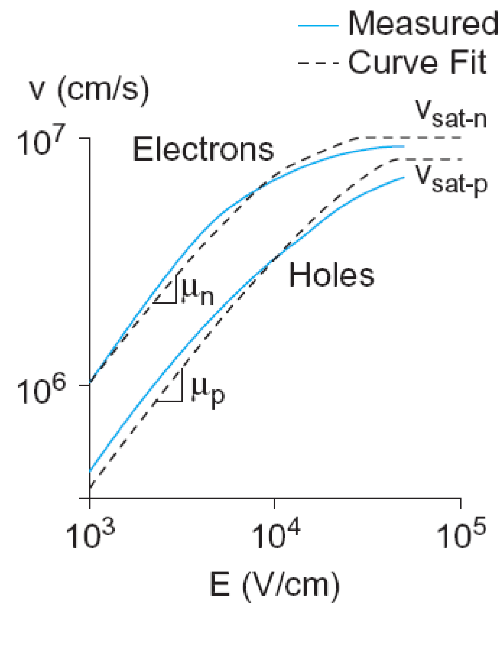
- High E_{vert} effectively reduces mobility
  - Collisions with oxide interface

$$\mu_{\text{eff}-n} = \frac{540 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}}{1 + \left(\frac{V_{gs} + V_t}{0.54 \frac{\text{V}}{\text{nm}} t_{\text{ox}}} \right)^{1.85}} \qquad \mu_{\text{eff}-p} = \frac{185 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}}{1 + \frac{\left| V_{gs} + 1.5 V_t \right|}{0.338 \frac{\text{V}}{\text{nm}} t_{\text{ox}}}}$$

Velocity Saturation

- At high E_{lat}, carrier velocity rolls off
  - Carriers scatter off atoms in silicon lattice
  - Velocity reaches v_{sat}
    - Electrons: 10^7 cm/s
    - Holes: 8×10^6 cm/s
  - Better model

$$v = \begin{cases} \dfrac{\mu_{\text{eff}} E}{1 + \dfrac{E}{E_c}} & E < E_c \\ v_{\text{sat}} & E \ge E_c \end{cases} \qquad E_c = \frac{2 v_{\text{sat}}}{\mu_{\text{eff}}}$$



Vel Sat I-V Effects

- Ideal transistor ON current increases with V_{DD}^2

$$I_{ds} = \mu C_{ox} \frac{W}{L} \frac{\left(V_{gs} - V_t\right)^2}{2} = \frac{\beta}{2}\left(V_{gs} - V_t\right)^2$$

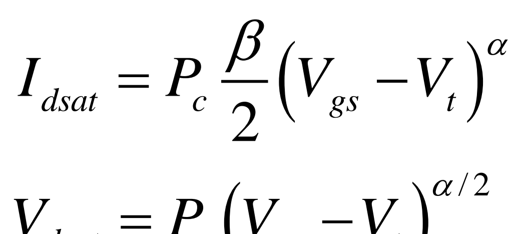
- Velocity-saturated ON current increases with V_{DD}

$$I_{ds} = C_{ox} W \left(V_{gs} - V_t\right) v_{max}$$

- Real transistors are partially velocity saturated
  - Approximate with α-power law model
  - $I_{ds} \propto V_{DD}^{\alpha}$
  - $1 < \alpha < 2$ determined empirically (≈ 1.3 for 65 nm)

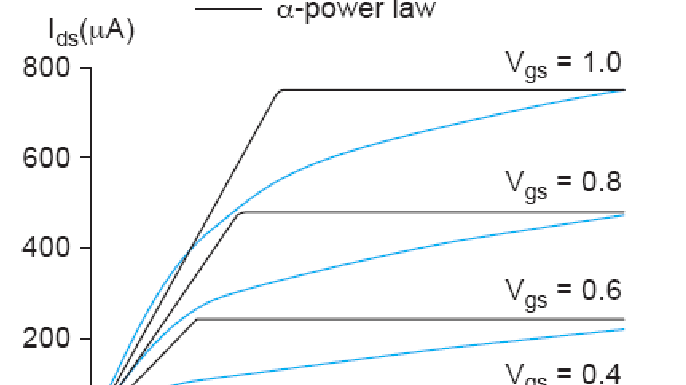
α-Power Model

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t \quad \text{cutoff} \\ I_{dsat} \dfrac{V_{ds}}{V_{dsat}} & V_{ds} < V_{dsat} \quad \text{linear} \\ I_{dsat} & V_{ds} > V_{dsat} \quad \text{saturation} \end{cases}$$



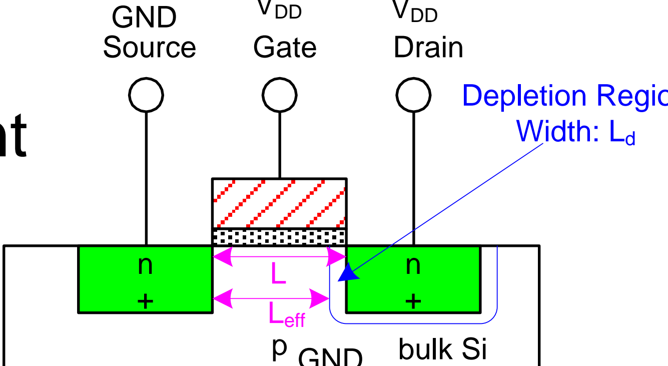
$$I_{dsat} = P_c \frac{\beta}{2}\left(V_{gs} - V_t\right)^{\alpha}$$

$$V_{dsat} = P_v \left(V_{gs} - V_t\right)^{\alpha/2}$$



Channel Length Modulation

- Reverse-biased p-n junctions form a *depletion region*
  - Region between n and p with no carriers
  - Width of depletion L_d region grows with reverse bias
  - $L_{eff} = L - L_d$
- Shorter L_{eff} gives _____ current
  - I_{ds} __________ with V_{ds}
  - Even in saturation



Chan Length Mod I-V

$$I_{ds} = \frac{\beta}{2}\left(V_{gs} - V_t\right)^2\left(1 + \lambda V_{ds}\right)$$

- λ = *channel length modulation coefficient*
  - not feature size
  - Empirically fit to I-V characteristics

Threshold Voltage Effects

- V_t is V_{gs} for which the channel starts to invert
- Ideal models assumed V_t is constant
- Really depends (weakly) on almost everything else:
  - Body voltage: *Body Effect*
  - Drain voltage: *Drain-Induced Barrier Lowering*
  - Channel length: *Short Channel Effect*

Body Effect

- Body is a fourth transistor terminal
- V_{sb} affects the charge required to invert the channel
  - Increasing V_s or decreasing V_b increases V_t

$$V_t = V_{t0} + \gamma\left(\sqrt{\phi_s + V_{sb}} - \sqrt{\phi_s}\right)$$

- ϕ_s = *surface potential* at threshold

$$\phi_s = 2v_T \ln\frac{N_A}{n_i}$$

  - Depends on doping level N_A
  - And intrinsic carrier concentration n_i
- γ = *body effect coefficient*

$$\gamma = \frac{t_{ox}}{\varepsilon_{ox}}\sqrt{2q\varepsilon_{si}N_A} = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C_{ox}}$$

Body Effect Cont.

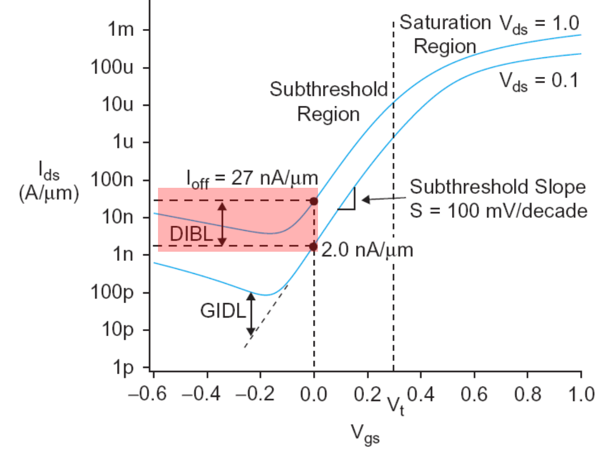
- For small source-to-body voltage, treat as linear

$$V_t = V_{t0} + k_\gamma V_{sb}$$

$$k_\gamma = \frac{\gamma}{2\sqrt{\phi_s}} = \frac{\sqrt{\dfrac{q\varepsilon_{si} N_A}{v_T \ln \frac{N_A}{n_i}}}}{2C_{ox}}$$

DIBL

- Electric field from drain affects channel
- More pronounced in small transistors where the drain is closer to the channel
- Drain-Induced Barrier Lowering
  - Drain voltage also affect V_t



$$V_t' = V_t - \eta V_{ds}$$

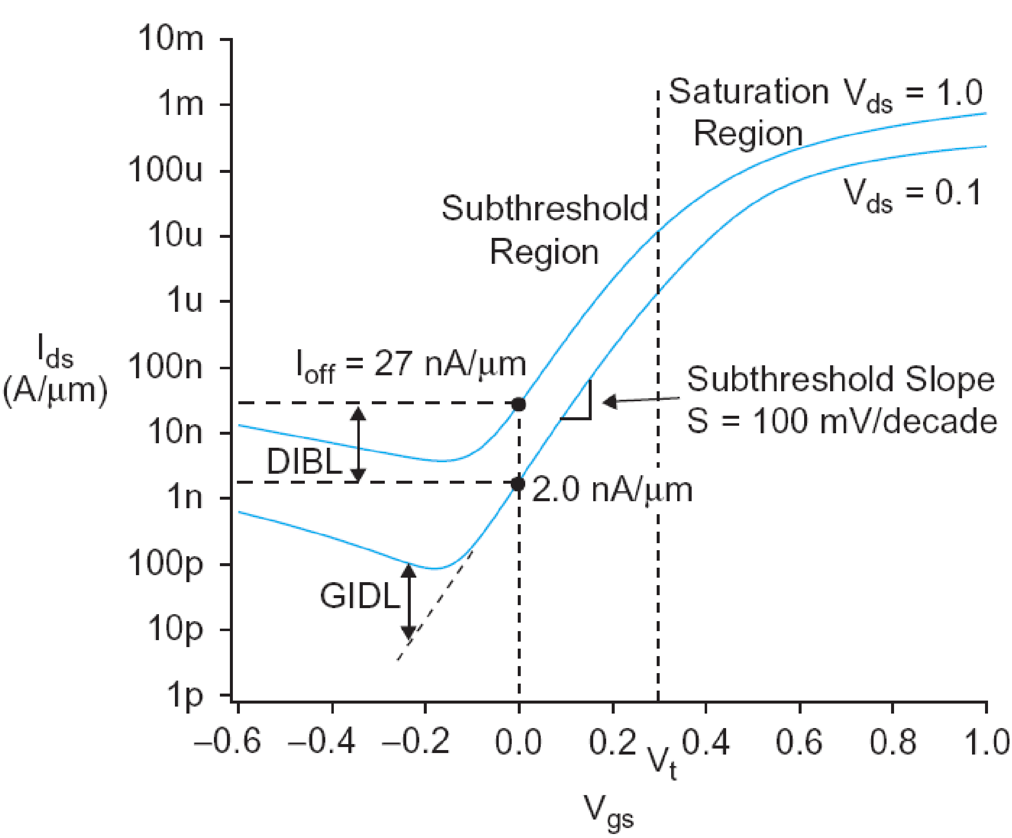
- High drain voltage causes current to ________

Short Channel Effect

- In small transistors, source/drain depletion regions extend into the channel
  - Impacts the amount of charge required to invert the channel
  - And thus makes V_t a function of channel length
- Short channel effect: V_t increases with L
  - Some processes exhibit a reverse short channel effect in which V_t decreases with L

Leakage

- What about current in cutoff?
- Simulated results
- What differs?
  - –



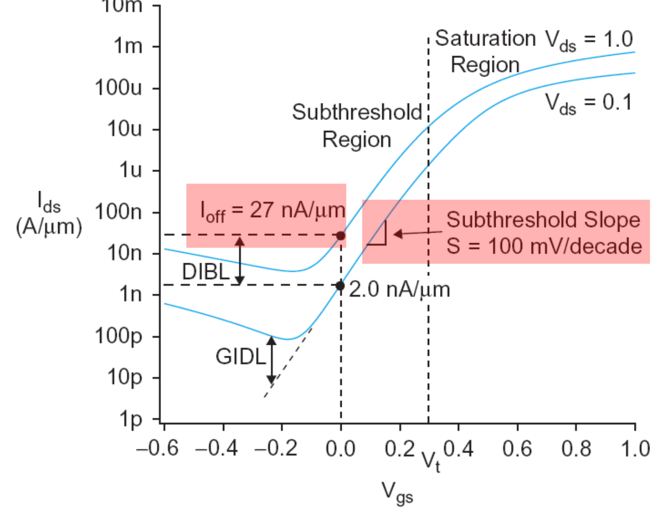
Leakage Sources

- Subthreshold conduction
  - Transistors can't abruptly turn ON or OFF
  - Dominant source in contemporary transistors
- Gate leakage
  - Tunneling through ultrathin gate dielectric
- Junction leakage
  - Reverse-biased PN junction diode current

Subthreshold Leakage

- Subthreshold leakage exponential with V_{gs}

$$I_{ds} = I_{ds0} e^{\frac{V_{gs} - V_{t0} + \eta V_{ds} - k_\gamma V_{sb}}{n v_T}} \left(1 - e^{\frac{-V_{ds}}{v_T}}\right)$$



- n is process dependent
  - typically 1.3-1.7
- Rewrite relative to I_{off} on log scale

$$I_{ds} = I_{off} 10^{\frac{V_{gs} + \eta(V_{ds} - V_{dd}) - k_\gamma V_{sb}}{S}} \left(1 - e^{\frac{-V_{ds}}{v_t}}\right)$$

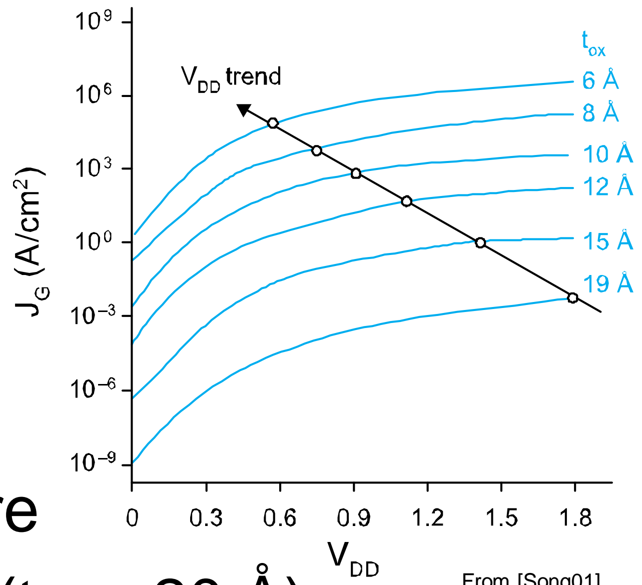
$$S = \left[\frac{d\left(\log_{10} I_{ds}\right)}{dV_{gs}}\right]^{-1} = n v_T \ln 10$$

- S ≈ 100 mV/decade @ room temperature

Gate Leakage

- Carriers tunnel thorough very thin gate oxides
- Exponentially sensitive to t_{ox} and V_{DD}

$$I_{gate} = WA\left(\frac{V_{DD}}{t_{ox}}\right)^2 e^{-B\frac{t_{ox}}{V_{DD}}}$$

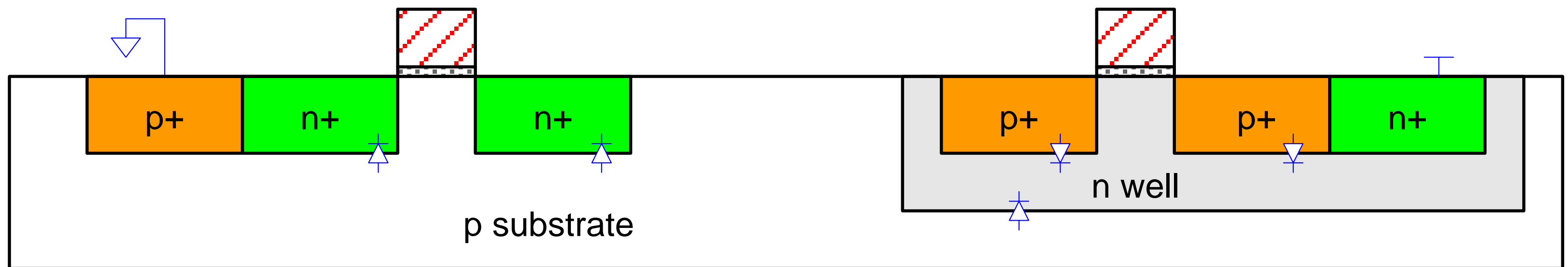


  - A and B are tech constants
  - Greater for electrons
    - So nMOS gates leak more

- Negligible for older processes (t_{ox} > 20 Å)
- Critically important at 65 nm and below (t_{ox} ≈ 10.5 Å)

Junction Leakage

- Reverse-biased p-n junctions have some leakage
  - Ordinary diode leakage
  - Band-to-band tunneling (BTBT)
  - Gate-induced drain leakage (GIDL)



Diode Leakage

- Reverse-biased p-n junctions have some leakage

$$I_D = I_S \left(e^{\frac{V_D}{v_T}} - 1 \right)$$

- At any significant negative diode voltage, $I_D = -I_s$
- I_s depends on doping levels
  - And area and perimeter of diffusion regions
  - Typically < 1 fA/μm^2 (negligible)

Band-to-Band Tunneling

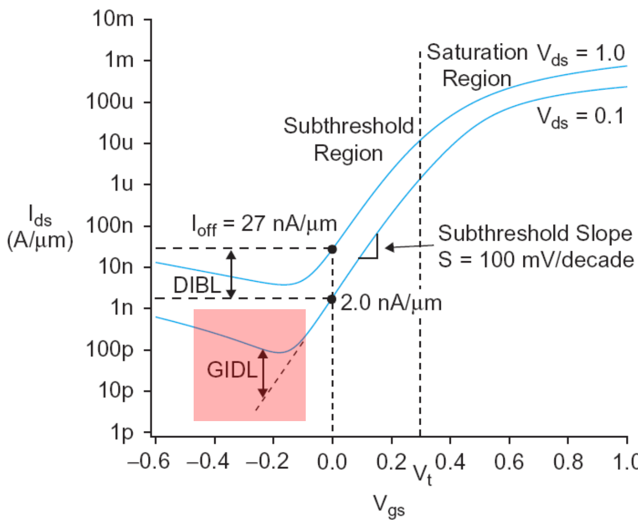
- Tunneling across heavily doped p-n junctions
  - Especially sidewall between drain & channel when *halo doping* is used to increase V_t
- Increases junction leakage to significant levels

$$I_{BTBT} = WX_j A \frac{E_j}{E_g^{0.5}} V_{dd} e^{-B\frac{E_g^{1.5}}{E_j}} \qquad E_j = \sqrt{\frac{2qN_{halo}N_{sd}}{\varepsilon\left(N_{halo}+N_{sd}\right)}\left(V_{DD}+v_T \ln\frac{N_{halo}N_{sd}}{n_i^2}\right)}$$

- X_j: sidewall junction depth
- E_g: bandgap voltage
- A, B: tech constants

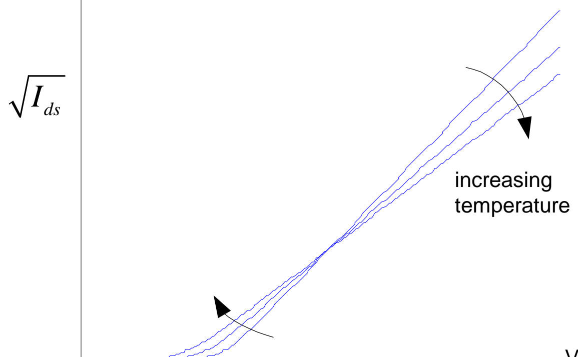
Gate-Induced Drain Leakage

- Occurs at overlap between gate and drain
  - Most pronounced when drain is at V_{DD}, gate is at a negative voltage
  - Thwarts efforts to reduce subthreshold leakage using a negative gate voltage



Temperature Sensitivity

- Increasing temperature
  - Reduces mobility
  - Reduces V_t
- I_{ON} ___________ with temperature
- I_{OFF} ___________ with temperature

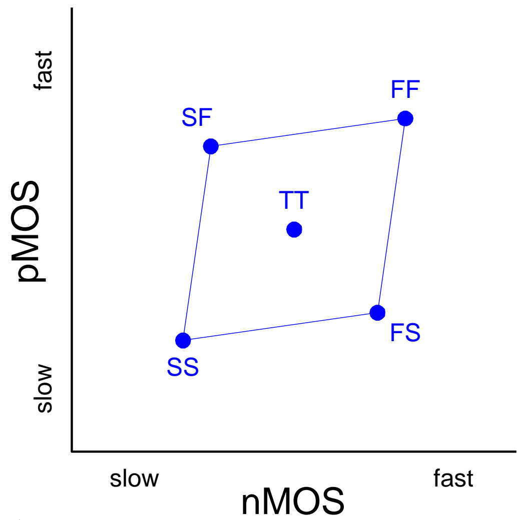


So What?

- So what if transistors are not ideal?
  - They still behave like switches.
- But these effects matter for…
  - Supply voltage choice
  - Logical effort
  - Quiescent power consumption
  - Pass transistors
  - Temperature of operation

Parameter Variation

- Transistors have uncertainty in parameters
  - Process: L_{eff}, V_t, t_{ox} of nMOS and pMOS
  - Vary around typical (T) values
- Fast (F)
  - L_{eff}: ________
  - V_t: ________
  - t_{ox}: ________
- Slow (S): opposite
- Not all parameters are independent for nMOS and pMOS



Environmental Variation

- V_{DD} and T also vary in time and space
- Fast:
  - V_{DD}: ________
  - T: ________

Corner	Voltage	Temperature
F		
T	1.8	70 C
S		

Process Corners

- Process corners describe worst case variations
  - If a design works in all corners, it will probably work for any variation.
- Describe corner with four letters (T, F, S)
  - nMOS speed
  - pMOS speed
  - Voltage
  - Temperature

Important Corners

- Some critical simulation corners include

Purpose	nMOS	pMOS	V_{DD}	Temp
Cycle time				
Power				
Subthreshold leakage				

Inverter Static Characteristics

Dr.Mohammad Abdel-Majeed

Assistant Professor

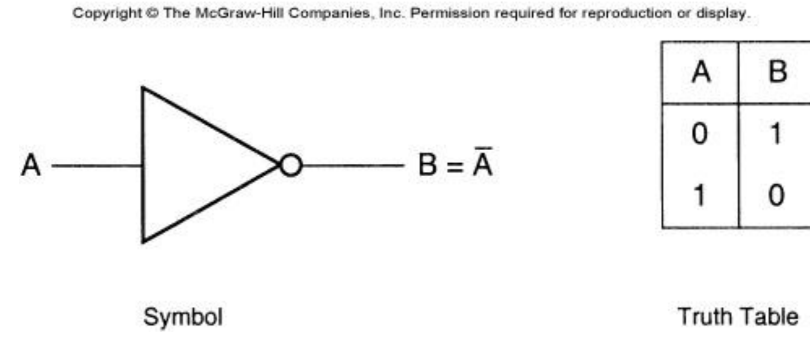
University of Jordan

Combinational Logic

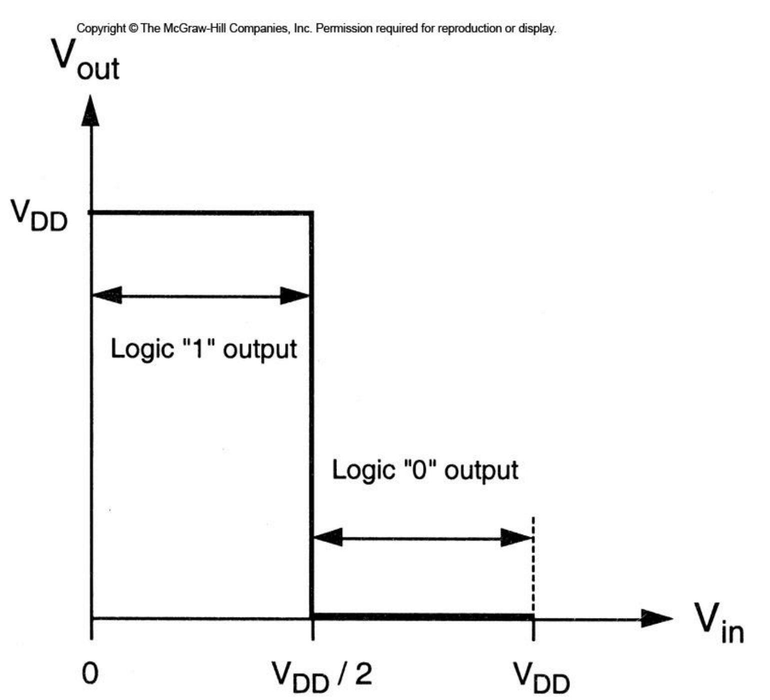
- A combinational logic cell, logic circuit or gate is generally a multiple input, single output system that performs a Boolean function
  • In the positive logic convention, logic 1 is shown by high voltage VDD and logic 0 by low voltage of zero

Ideal Inverter and Inverter Threshold Voltage

- Logic symbol & truth table

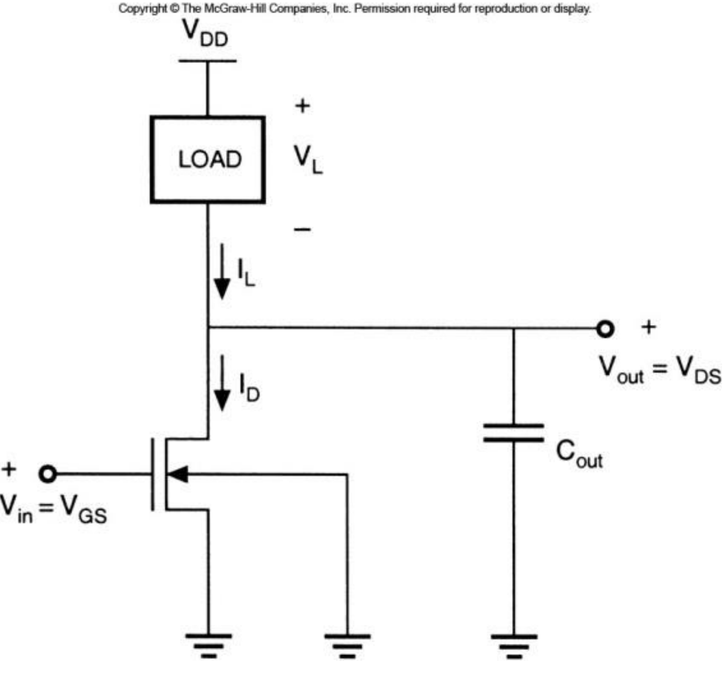


A	B
0	1
1	0

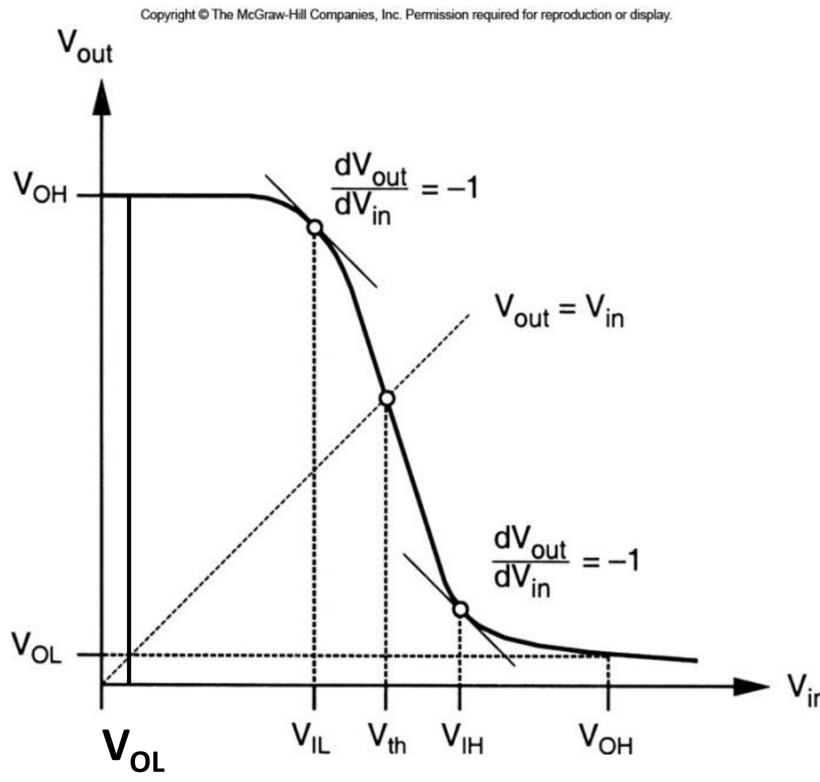


VTC of an NMOs inverter

- Vin is ______
- Vout is ________
- VSB = ________
- The circuit connected to the output node can be represented using the capacitance C_{out}
- $I_D = I_L$ (__________). By solving this equation we can derive the VTC



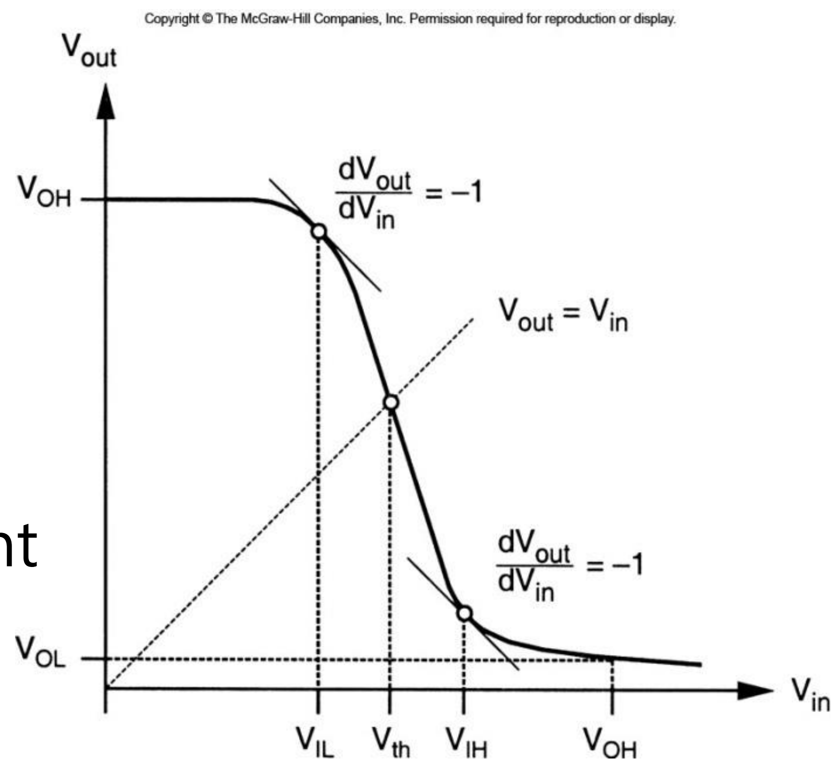
VTC



V_{OL}

nMOS Inverter: Schematic & VTC

- V_{OH} : $V_{OUT,MAX}$ when the output level is logic "1"
- V_{OL} : $V_{OUT,MIN}$ when the output level is logic "0"
- V_{IL} : $V_{IN,MAX}$ which can be *interpreted* as logic "0"
- V_{IH} : $V_{IN,MIN}$ which can be *interpreted* as logic "1"
- Inverter threshold Voltage V_M is defined as the point where $V_{in} = V_{out}$



- _____________________<= Vin <=____
  - Vin is interpreted as “0”
  - This means that V_{IL} is low enough to ensure a logic 1 output

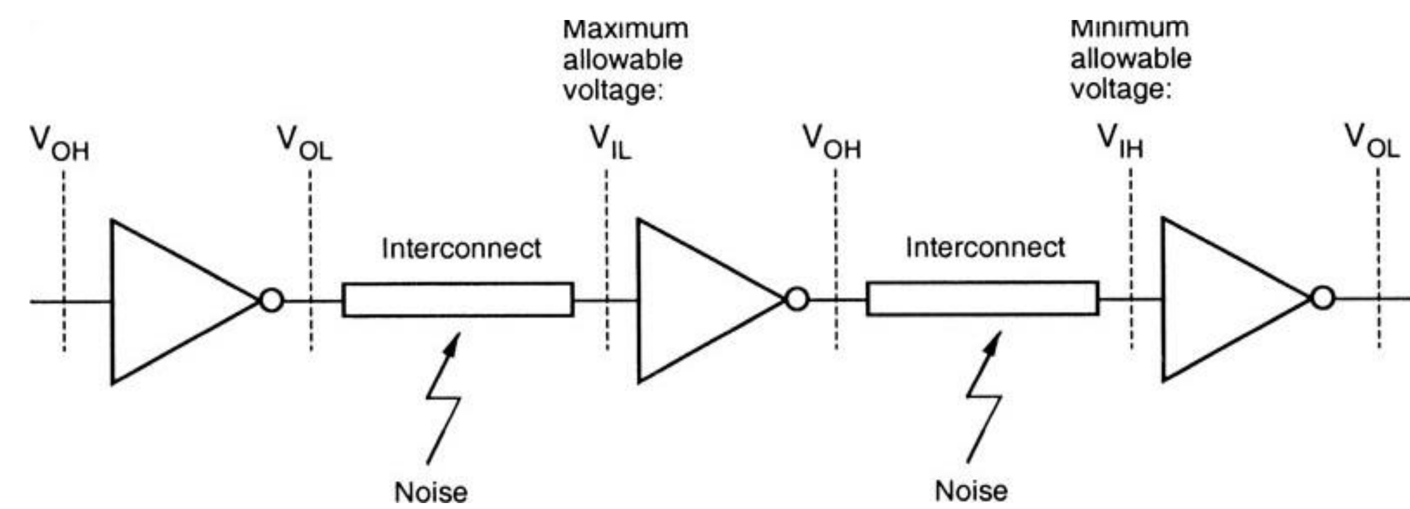
- _____<= Vin <=____
  - Vin is interpreted as “1”
  - This means that V_{IH} is low enough to ensure a logic 0 output

Tolerance

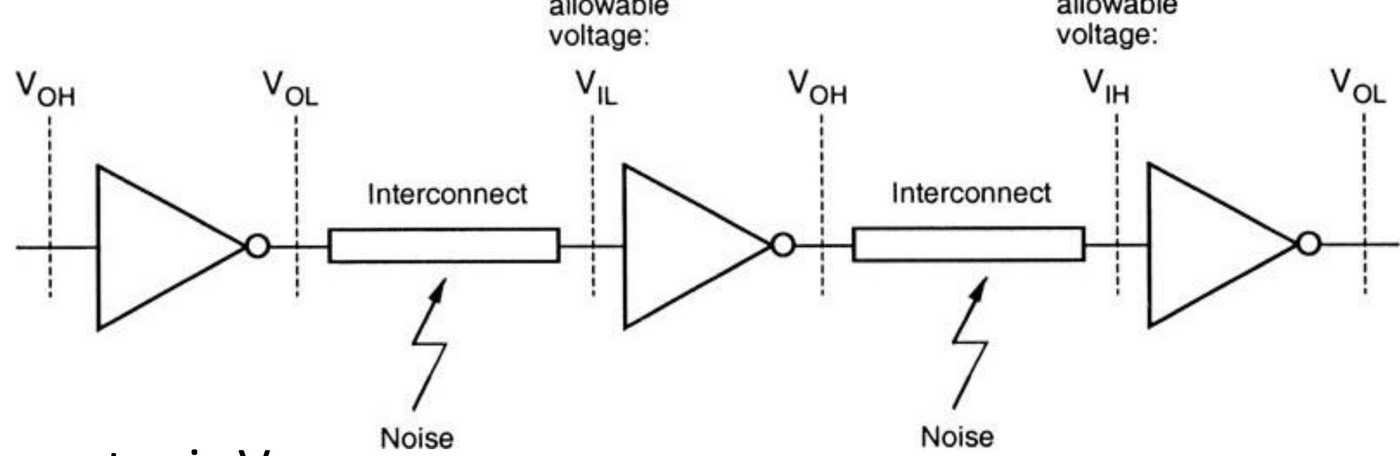
- Ability to interpret range of values as logic 0 or logic 1 allows the circuits to operate with certain ___________

- Noise sources
  - Unwanted capacitive coupling
  - Radiations
  - …

Noise Margin

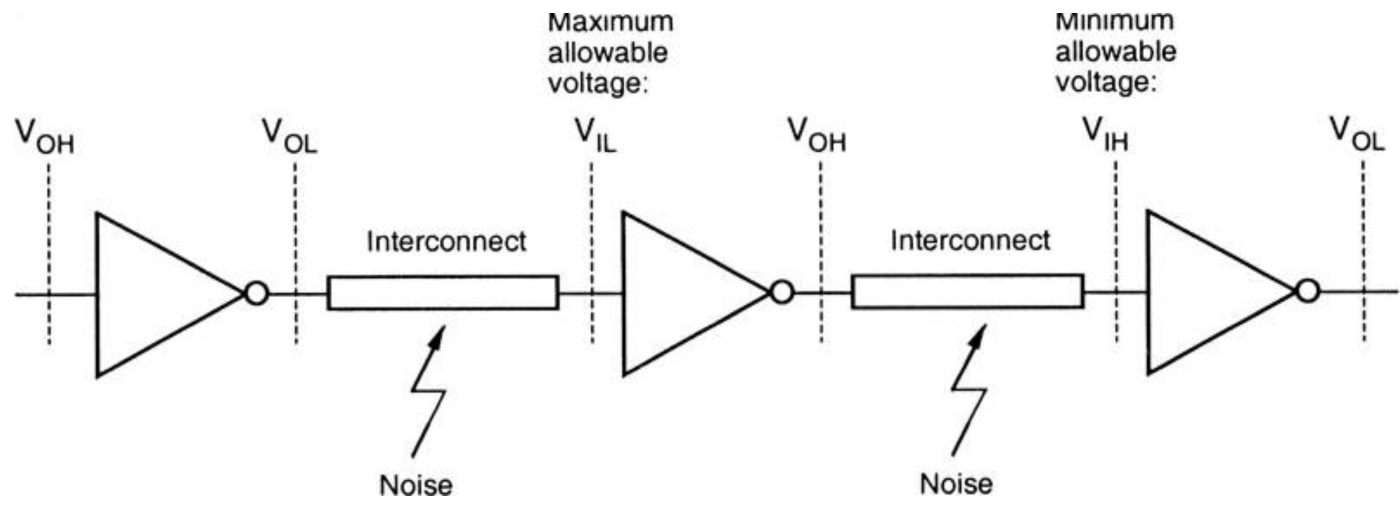


Input Noise Margin

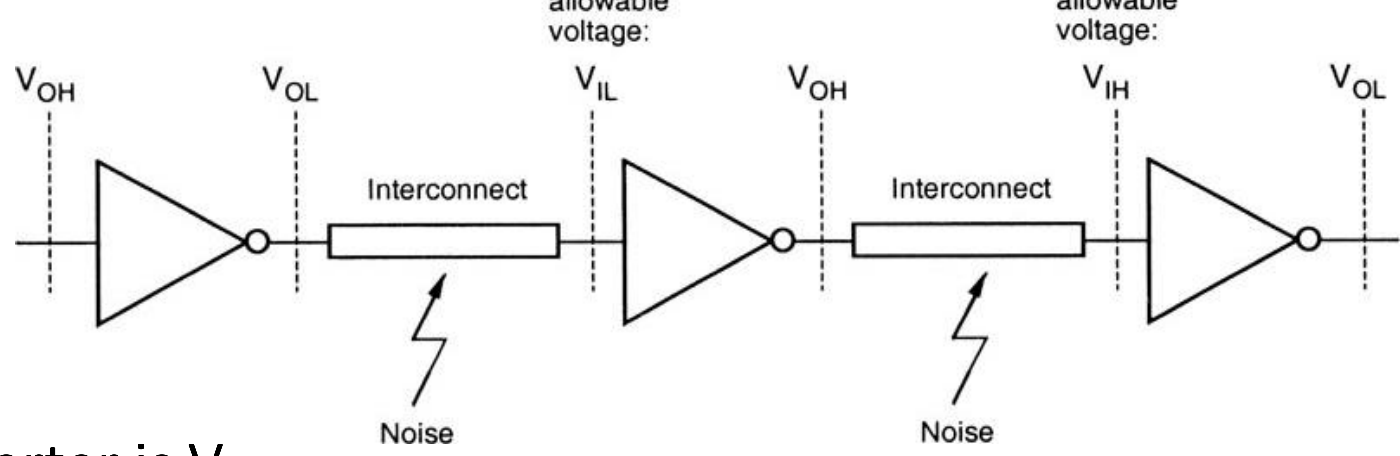


- By definition the output of the first inverter is V_{OL}
- The output signal of the 1st inverter will be perturbed during transmission because of on-chip interconnect
- If the input voltage of the 2nd inverter is smaller than V_{OL}, then this signal will be interpreted correctly as a logic "0" input by 2nd inverter
- But if the input voltage becomes larger than V_{IL}, then it may not be interpreted correctly by the inverter
- Thus, V_{IL} is the maximum allowable voltage at the input of the 2nd inverter

Output Noise Margin



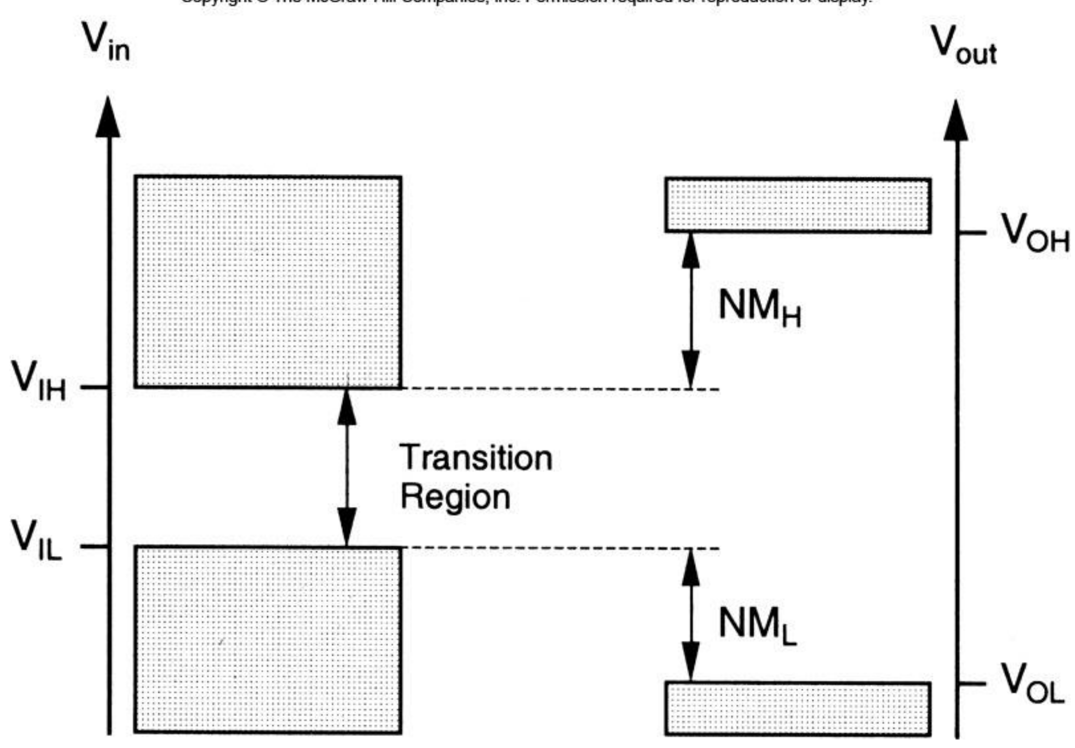
Input Noise Margin



- Assume the output of the 2nd inverter is V_{OH}
- Input of the 3rd inverter will be different from V_{OH} due to noise interference
- If the input voltage of the 3rd inverter is larger than V_{OH} this signal is interpreted correctly as a logic "1"
- If the voltage level drops below V_{IH} , the input may not be interpreted as a logic "1"
- Therefore, V_{IH} is the minimum allowable voltage at the input of the 3rd inverter

Noise Margin

- NM_L =
- NM_H =



Power

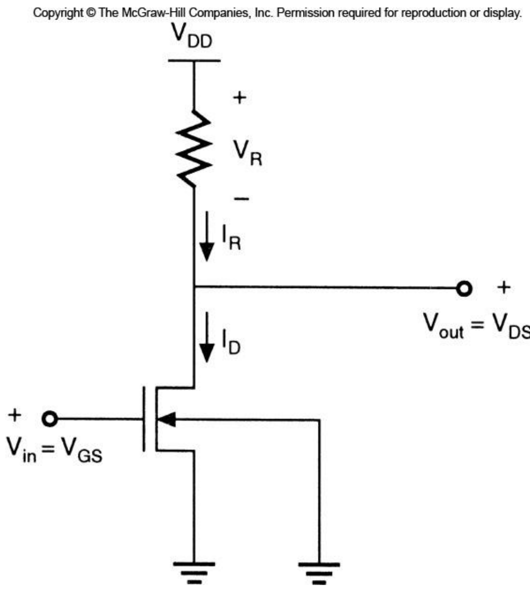
- Lower consumption extends battery-based operation time for portable systems
- DC power dissipation: $P_{DC} = V_{DD}.I_{DC}$
- DC power is input and output voltage dependent
- To calculate avg DC power assume 50% of the time input is at V_{in} = low and 50% at V_{in}= high
  - $P_{DC} = 0.5\ VDD\ \{I_{DC}\ (V_{in}=low) + I_{DC}\ (V_{in}=high)\}$
- Inverter type and its design affects the power dissipation significantly

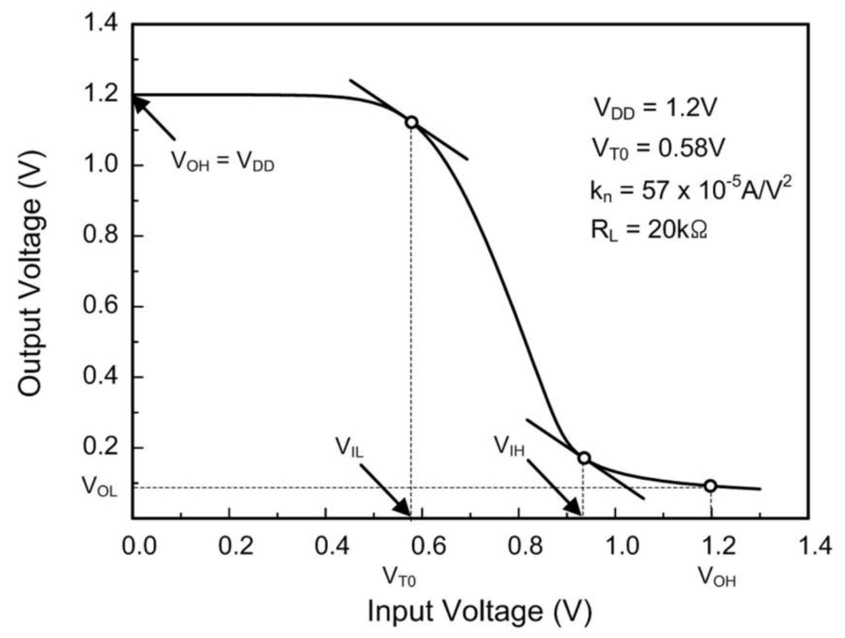
Inverter

- Resistive Load Inverter
- NMOS depletion load inverter
- CMOS inverter

Resistive-Load Inverter

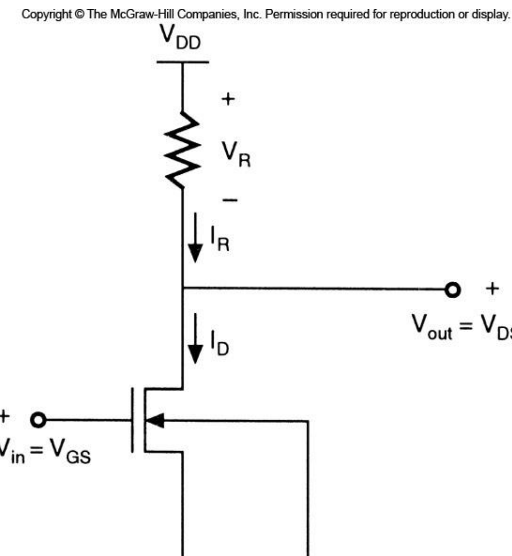
- Resistive-load inverter circuit & its VTC





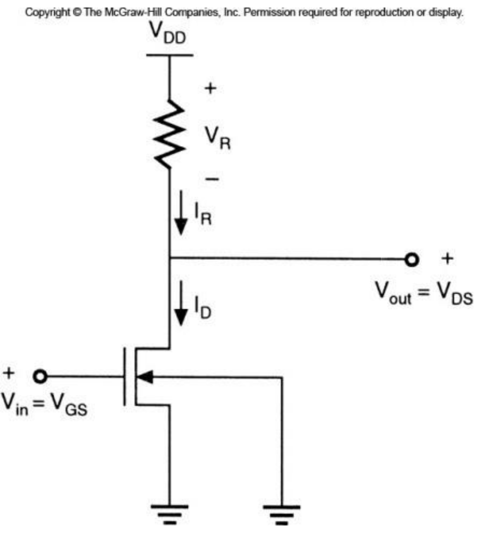
Resistive-Load Inverter

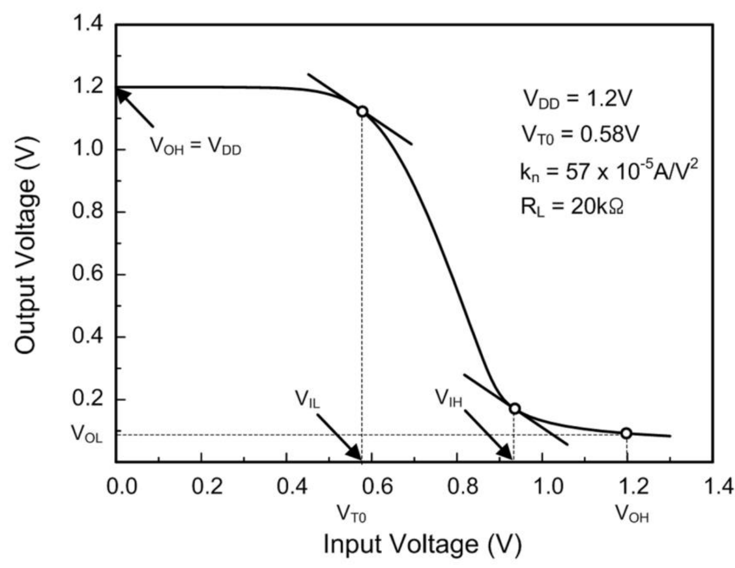
- Resistive-load inverter circuit & its VTC
- $V_{in}<V_{T0}$, nMOS off
  - I_D= ______
  - V_{out} = ______
- $V_{in}>V_{T0}$, nMOS in sat.
  - $V_{ds} =$ _____ $> V_{in} - V_{T0}$
  - $I_R =$
- $V_{in}>V_{T0}$ & $V_{out}<V_{DSAT}$, nMOS in lin.
  - $I_R =$



Calculation of V_{OH}

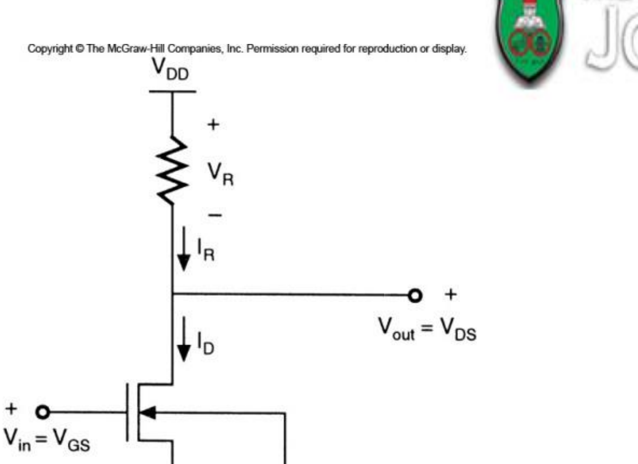
- $V_{out} = V_{dd} – R_L.I_R$
- When $V_{in} < V_{T0}$
  - Transistor is in _____
  - $I_R = I_D = 0$
  - V_{OH} = _____

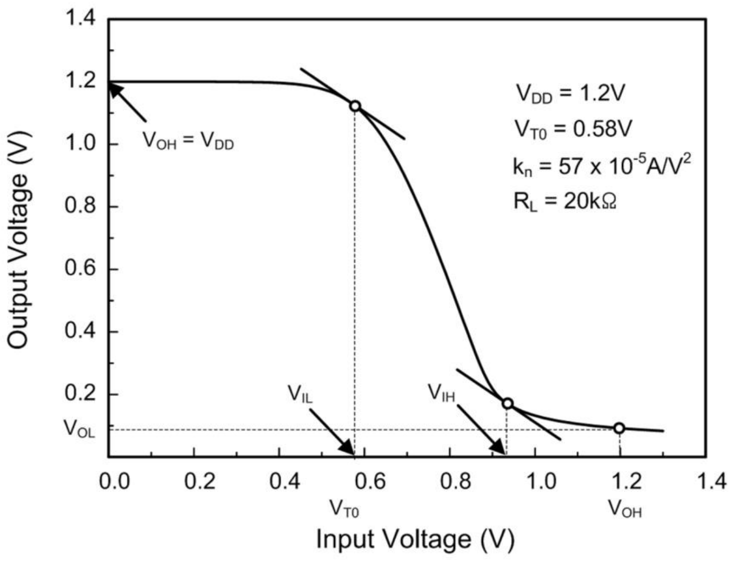




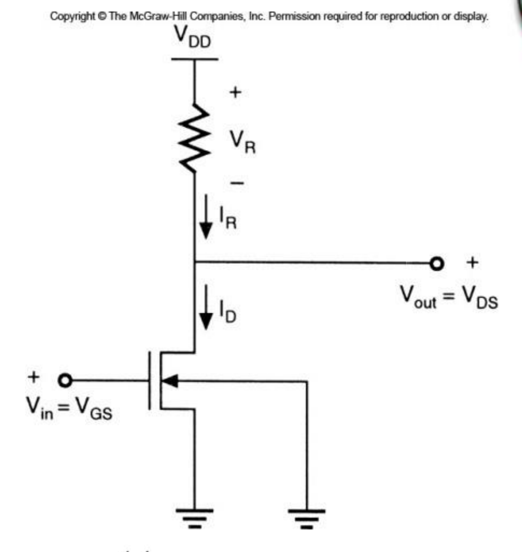
Calculation of V_{OL}

- V_{OL} is obtained when $V_{in} = V_{OH} =$ _____
  - V_{in}
  - NMOS is in _________
  - Using KCL $I_R = I_D$



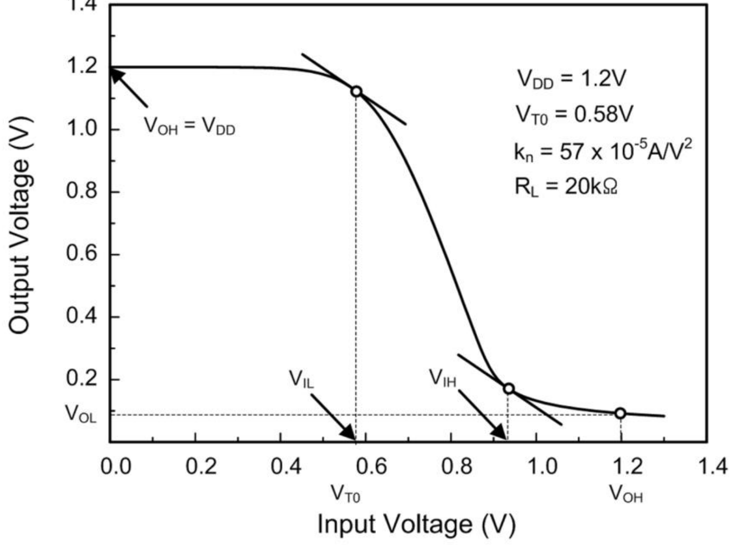


Calculation of V_{OL}



- V_{OL} is obtained when $V_{in} = V_{OH} =$ _____
  - V_{in}
  - NMOS is in _________
  - Using KCL $I_R = I_D$
  - $(Vdd - VOL) / \mathrm{RL} = \mathrm{B}/2 \,.\, (2.(\mathrm{V_{dd}} - \mathrm{VT_0}).\mathrm{VOL} - \mathrm{V^2}$

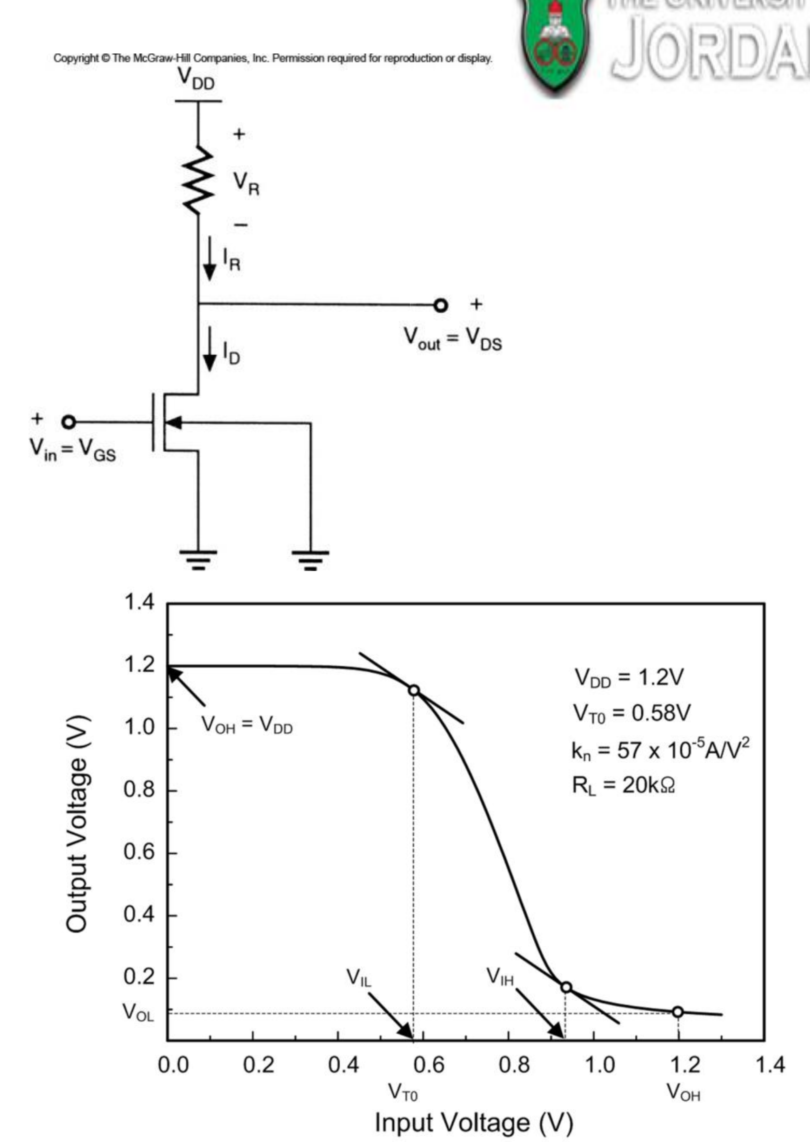
$$V_{OL} = V_{DD} - V_{T0} + \frac{1}{k_n R_L} - \sqrt{(V_{DD} - V_{T0} + \frac{1}{k_n R_L})^2 - \frac{2V_{DD}}{k_n R_L}}$$



  - Use the KCL and V_{OL} definition to to find the value of V_{OL}

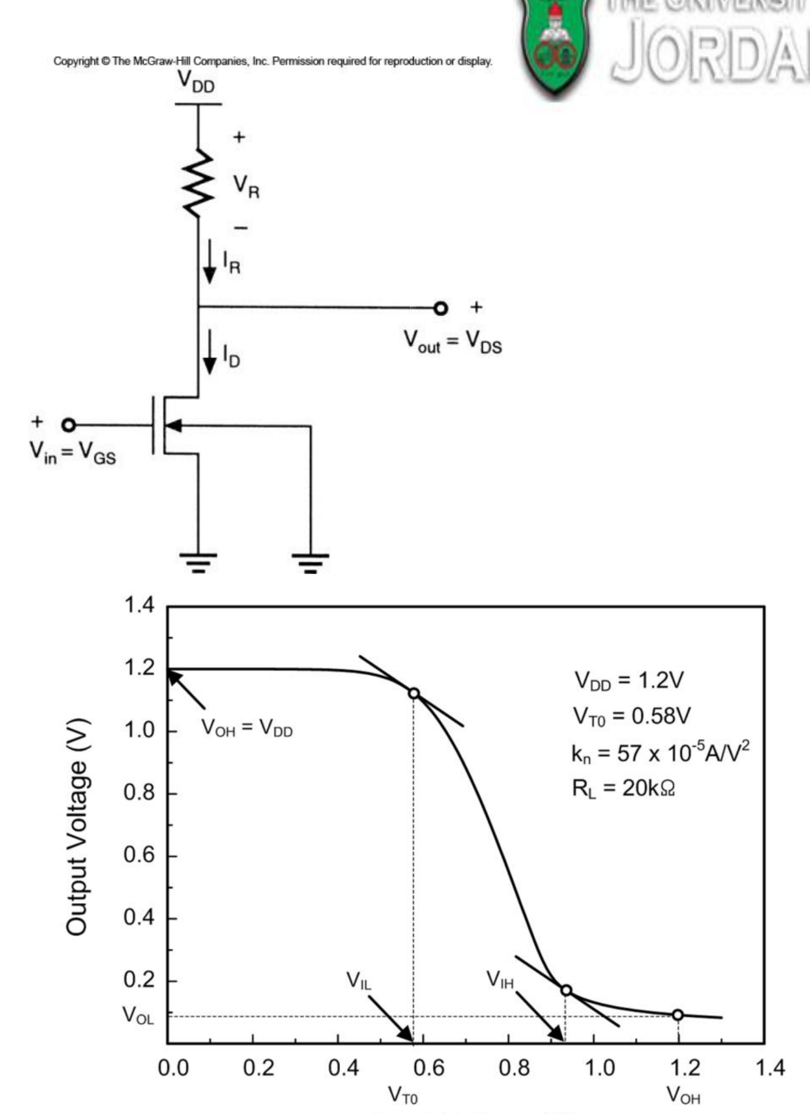
Calculation of V_{IL}

- At $V_{in}=V_{IL}$ $dV_{out}/dV_{in} = -1$
- NMOS is in __________
  - Using KCL $I_R = I_D$



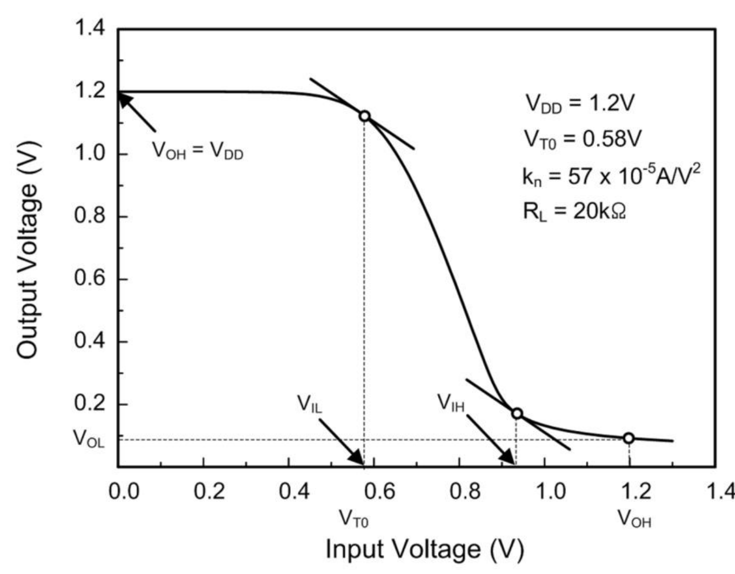
Calculation of V_{IL}

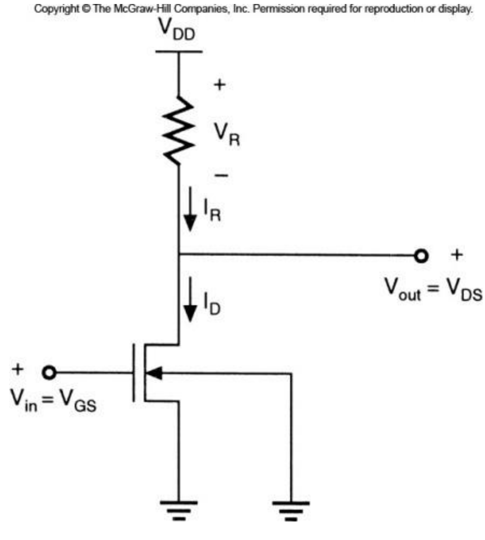
- At $V_{in}=V_{IL}$ $dV_{out}/dV_{in} = -1$
- NMOS is in __________
  - Using KCL $I_R = I_D$
  - $(Vdd - Vout)/ \mathrm{RL} = \mathrm{B}/2\ (\mathrm{Vin} - \mathrm{VT}_0)^2$ ....(eq1)
  - Differentiate both side
    - $-1/R_L \cdot (dV_{out}/dV_{in}) = B(V_{IL} - V_{T0})$
  - substitute $dV_{out}/dV_{in} = -1$
    - $-1/R_L \cdot (-1) = B(V_{IL} - V_{T0})$
    - $V_{IL} = V_{T0} + 1/(B.R_L)$ ....(eq2)



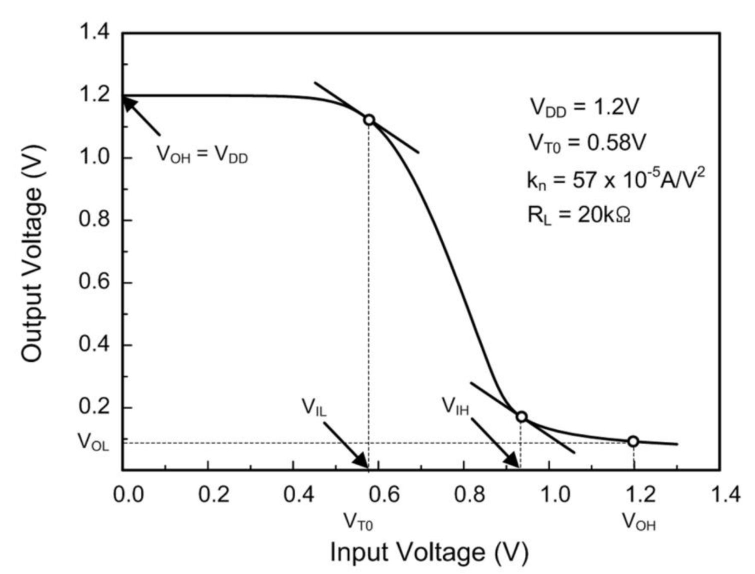
Calculation of V_{IH}

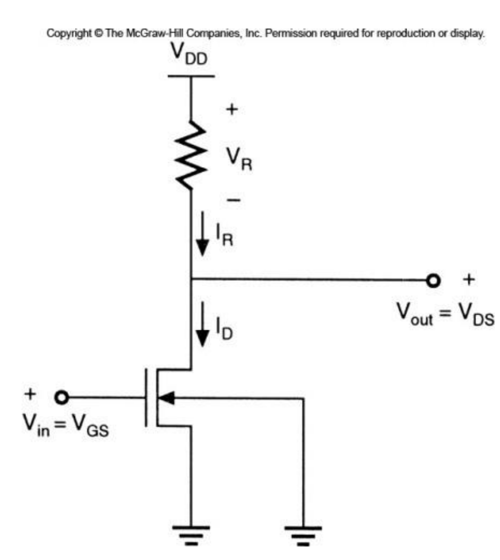
- At $V_{in}=V_{IH}$ $dV_{out}/dV_{in} = -1$
- V_{in} is slightly larger than V_{OL}
- NMOS is in ________





Calculation of V_{IH}





- At $V_{in}=V_{IH}$ $dV_{out}/dV_{in} = -1$
- V_{in} is slightly larger than V_{OL}
- NMOS is in __________
  - Using KCL $I_R = I_D$
  - $(Vdd - Vout)/\ \mathrm{RL} = \mathrm{B}/2\ .\ (2.\ (\mathrm{V_{in}} - \mathrm{VT_0}).\ \mathrm{VOUT} - \mathrm{V^2}_{\mathrm{OUT}})$ ….eq(1)
  - Differentiate both side
    - $-1/R_L\ .\ (dV_{out}/dV_{in}) = \mathrm{B}/2\ .\ (2.\ (\mathrm{V_{in}} - \mathrm{VT_0}).\ \mathrm{dV_{OUT}/dVin} - 2\mathrm{V_{OUT}}\ .\ \mathrm{dV_{OUT}/dVin}$ ….eq(2)

  - Solve eq1 and eq2 and substitute $dV_{out}/dV_{in} = -1$ and V_{in} with V_{IH}

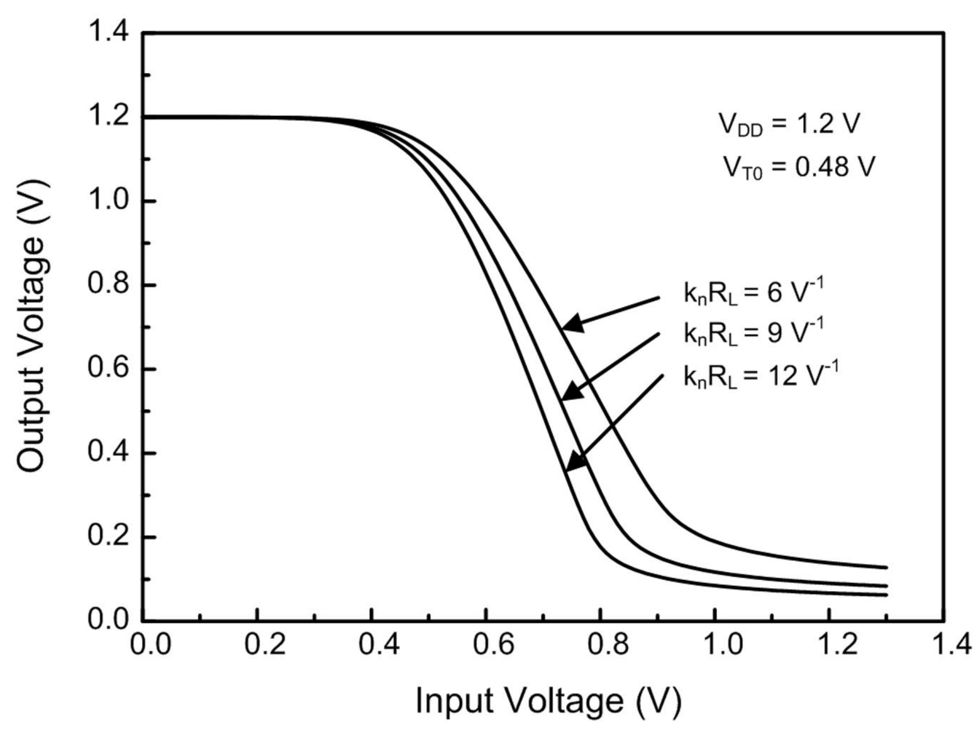
$$V_{IH} = V_{T0} + \sqrt{\frac{8}{3} \cdot \frac{V_{DD}}{k_n R_L}} - \frac{1}{k_n R_L}$$

VTC, Power & Chip Area

- VTC of the resistive-load inverter for different $(k_n \cdot R_L)$

$$V_{IH} = V_{T0} + \sqrt{\frac{8}{3} \cdot \frac{V_{DD}}{k_n R_L}} - \frac{1}{k_n R_L}$$

$$V_{IL} = V_{T0} + \frac{1}{k_n R_L}$$



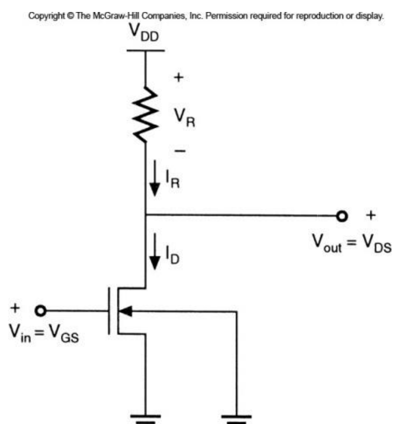
Power

- $V_{in} = V_{OH}$
  - NMOS is in Linear
  - $I_D = I_R = (V_{dd} – V_{OL})/R_L$
- $V_{in} = V_{OL}$
  - NMOS is cutoff
  - ID = 0
- P_{DC}(Average) = $(V_{dd}/2).(V_{dd} – V_{OL})/R_L$

Resistor

- Two possibilities for fabricating resistors:
  - Diffused resistor: an isolated n or p-type diffusion region.
    - 20-100 ohm/square: needs large area not practical for VLSI
  - Metal Resistor
  - Undoped polysilicon resistor:
    - mask off during poly doping to create about 10M ohm/square.

Example 5.1 - Inverter Design

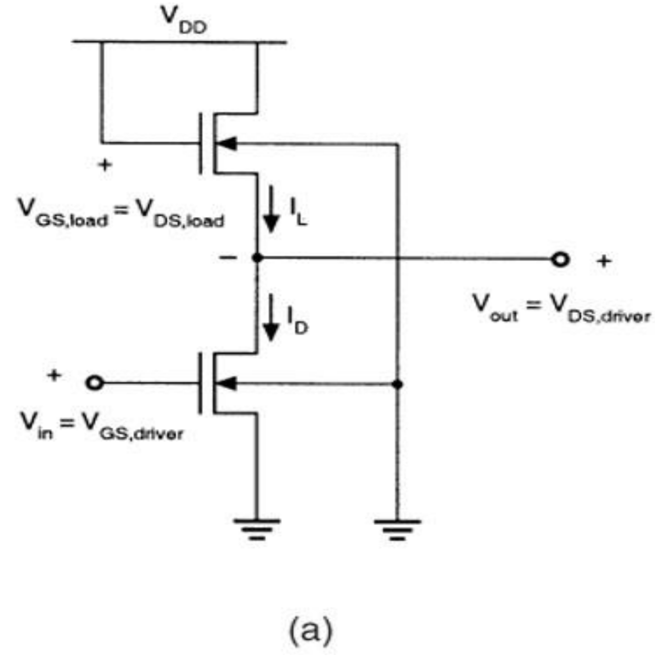


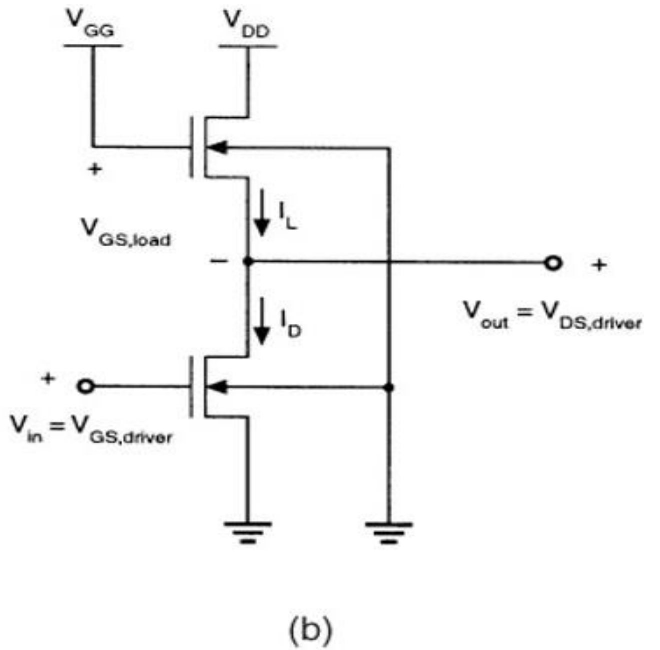
- Resistive-load inverter circuit
  - V_{DD}=5V, $\mu*C_{ox}$=30μA/V^2, V_{T0}=1V
  - V_{OL}=200mV
- Determine (W/L) ratio of the driver Tr. And R_L to obtain the required V_{OL}
  - V_{out} = ______, V_{in} =_______. NMOS is in ______

(W/L)-Ratio	Load resistor (R_L[kΩ])	DC power consumption ($P_{DC,avg.}$[μW])
1		
2		
3		
4		
5		
6		

Inverters with MOSFET Load

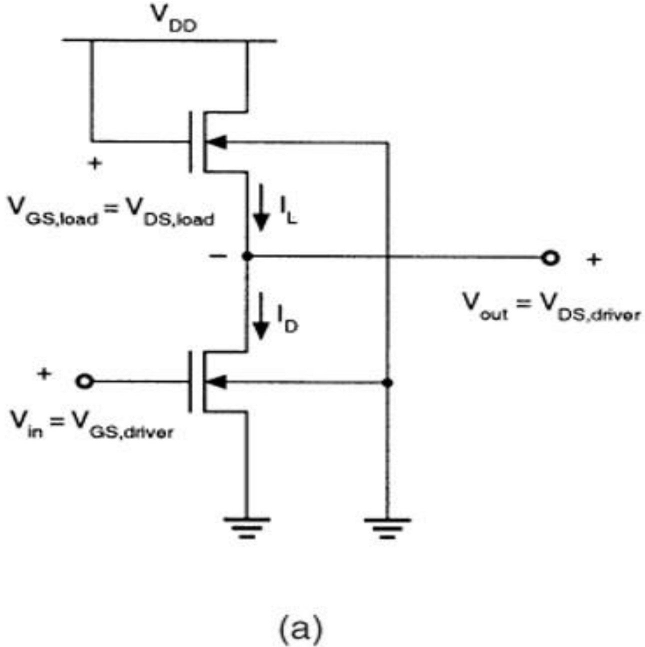
- Enhancement-Load Inverter





Inverters with MOSFET Load

- Enhancement-Load Inverter



(a)

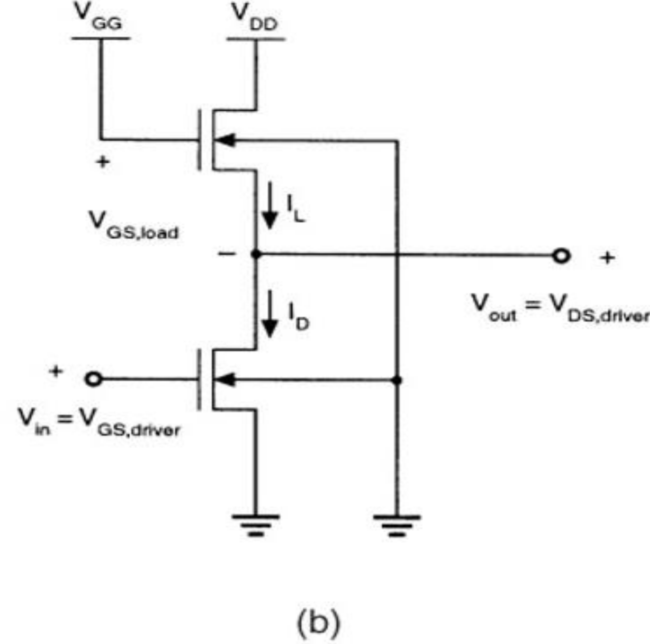
Load NMOS operates in ______

V_{OH} = ________

Power

Inverters with MOSFET Load

- Enhancement-Load Inverter



(b)

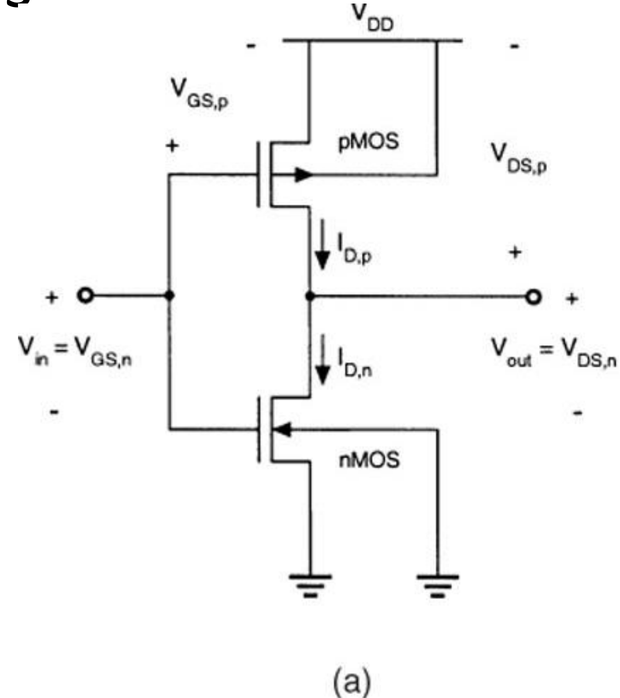
Load NMOS operates in _______

V_{OH} = _________

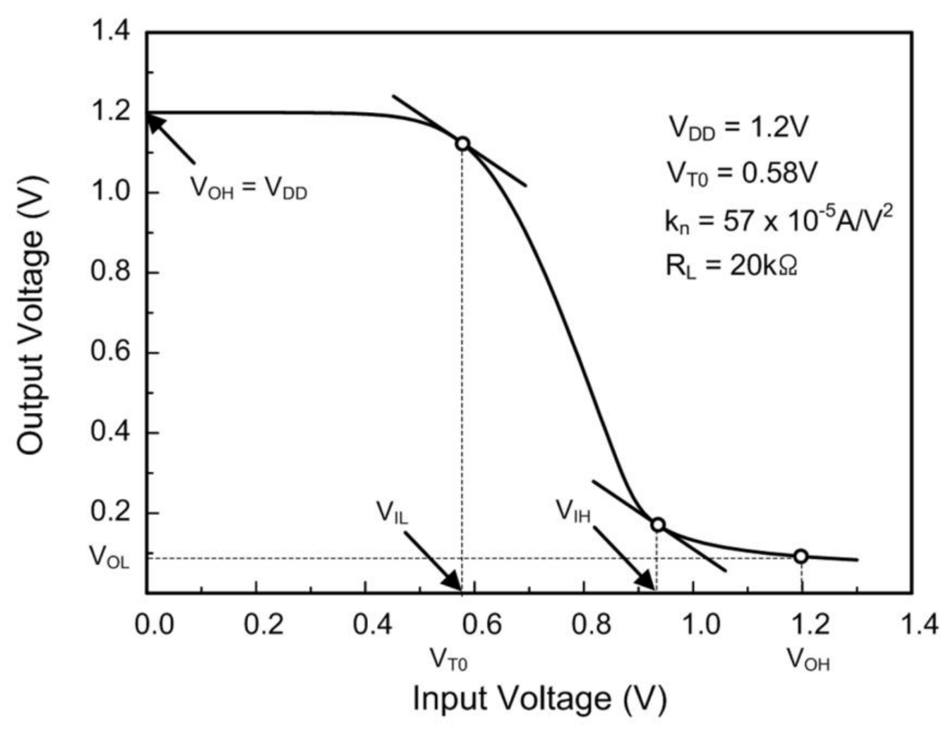
Power

CMOS Inverter

- Has NMOS and PMOS Transistors
- Cons
  - Complexity
- Pros
  - Power
  - VTC is sharp

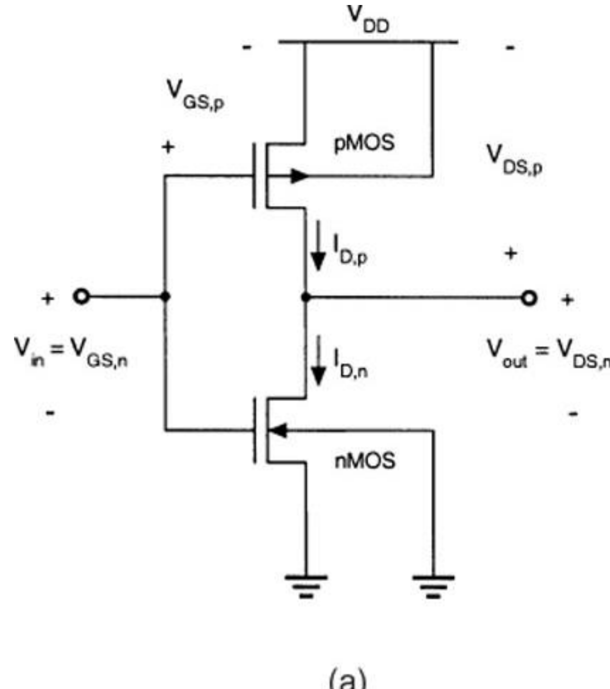


THE UNIVERSITY OF
JORDAN


Output Voltage (V)
Input Voltage (V)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
V_OL
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
V_T0
V_OH
V_OH = V_DD
V_IL
V_IH
V_DD = 1.2V
V_T0 = 0.58V
k_n = 57 x 10^-5 A/V^2
R_L = 20kΩ

Operation

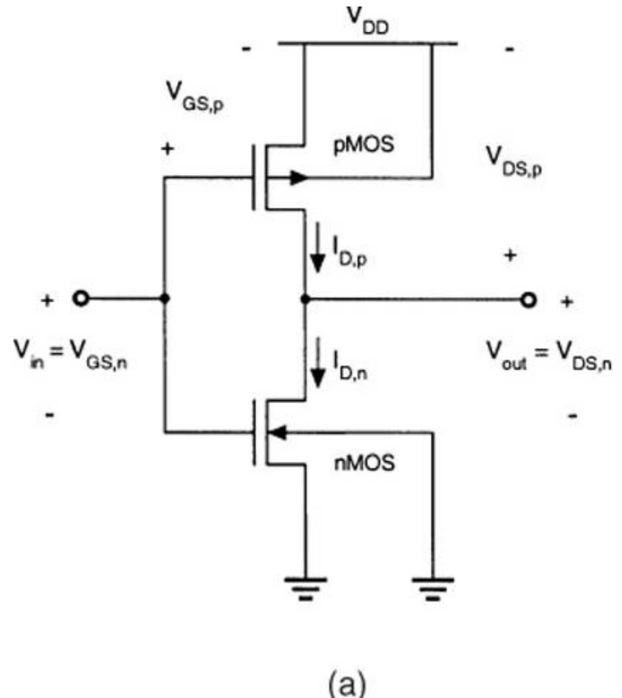
- $V_{GS,n}$ ____________________
- $V_{GS,p}$ ____________________
- $V_{DS,n}$ ____________________
- $V_{DS,p}$ ____________________



(a)

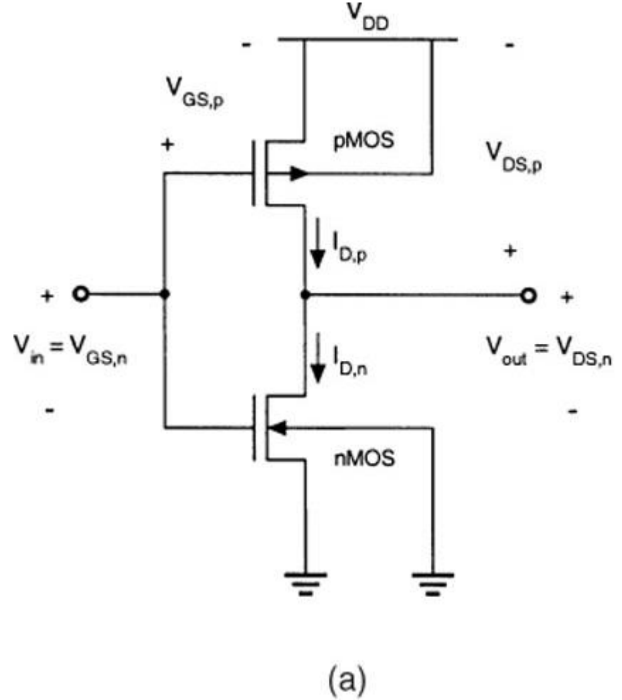
Operation

- $V_{GS,n} < V_{th,n}$
- NMOS is OFF
- PMOS operates in ________



Operation

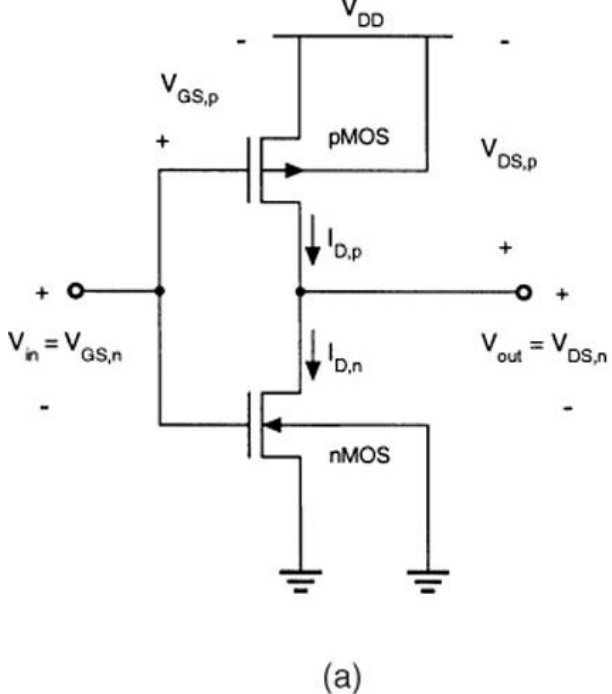
- $V_{GS,p} > V_{th,p}$
- PMOS is OFF
- NMOS operates in ________



(a)

Operation

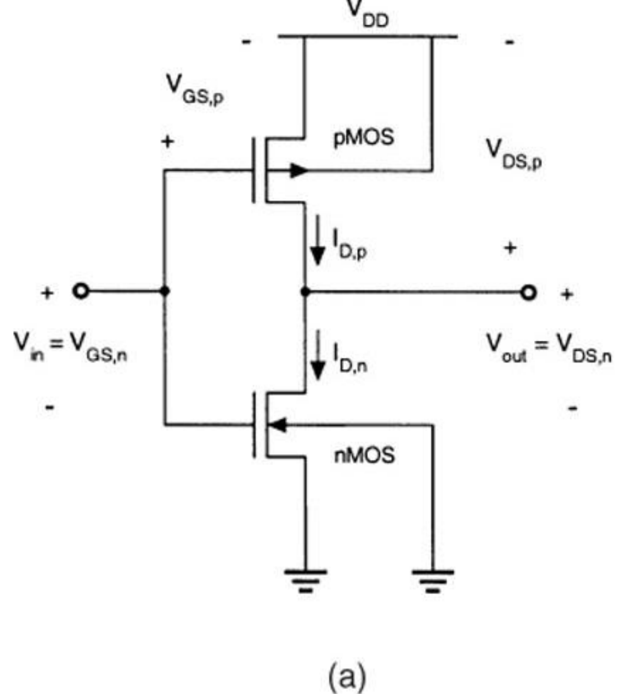
- $V_{in} = V_{IL}$
- NMOS is in _______
- PMOS is in _______



(a)

Operation

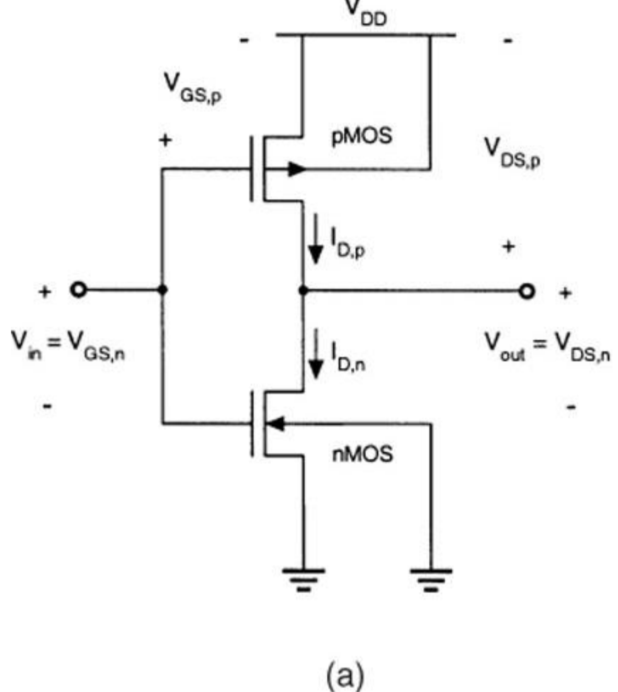
- $V_{in} = V_M = Vdd/2$
- NMOS is in _______
- PMOS is in _______



(a)

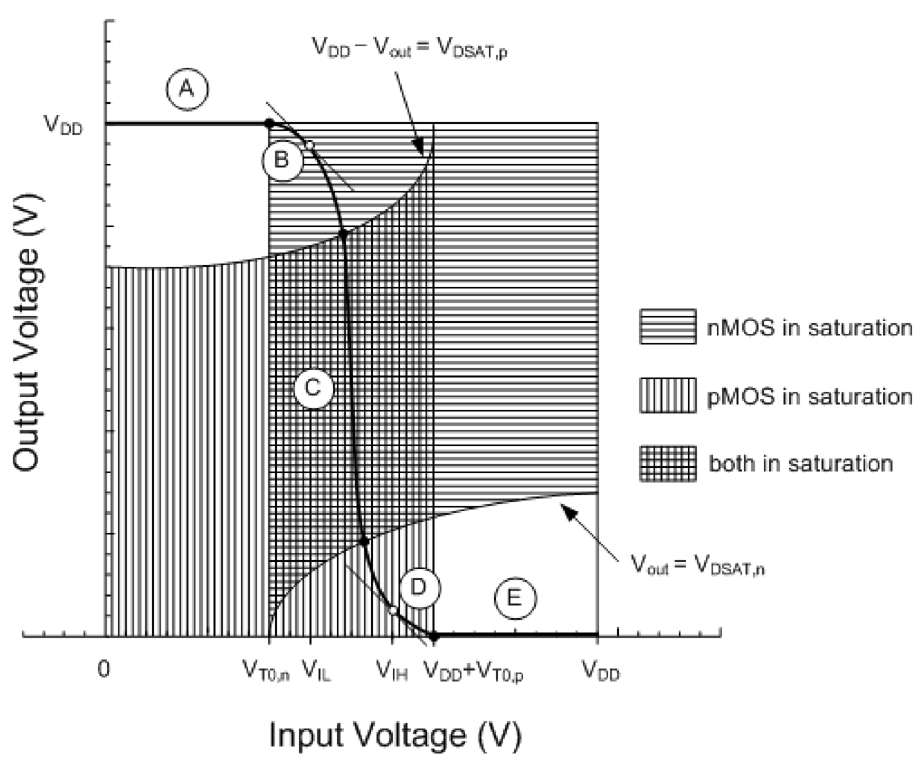
Operation

- $V_{in} = V_{IH}$
- NMOS is in _______
- PMOS is in _______



(a)

THE UNIVERSITY OF
JORDAN


$V_{DD} - V_{out} = V_{DSAT,p}$
A
B
C
D
E
V_{DD}
Output Voltage (V)
nMOS in saturation
pMOS in saturation
both in saturation
$V_{out} = V_{DSAT,n}$
0
$V_{T0,n}$
V_{IL}
V_{IH}
$V_{DD}+V_{T0,p}$
V_{DD}
Input Voltage (V)

Region	V_{in} in the region	V_{out}	NMOS	PMOS
A	$< V_{th,n}$	V_{OH}	Cut-off	Linear
B	V_{IL}	$\sim V_{OH}$	Sat.	Linear
C	V_M	V_M	Sat.	Sat.
D	V_{IH}	$\sim V_{OL}$	Linear	Sat.
E	$>V_{tdd} + V_{th,p}$	V_{OL}	Linear	Cut-off

V_{OH} & V_{OL}

- V_{OH} = __________

- V_{OL} = ___________

V_{IL}

- $V_{in} = V_{IL} \rightarrow dV_{out}/d\,V_{in} = -1$
- NMOS is in _________
- PMOS is in _________

V_{IL}

- $V_{in} = V_{IL} \rightarrow dV_{out}/d V_{in} = -1$
- NMOS is in _________
- PMOS is in _________

- $B_n/2(V_{GS,n} – V_{th,n})^2 = B_p/2 (2.(V_{GS,p} – V_{th,p}). V_{DS,p} - V_{DS,p}^2)$

$$V_{IL} = \frac{2V_{out} + V_{T0,p} - V_{DD} + k_R V_{T0,n}}{1 + k_R}$$

- $K_R = K_n/K_p$

V_{IH}

- $V_{in} = V_{IH} \rightarrow dV_{out}/d\,V_{in} = -1$
- NMOS is in _________
- PMOS is in _________

V_{IH}

- $V_{in} = V_{IH} \rightarrow dV_{out}/d V_{in} = -1$
- NMOS is in _________
- PMOS is in _________

- $B_p/2(V_{GS,p} - V_{th,p})^2 = B_n/2\,(2.(V_{GS,n} - V_{th,n}).\, V_{DS,n} - V_{DS,n}^2)$

$$V_{IH} = \frac{V_{DD} + V_{T0,p} + k_R \cdot \left(2V_{out} + V_{T0,n}\right)}{1 + k_R}$$

- $K_R = K_n/K_p$

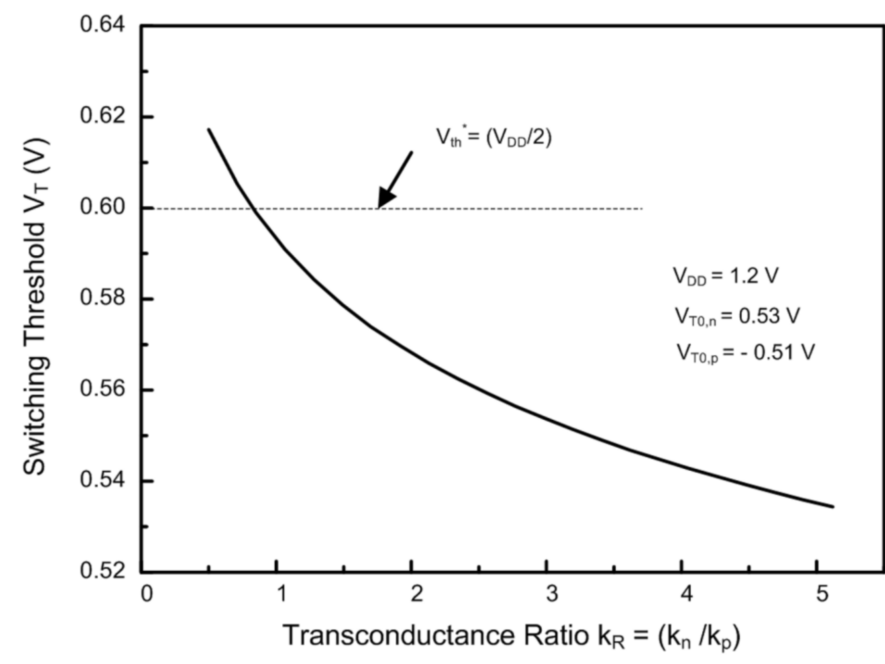
Inverter threshold voltage

- $V_{in} = V_{out}$
- NMOS and PMOS are in Saturation

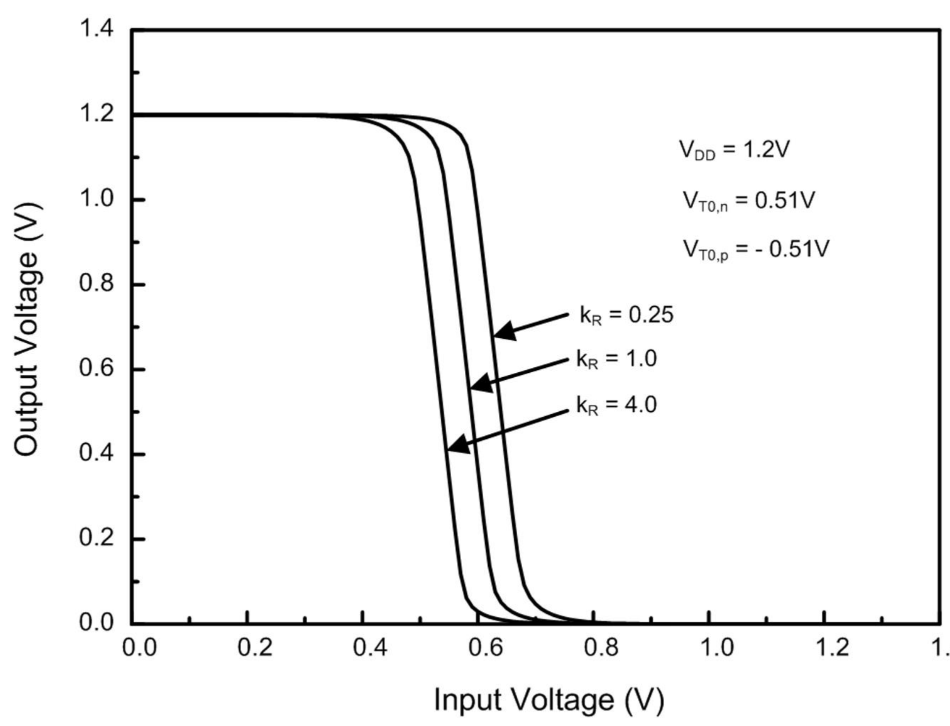
Inverter threshold voltage

- $V_{in} = V_{out}$
- NMOS and PMOS are in Saturation

$$V_{th} = \frac{V_{T0,n} + \sqrt{\frac{1}{k_R}} \cdot \left(V_{DD} - \left|V_{T0,p}\right|\right)}{\left(1 + \sqrt{\frac{1}{k_R}}\right)}$$



Design of CMOS Inverter



Design of CMOS Inverters (2)

- If symmetric CMOS inverter with $V_{T0,n} = |V_{T0,p}|$ and $k_R = 1$,

$$V_{IL} = \frac{1}{8} \cdot (3V_{DD} + 2V_{T0,n})$$

$$V_{IH} = \frac{1}{8} \cdot (5V_{DD} - 2V_{T0,n})$$

→ $V_{IL} + V_{IH} = V_{DD}$

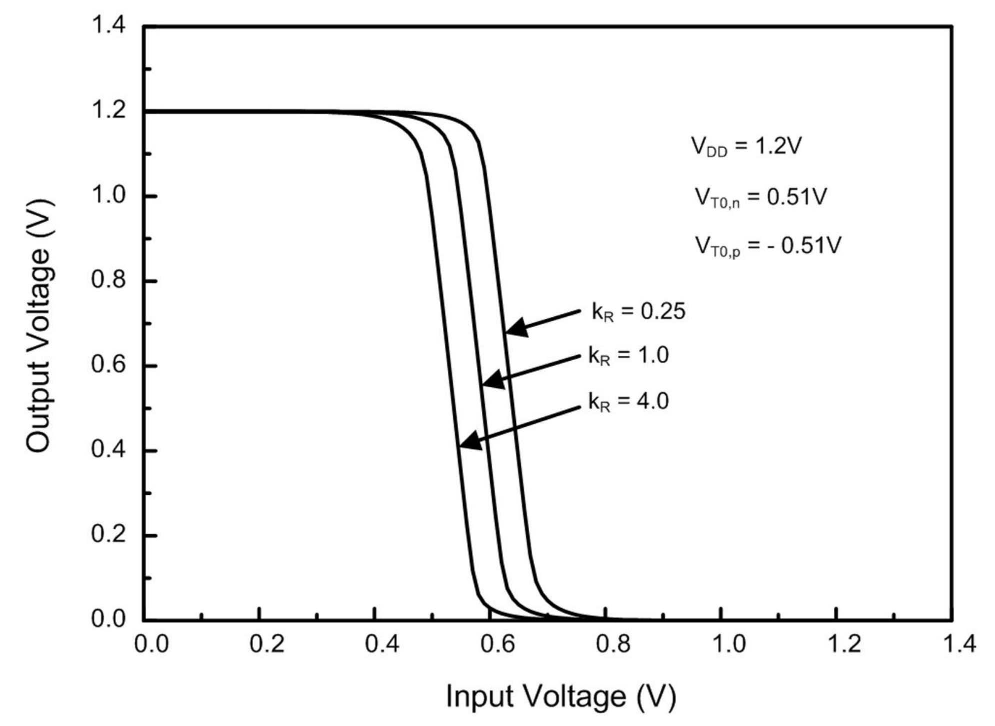
in a symmetric inverter

$$NM_L = V_{IL} - V_{OL} = V_{IL}$$

→

$$NM_H = V_{OH} - V_{IH} = V_{DD} - V_{IH}$$

→ $NM_L = NM_H = V_{IL}$



THE UNIVERSITY OF
JORDAN

Inverter Transient Characteristics

Dr.Mohammad Abdel-Majeed

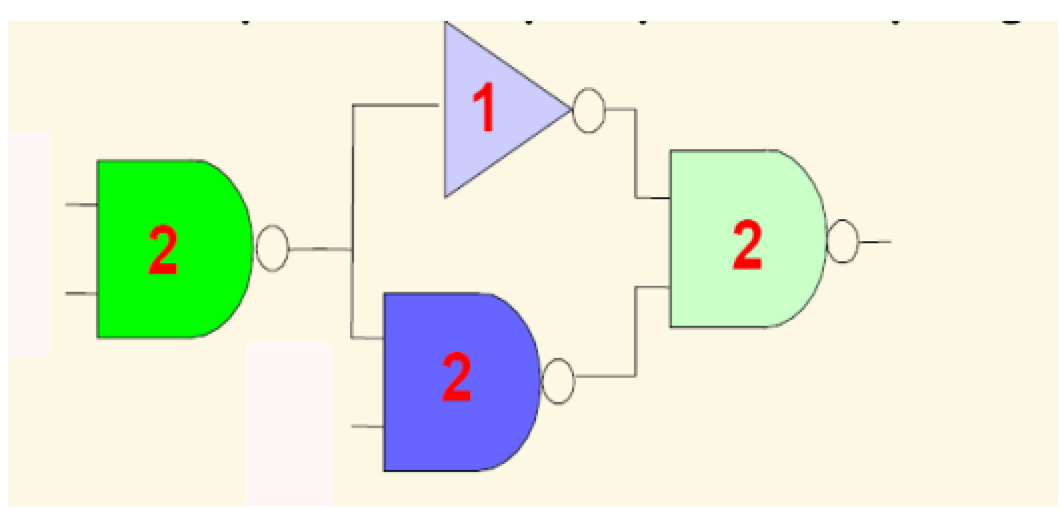
Assistant Professor

University of Jordan

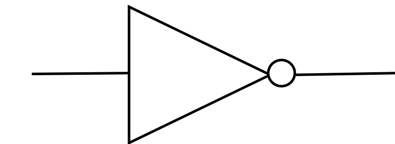
Timing Analysis

- Delay models are required for _________ and the ________

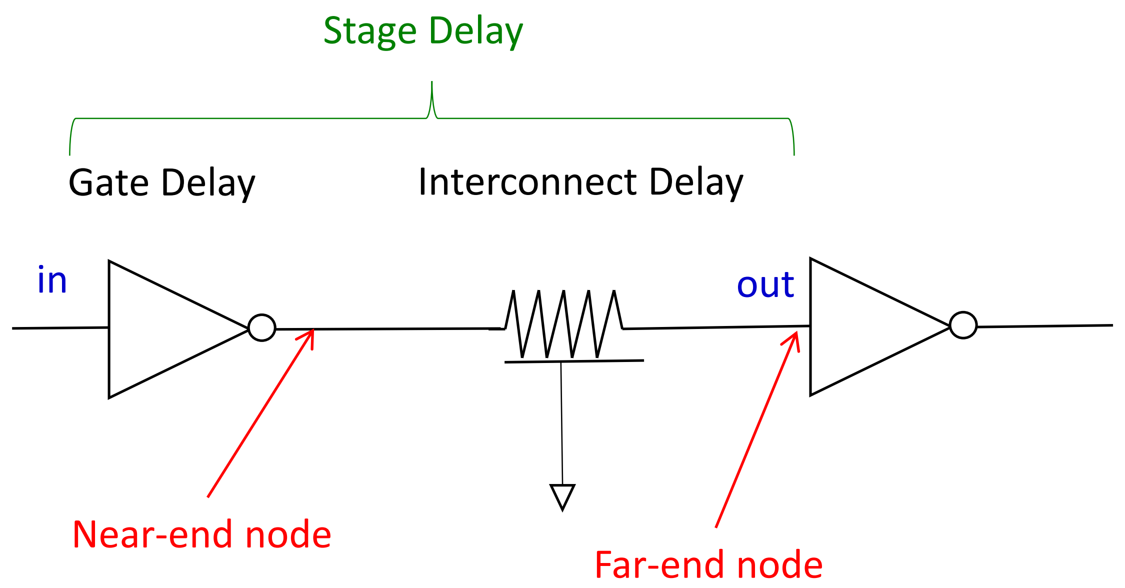
Example



THE UNIVERSITY OF
JORDAN



Inverter delay



Stage Delay

Gate Delay

Interconnect Delay

in

out

Near-end node

Far-end node

V_{DD}

pMOS

$i_{D,p}$

V_{in}

V_{out}

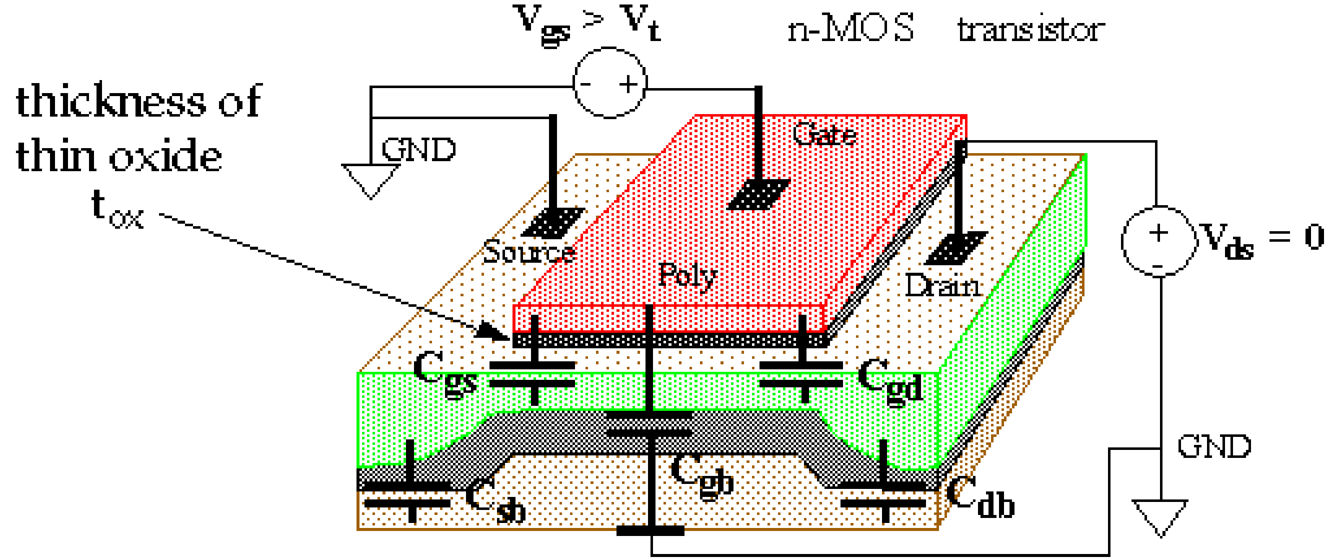
$i_{D,n}$

i_C

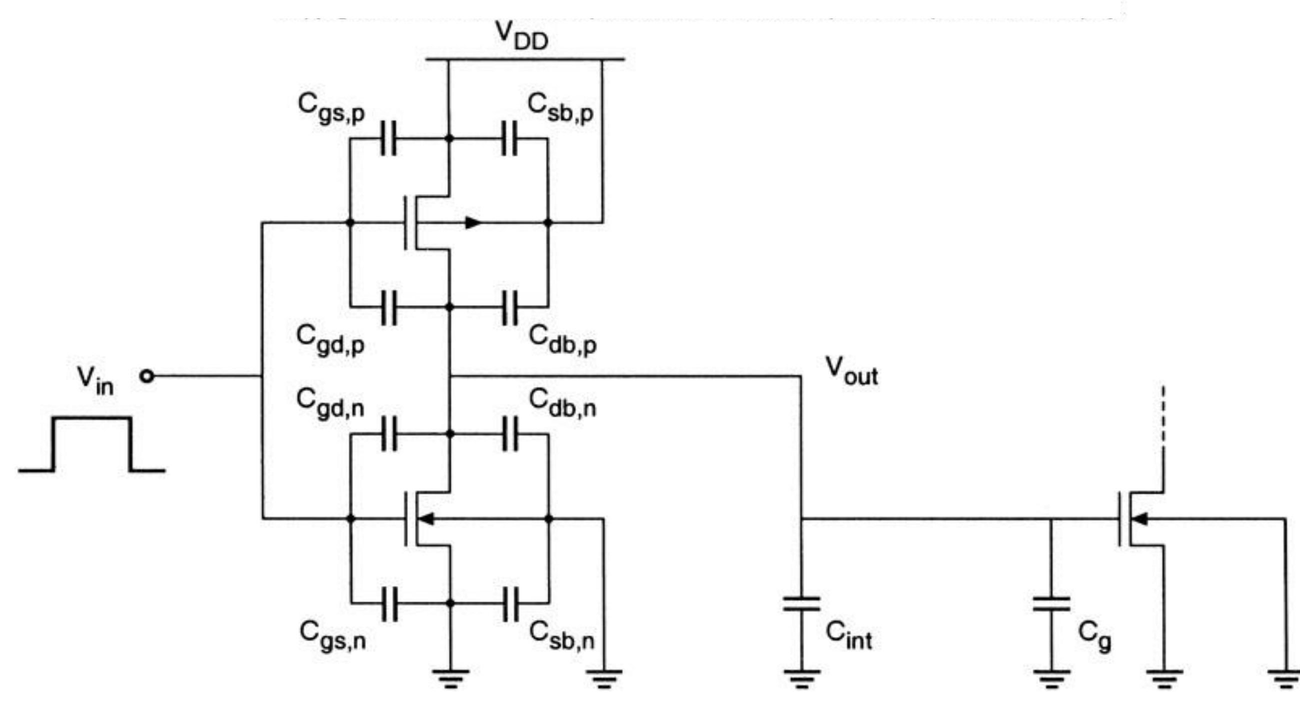
C_{load}

nMOS

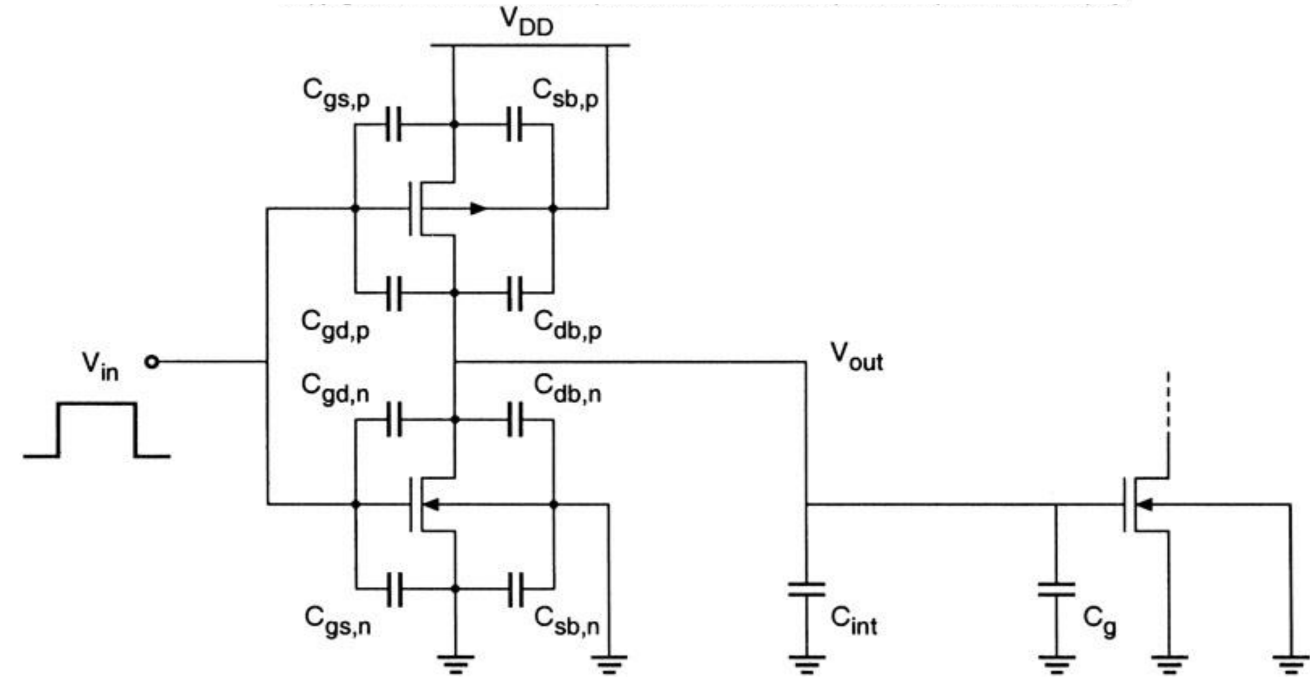
THE UNIVERSITY OF
JORDAN


Vgs > Vt
n-MOS transistor
thickness of
thin oxide
tox
GND
Gate
Source
Poly
Drain
Vds = 0
Cgs
Cgd
Cgb
Csb
Cdb
GND

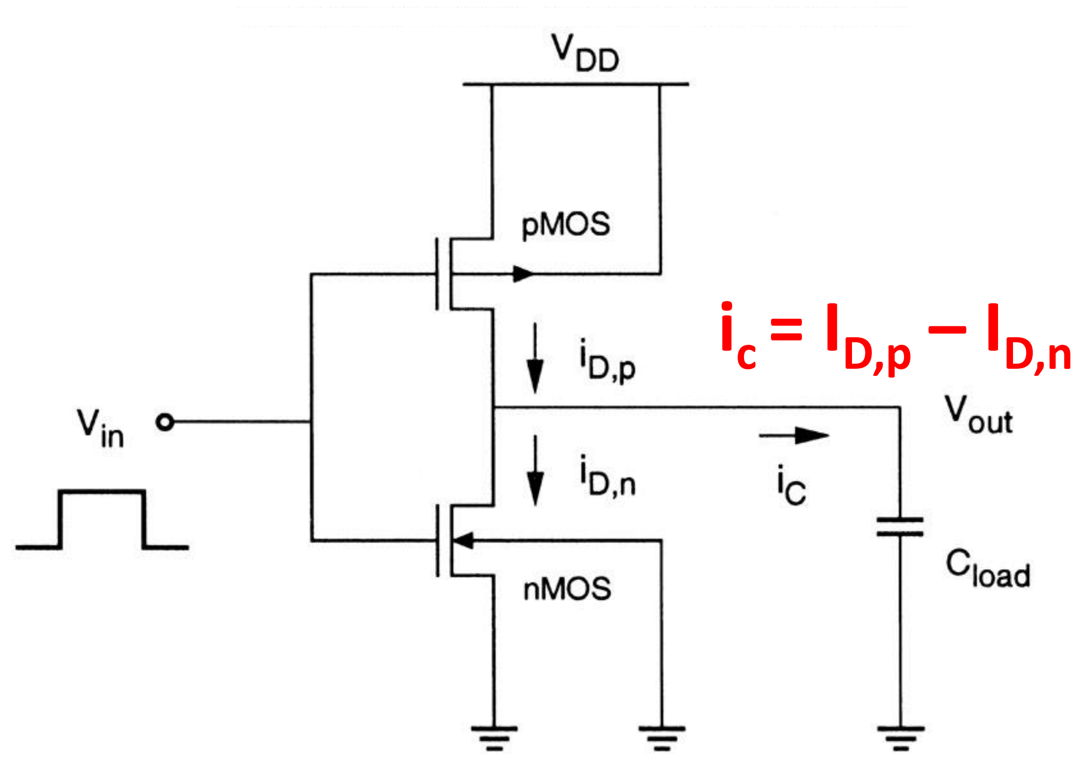
- Cload = _____



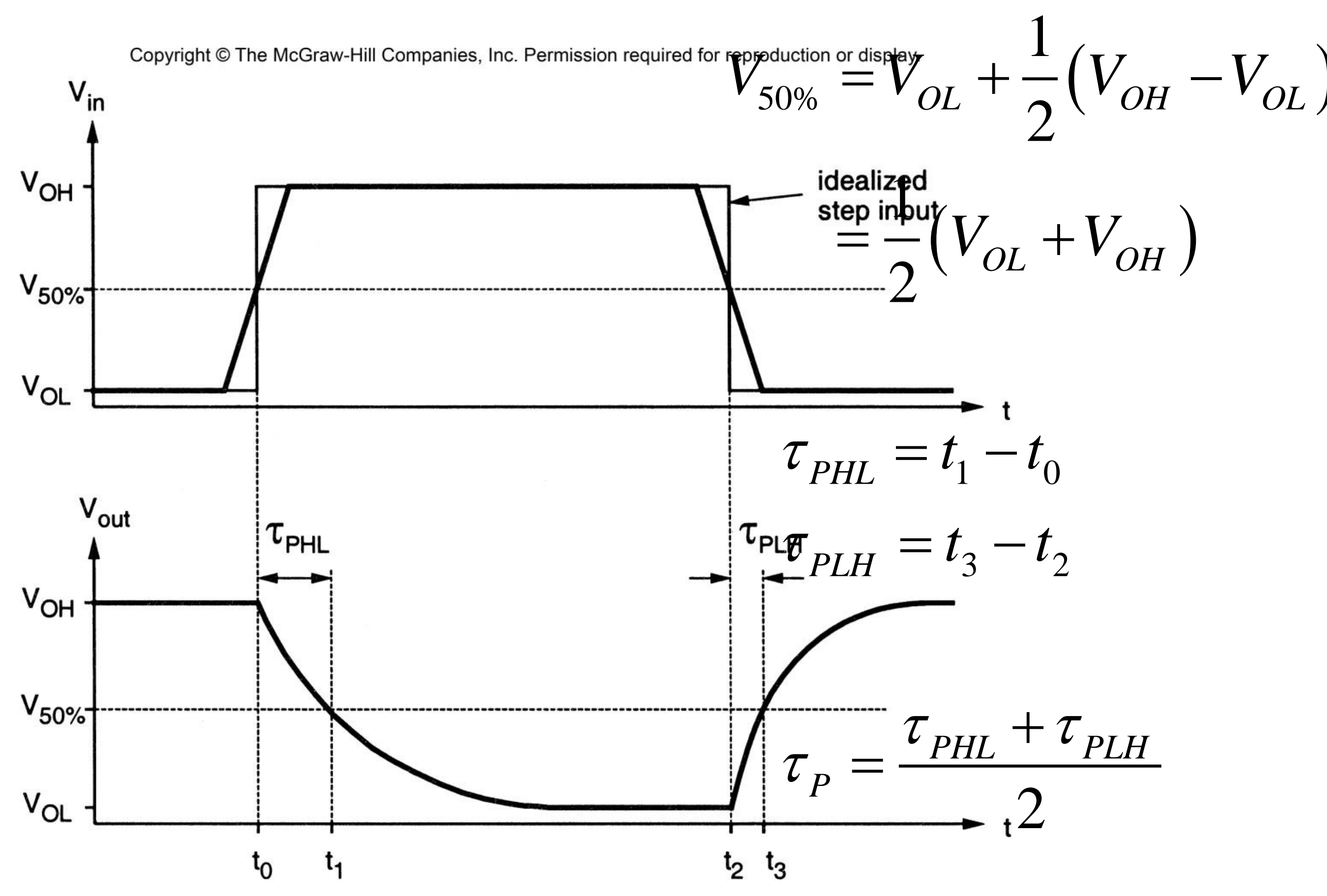
- Cload = $C_{load} = C_{gd,n} + C_{gd,p} + C_{db,n} + C_{db,p} + C_{int} + C_g$



THE UNIVERSITY OF
JORDAN


VDD
pMOS
iD,p
Vin
Vout
iD,n
iC
Cload
nMOS
ic = ID,p – ID,n

Delay-Time Definitions: Propagation Delays



$$V_{50\%} = V_{OL} + \frac{1}{2}(V_{OH} - V_{OL})$$

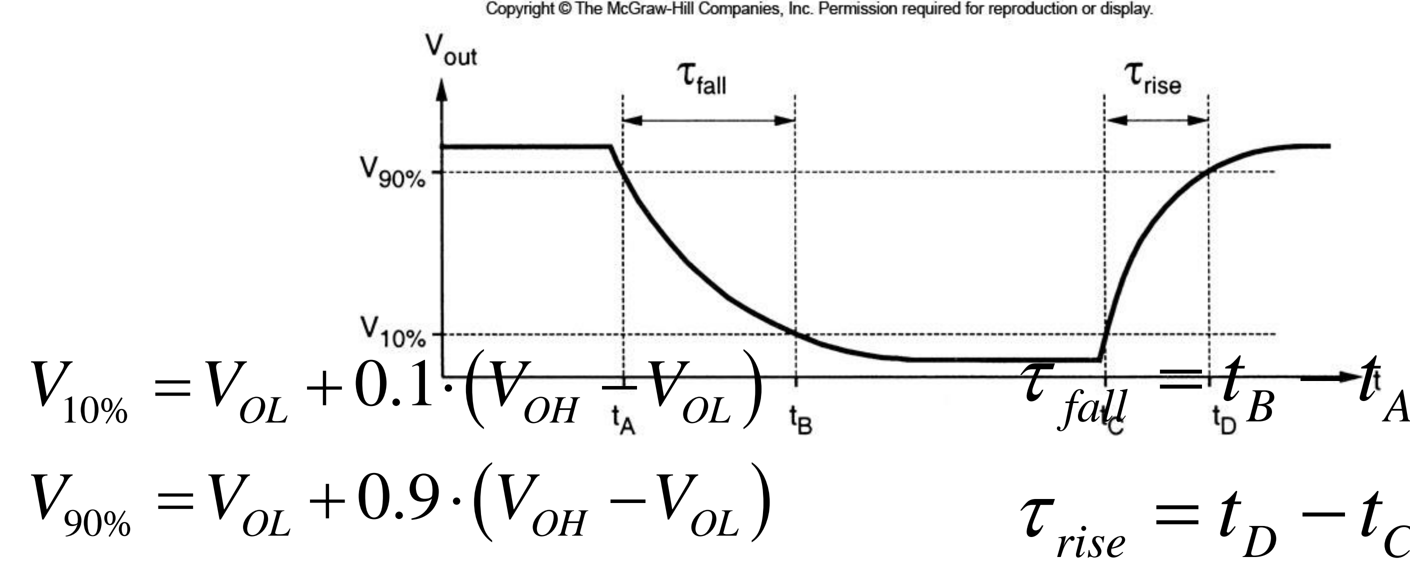
$$= \frac{1}{2}(V_{OL} + V_{OH})$$

$$\tau_{PHL} = t_1 - t_0$$

$$\tau_{PLH} = t_3 - t_2$$

$$\tau_P = \frac{\tau_{PHL} + \tau_{PLH}}{2}$$

Delay-Time Definitions: Rise & Fall Times



$$V_{10\%} = V_{OL} + 0.1 \cdot (V_{OH} - V_{OL})$$

$$V_{90\%} = V_{OL} + 0.9 \cdot (V_{OH} - V_{OL})$$

$$\tau_{fall} = t_B - t_A$$

$$\tau_{rise} = t_D - t_C$$

Calculation of Propagation Delay (AVG current method)

- Simplest method : estimating the average capacitance current

Calculation of Propagation Delay (AVG current method)

- Simplest method : estimating the average capacitance current

$$\tau_{PHL} = \frac{C_{load} \cdot \Delta V_{HL}}{I_{avg,HL}} = \frac{C_{load} \cdot (V_{OH} - V_{50\%})}{I_{avg,HL}}$$

$$\tau_{PLH} = \frac{C_{load} \cdot \Delta V_{LH}}{I_{avg,LH}} = \frac{C_{load} \cdot (V_{50\%} - V_{OL})}{I_{avg,LH}}$$

Calculation of Propagation Delay (AVG current method)

- Simplest method : estimating the average capacitance current

$$\tau_{PHL} = \frac{C_{load} \cdot \Delta V_{HL}}{I_{avg,HL}} = \frac{C_{load} \cdot (V_{OH} - V_{50\%})}{I_{avg,HL}}$$

$$\tau_{PLH} = \frac{C_{load} \cdot \Delta V_{LH}}{I_{avg,LH}} = \frac{C_{load} \cdot (V_{50\%} - V_{OL})}{I_{avg,LH}}$$

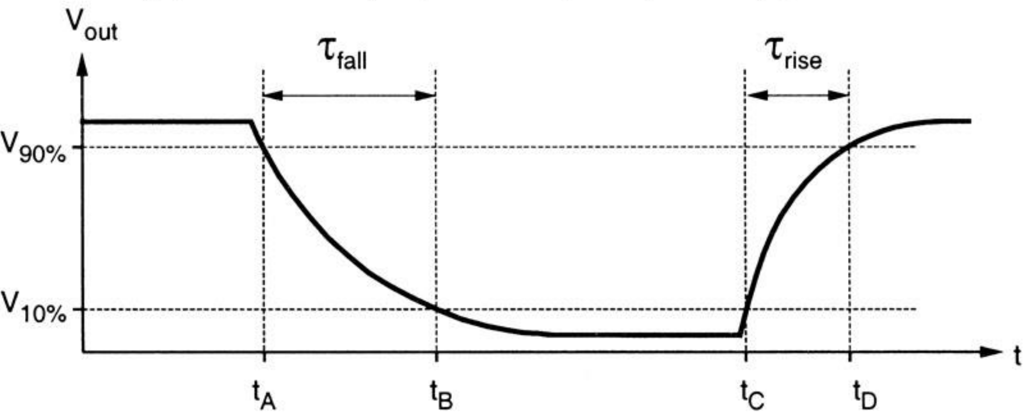
- The average current

$$I_{avg,HL} = \frac{1}{2}\left[i_C(V_{in} = V_{OH}, V_{out} = V_{OH}) + i_C(V_{in} = V_{OH}, V_{out} = V_{50\%})\right]$$

$$I_{avg,LH} = \frac{1}{2}\left[i_C(V_{in} = V_{OL}, V_{out} = V_{50\%}) + i_C(V_{in} = V_{OL}, V_{out} = V_{OL})\right]$$

Calculation of Rise and fall time (AVG current method)

- $T_{rise} = t_D - tC = \frac{C_{load}(V_{90\%} - V_{10\%})}{I_{avg,rise}}$
- $I_{avg,rise} = \frac{I_C + ID}{2}$
- $I_{avg,rise} = \frac{I_{sd,p}(Vin=VOL, Vout=V_{10\%}) + Isdp(Vin=VOL, Vout=V_{10\%})}{2}$

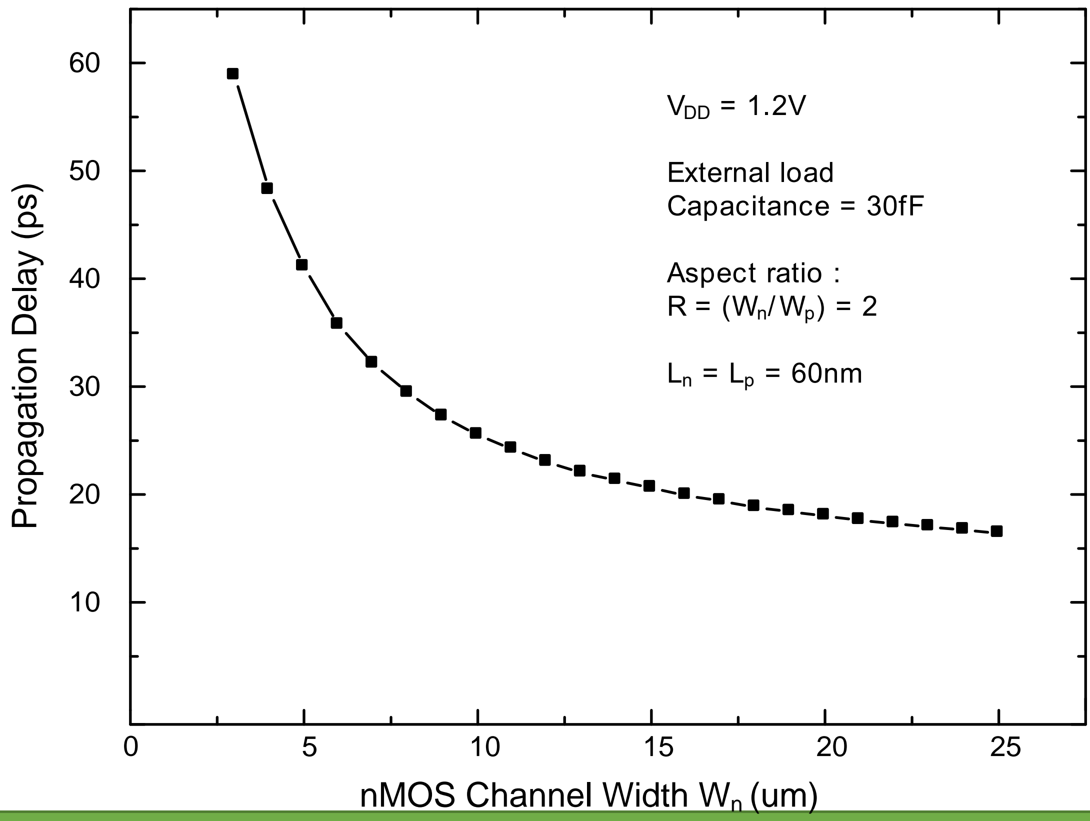


Example

- CMOS inverter with NMOS μC_{ox} = 20uA/V^2 , $(W/L)_n$ =10, $V_{t,n}$ =1V, C_{load} 1_{pF} and V_{DD}=5V
  - Calculate output fall time

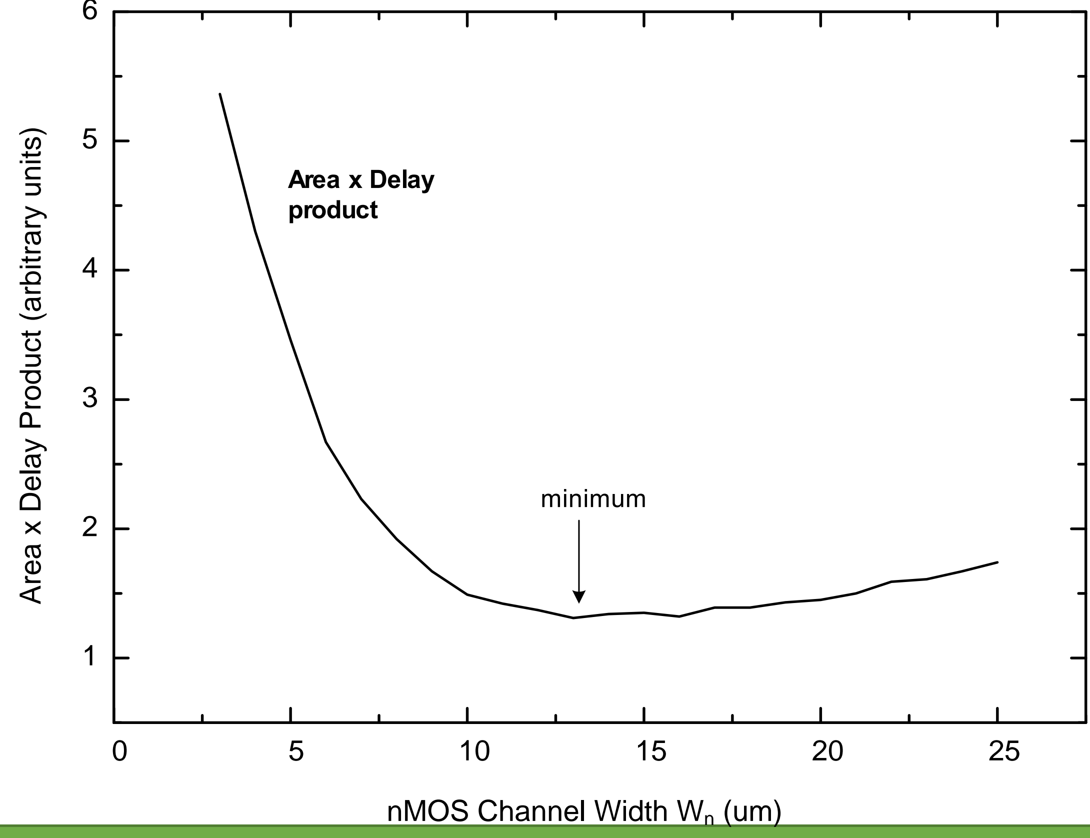
Example 6.4 (2)

- The falling-output propagation delay



Example 6.4 (3)

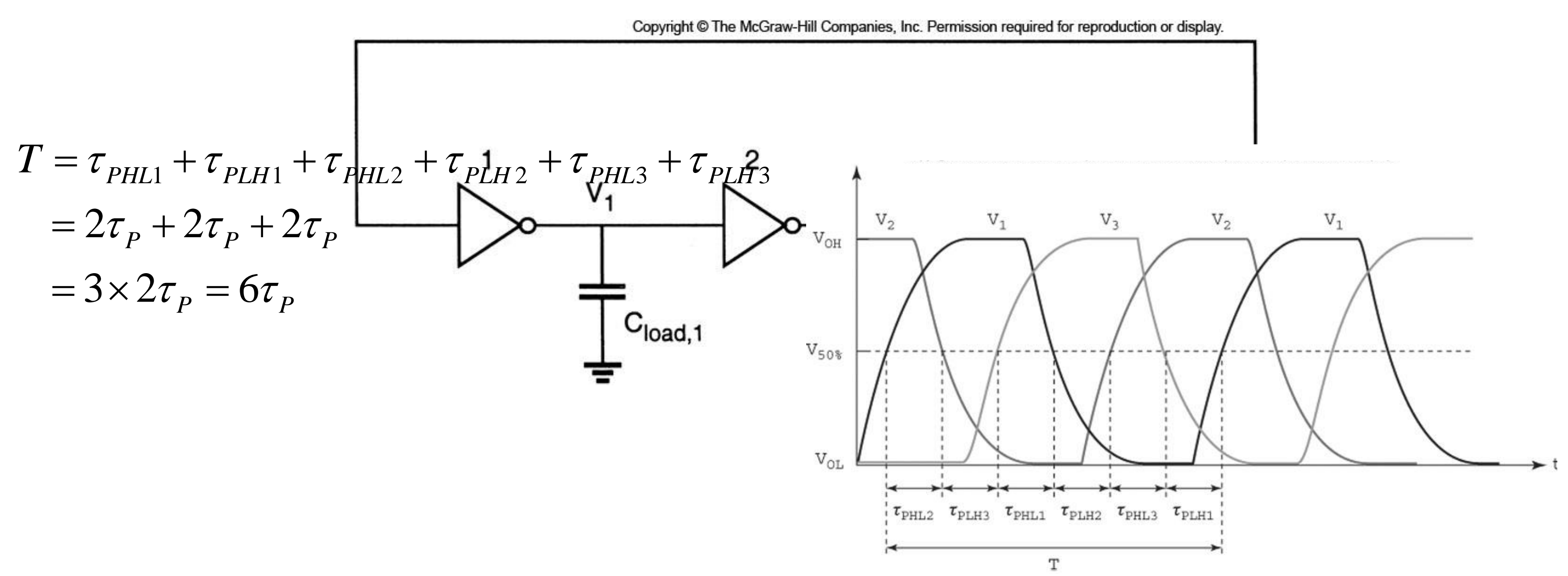
- The (Area x Delay) product



Ring oscillator

CMOS Ring Oscillator Circuit

$$T = \tau_{PHL1} + \tau_{PLH1} + \tau_{PHL2} + \tau_{PLH2} + \tau_{PHL3} + \tau_{PLH3}$$
$$= 2\tau_P + 2\tau_P + 2\tau_P$$
$$= 3 \times 2\tau_P = 6\tau_P$$



MOSFET Capacitor

- The on-chip capacitance found in MOS circuit are in general complicated functions of the layout geometries and the manufacturing processes.
- We will develop simple approximations for the on-chip MOSFET capacitances.

MOSFET Capacitor (2)

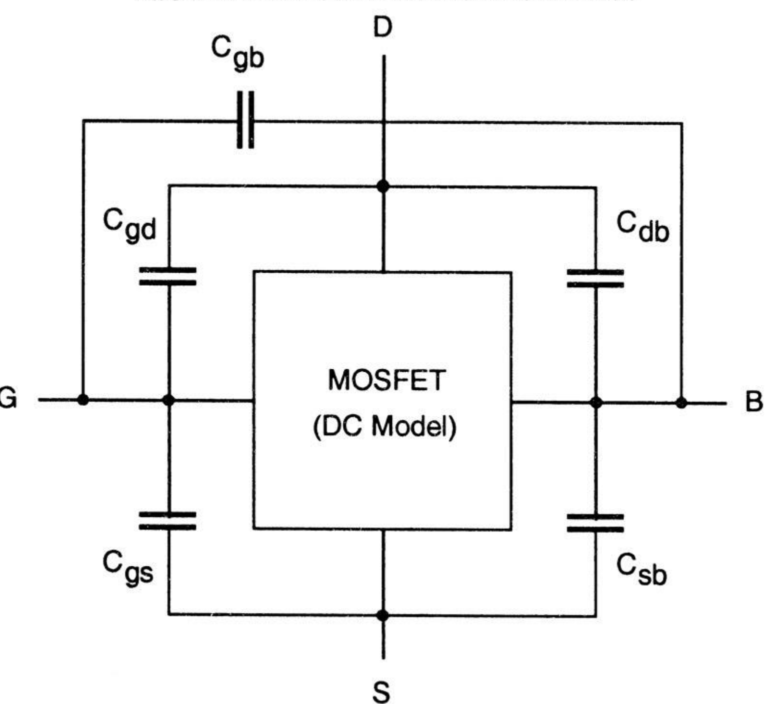
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

- The channel length is given by

$$L = L_M - 2 \cdot L_D \quad (3.117)$$

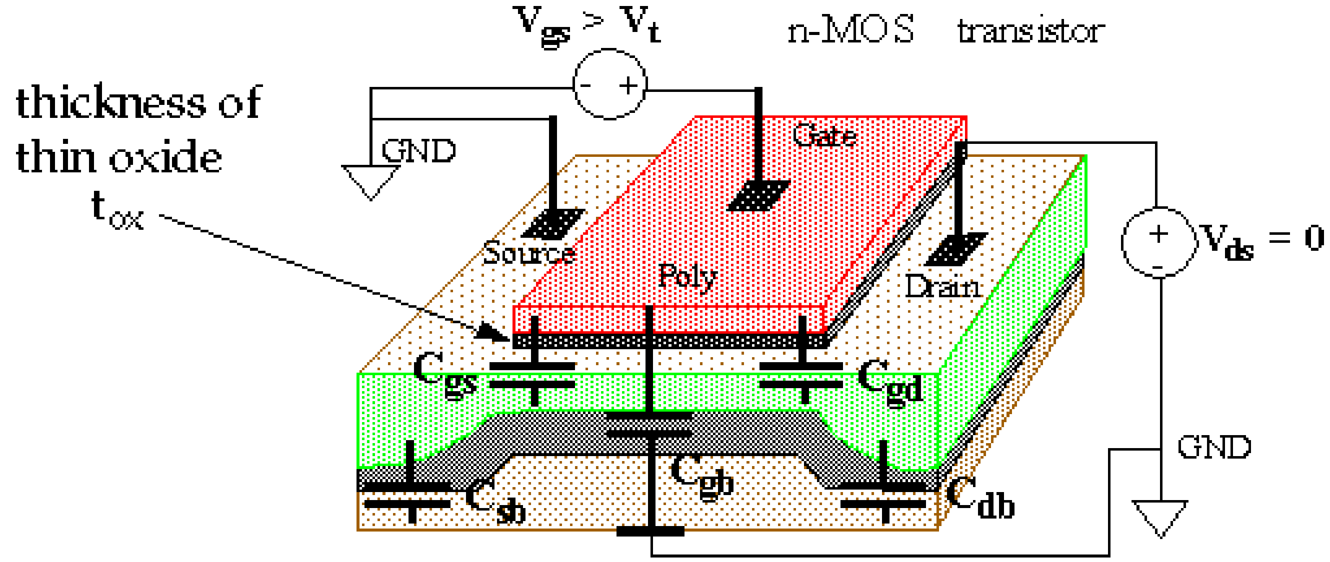
- The p+ regions around the source and drain, namely the channel-stop implants, are used to prevent the formation of any unwanted channels between two neighboring n+ diffusion regions

MOSFET Capacitor (3)



- MOSFET parasitic capacitances are observed bet terminals.
- Most of the capacitances are distributed and the complex.
- Capacitances can be models as
  - Lumped
  - distributed
- Parasitic device capacitances can be classified into two major group
  - Oxide-related capacitance
  - Junction capacitance

THE UNIVERSITY OF
JORDAN


Vgs > Vt
n-MOS transistor
thickness of
thin oxide
tox
GND
Gate
Source
Poly
Drain
Vds = 0
Cgs
Cgd
Cgb
Csb
Cdb
GND

Oxide-related Capacitances (1)

- The gate electrode overlaps both the source region and the drain region at the edges.
  - The two overlap capacitances that arise as a result of this structural arrangement.
    - C_{GD}(overlap)
    - C_{GS}(overlap)
  - The overlap capacitances can be found as

$$C_{GS}(overlap) = C_{ox} \cdot W \cdot L_D$$

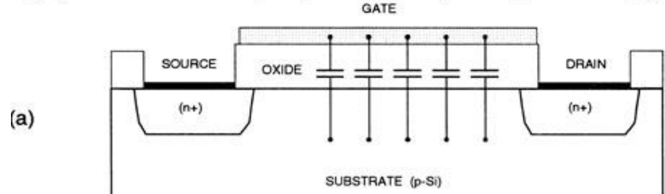
$$C_{GD}(overlap) = C_{ox} \cdot W \cdot L_D$$

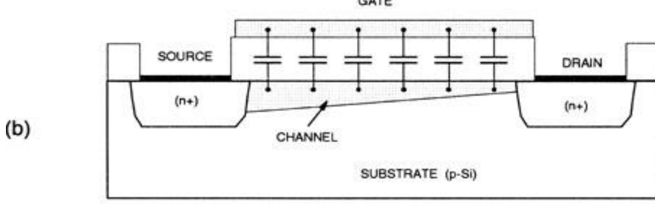
with $$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$

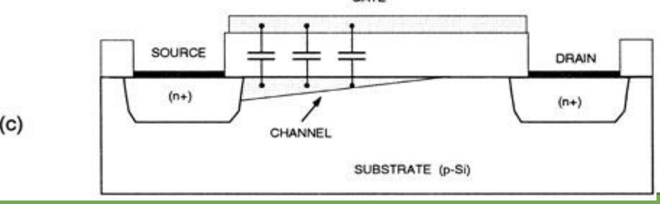
Oxide-related Capacitances (2)

- Capacitances which result from the interaction between the gate voltage and the channel charge.
  - C_{gs} ,C_{gd} ,C_{gb}

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



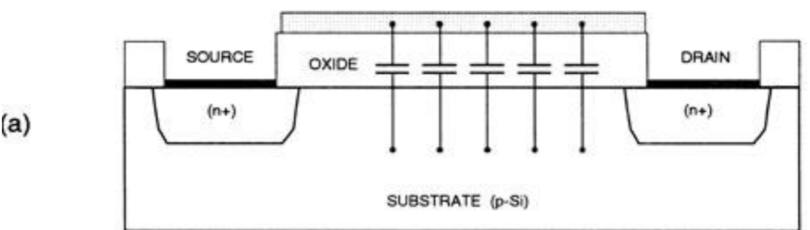




Oxide-related Capacitances (2)

- Cut-off mode
  - The surface is not inverted.
  - No conducting channel between source and drain
    - $C_{gs} = C_{gd} = 0$
  - The gate-to-substrate capacitance can be approximated by

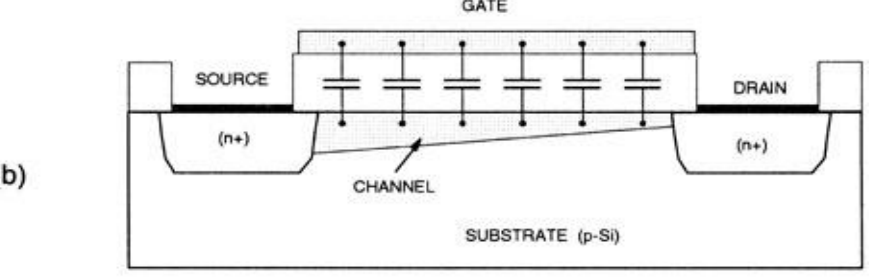
$$C_{gb} = C_{ox} \cdot W \cdot L \qquad (3.120)$$



Oxide-related Capacitances (2)

- Linear mode
  - The inverted channel extends across the MOSFET.
  - Conducting inversion layer shields the substrate from the gate electric field : $C_{gb} = 0$
  - The distributed gate-to-channel capacitance (equal S,D)

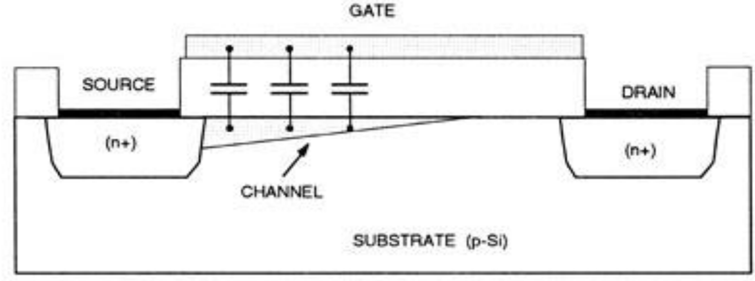
$$C_{gs} \approx C_{gd} \cong \frac{1}{2} \cdot C_{ox} \cdot W \cdot L$$



Oxide-related Capacitances (3)

- Saturation mode
  - The inversion region is pinched off.
  - The gate-to-drain capacitance component is equal to zero
    - $C_{gd} = 0$
  - Source still linked to the conducting channel.
    - Shielding effect still remain : $C_{gb} = 0$
  - The distributed gate-to-channel capacitance as seen between the gate and the source can be approximated by

$$C_{gs} \cong \frac{2}{3} \cdot C_{ox} \cdot W \cdot L$$



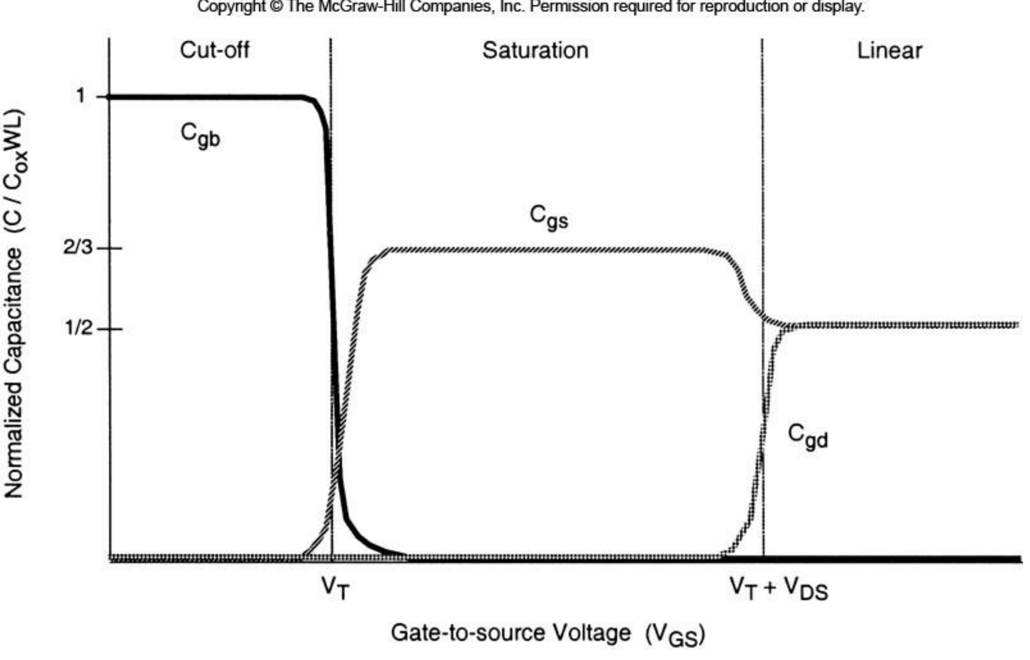
Oxide-related Capacitances (4)

Capacitance	Cut-off	Linear	Saturation
C_{gb} (total)	$C_{ox}WL$	0	0
C_{gd} (total)	$C_{ox}WL_D$	$1/2C_{ox}WL + C_{ox}WL_D$	$C_{ox}WL_D$
C_{gs} (total)	$C_{ox}WL_D$	$1/2C_{ox}WL + C_{ox}WL_D$	$2/3C_{ox}WL + C_{ox}WL_D$

- We have to combine the distributed C_{gs} and C_{gd} values found here with the relevant overlap capacitance values, in order to calculate the total capacitance between the external device terminals.

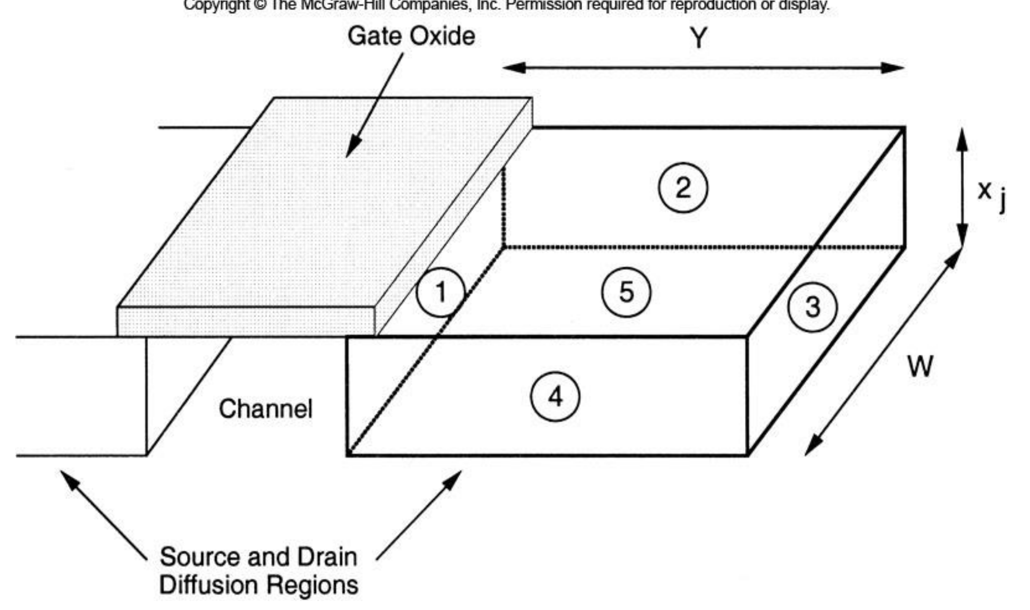
Oxide-related Capacitances (5)

- Variation of the distributed (gate-to-channel) oxide capacitances as function of gate-to-source voltage.



Junction Capacitances (1)

- Consider the voltage-dependent source-substrate and drain-substrate junction capacitances : C_{sb} , C_{db}



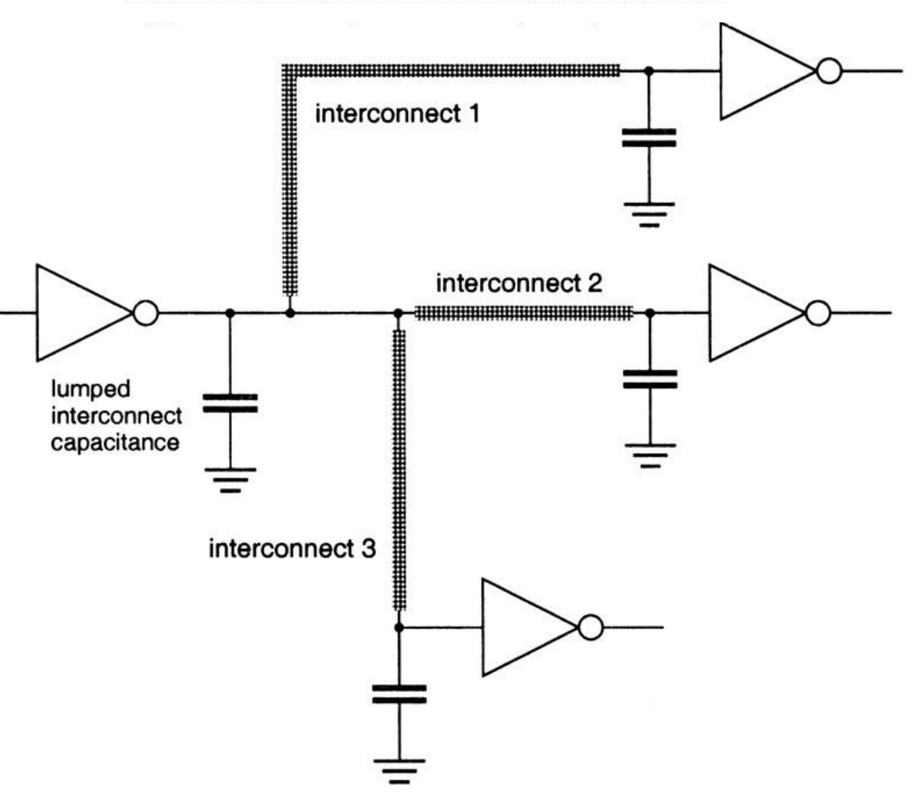
Junction	Area	Type
1	$W \cdot x_j$	n+/p
2	$Y \cdot x_j$	n+/p+
3	$W \cdot x_j$	n+/p+
4	$Y \cdot x_j$	n+/p+
5	$W \cdot Y$	n+/p

Junction Capacitances (2)

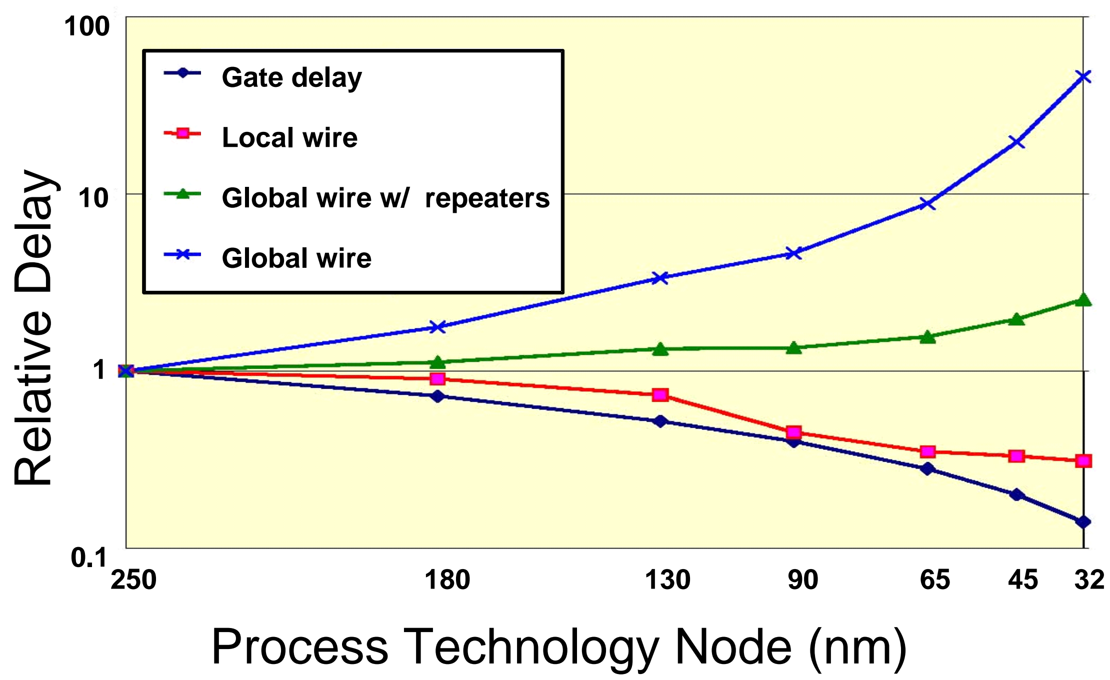
- C_{sb} and C_{db} are due to the depletion charge surrounding the respective source or drain diffusion regions embedded in the substrate.
  - both of these junctions are reverse-biased under normal operating conditions.
  - The amount of junction capacitance is a function of the applied terminal voltages.
  - Junction capacitances associated with sidewalls (2,3,4) will be different from the other junction capacitance.

Estimation of Interconnect Parasitics

- Main components of the output lc
  - Internal parasitic capacitances of tr
  - Interconnect capacitances
  - Input capacitances of the fan-out g



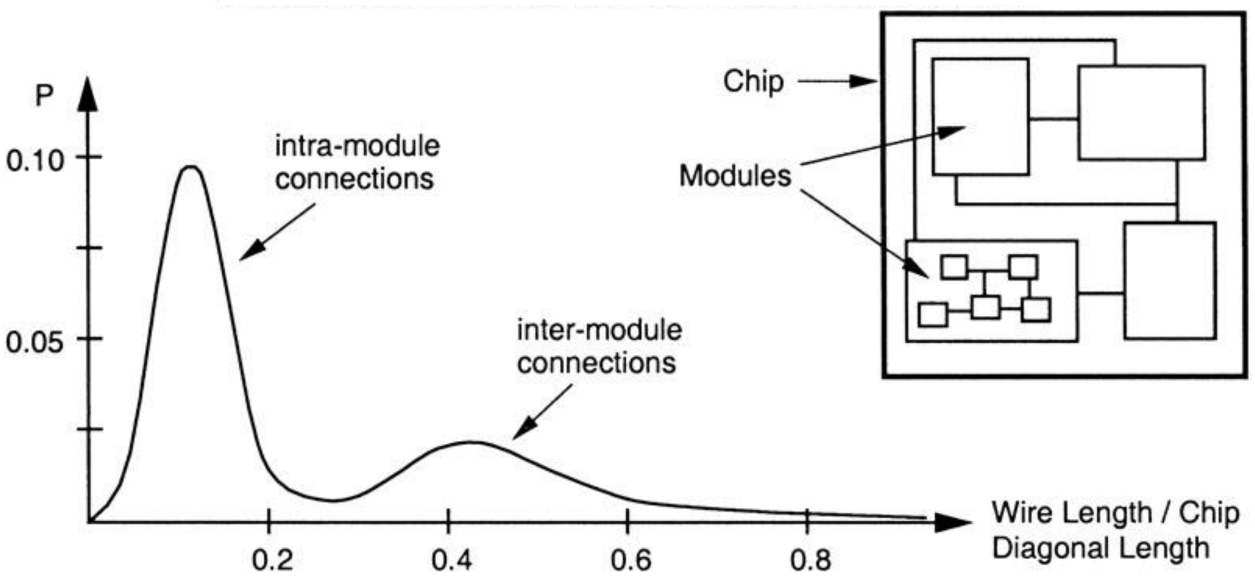
Interconnect Delay



- Dealing with the implications and optimizing a system for speed
  - Estimating the interconnect parasitics in a large chip
  - Simulating the transient effects.

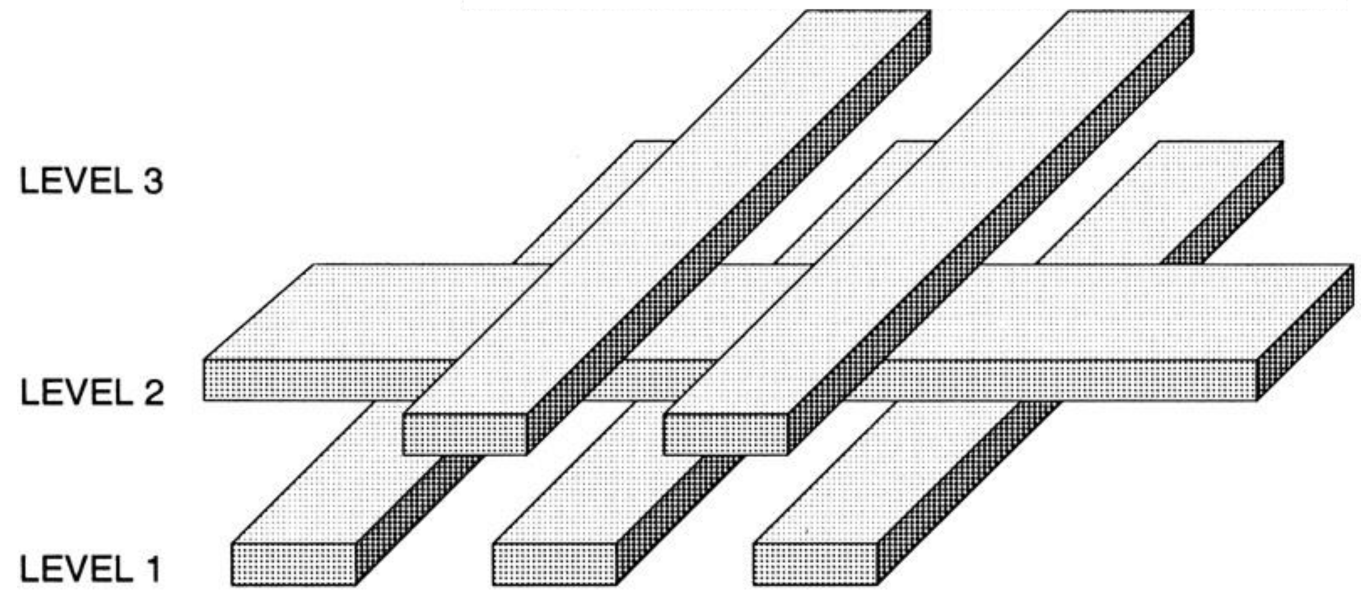
Statistical distribution

- Statistical distribution of interconnection length on a typical chip



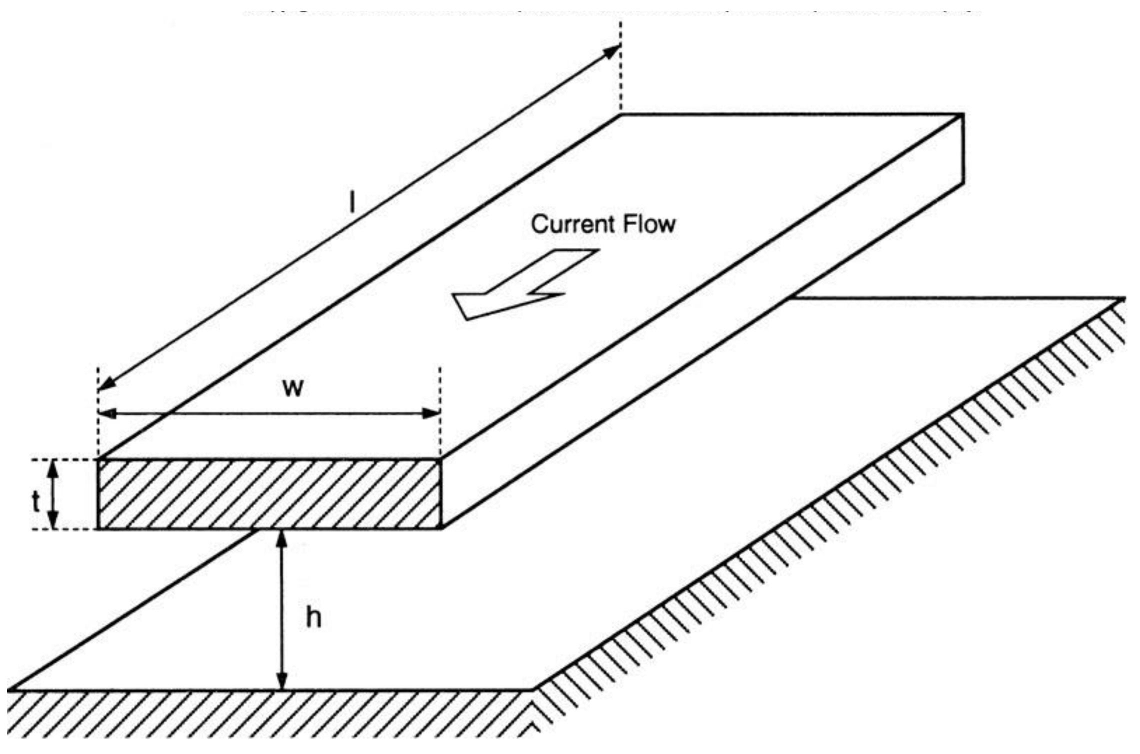
Interconnect Capacitance Estimation (1)

- A simplified view of six interconnections on three different levels



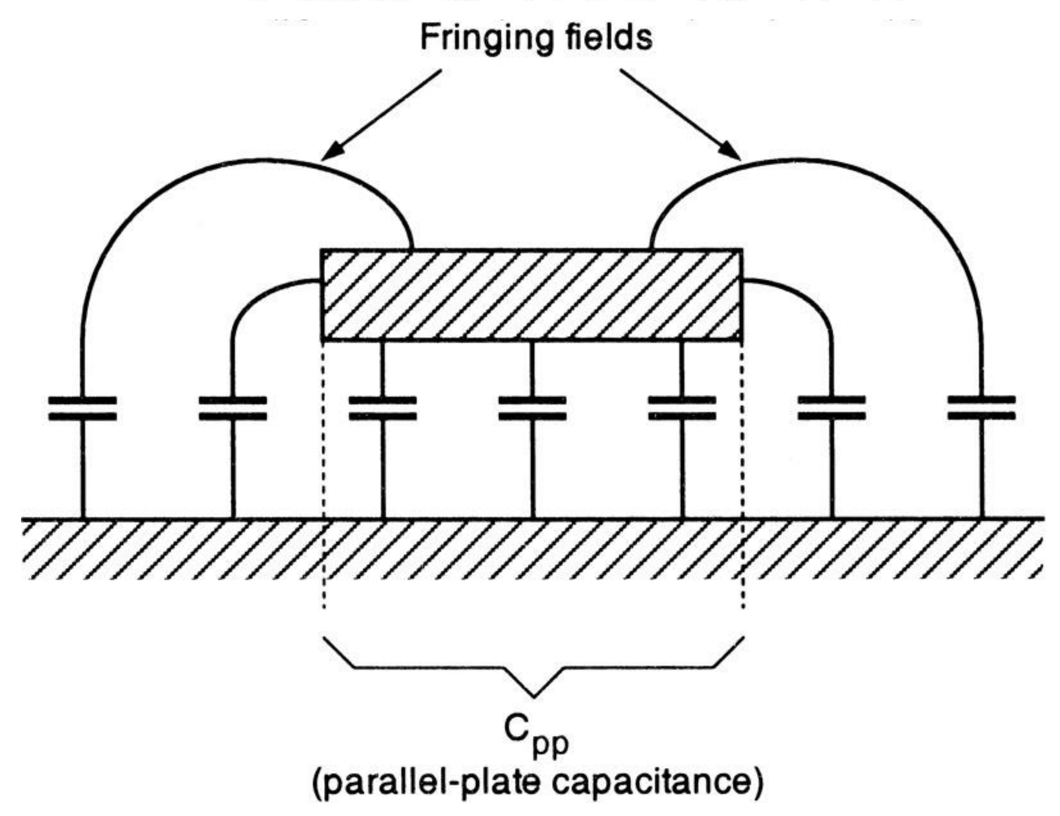
Interconnect Capacitance Estimation (2)

- The section of a single interconnect



Interconnect Capacitance Estimation (3)

- The total parasitic capacitance



Thickness value of different layers

Field oxide thickness	3µm	
Gate oxide thickness	2.6nm	
Polysilicon thickness	1µm	(minimum width 0.06µm)
Poly-metal oxide thickness	1.1µm	
Metal 1 thickness	1.8µm	(minimum width 0.09µm)
Metal 2~7 thickness	2.2µm	(minimum width 0.1µm)
Metal 8~9 thickness	9µm	(minimum width 0.4µm)
Via oxide thickness (PO-M1)	1.75µm	
Via oxide thickness (M1-M6)	2.2µm	
Via oxide thickness (M6-M9)	9µm	
n^+ junction depth	23nm	
p^+ junction depth	28nm	
n-well junction depth	3µm	

Interconnect Resistance Estimation

- Total resistance in indicated current direction

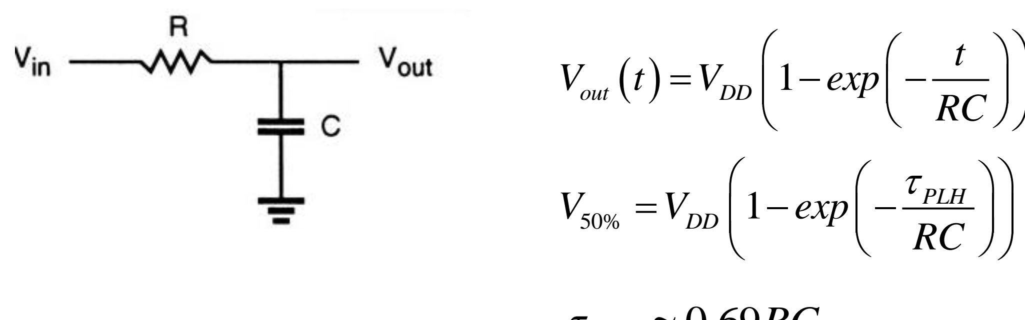
$$R_{wire} = \rho \cdot \frac{l}{w \cdot t} = R_{sheet}\left(\frac{l}{w}\right)$$

- The sheet resistivity of the line

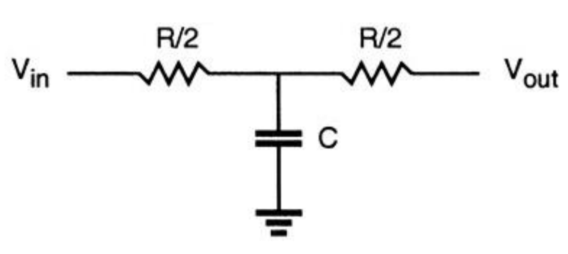
$$R_{sheet} = \left(\frac{\rho}{t}\right)$$

RC Delay Models

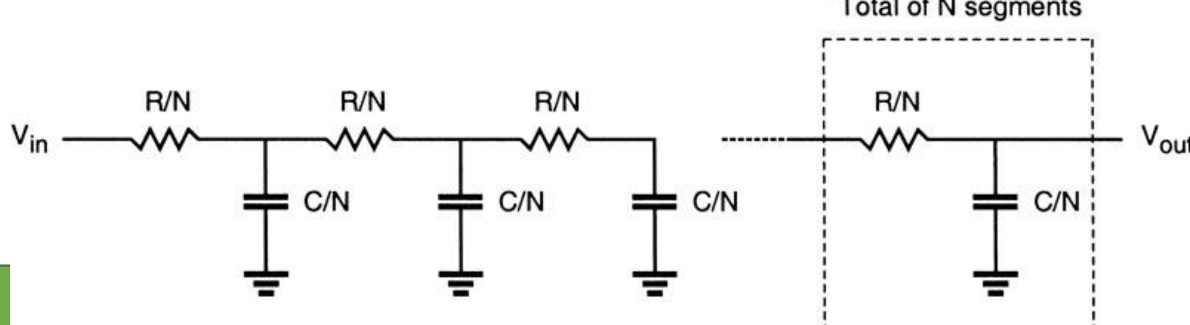
- Simple lumped RC model & T-model



$$\tau_{PLH} \approx 0.69RC$$



- Distributed RC ladder network model



Various RC Models

V_{in} R V_{out} C

(a) lumped RC model

V_{in} R/2 R/2 V_{out} C

(b) T-model

V_{in} R/4 R/2 R/4 V_{out} C/2 C/2

(c) T2-model

V_{in} R V_{out} C/2 C/2

(d) π-model

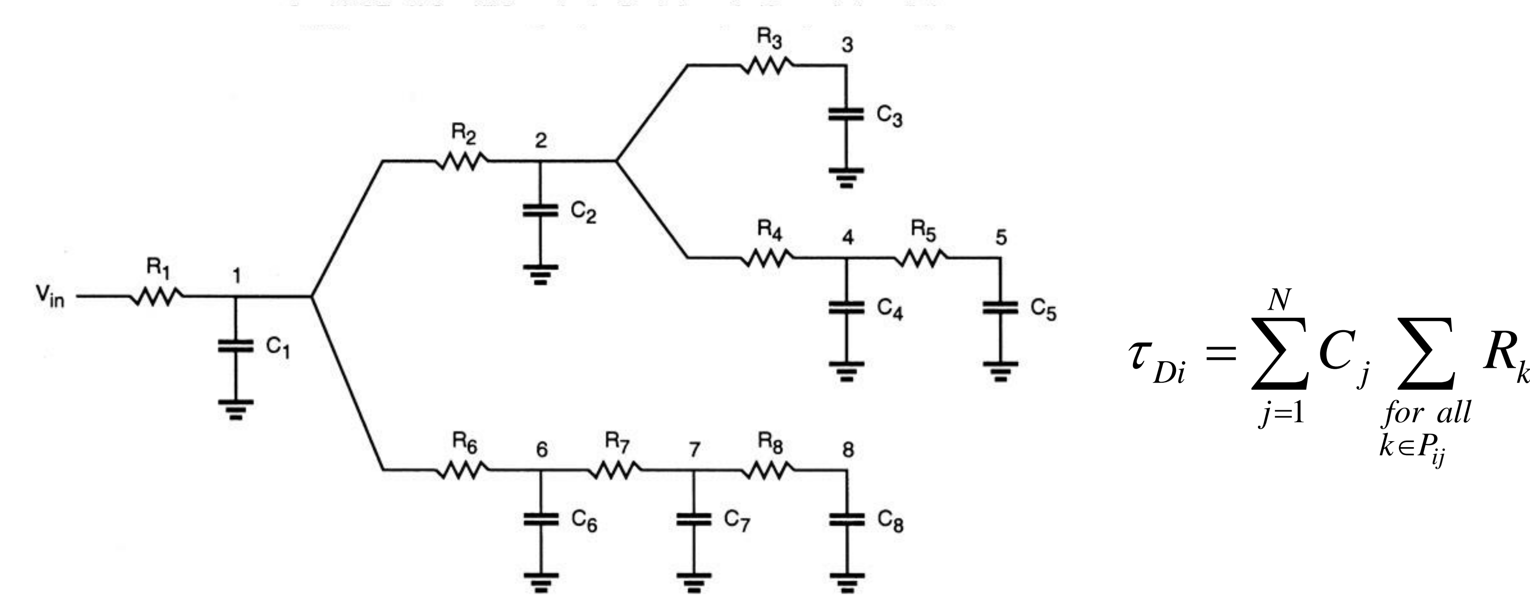
V_{in} R/2 R/2 V_{out} C/4 C/2 C/4

(e) π2-model

V_{in} R/3 R/3 R/3 V_{out} C/6 C/3 C/3 C/6

(f) π3-model

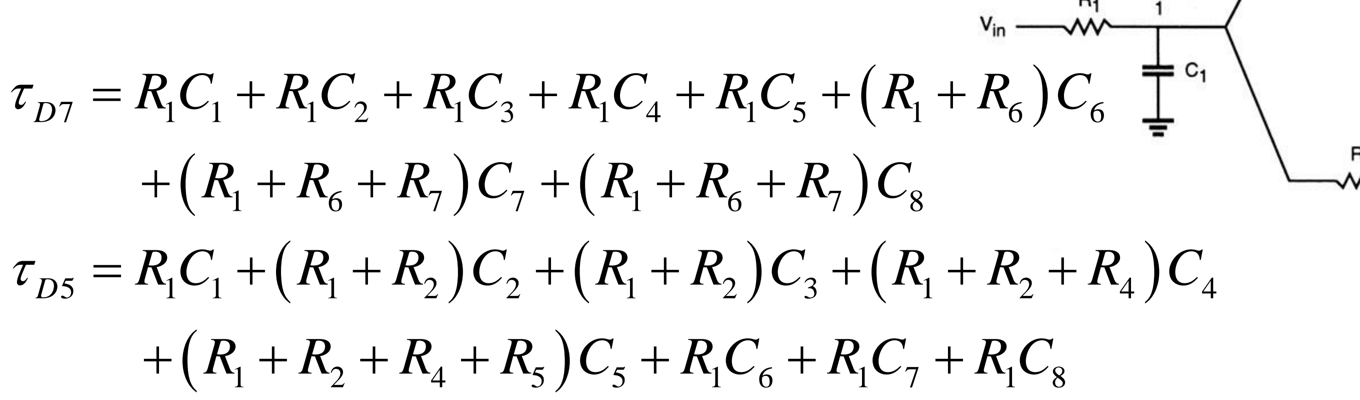
The Elmore Delay (1)



$$\tau_{Di} = \sum_{j=1}^{N} C_j \sum_{\substack{for\ all \\ k \in P_{ij}}} R_k$$

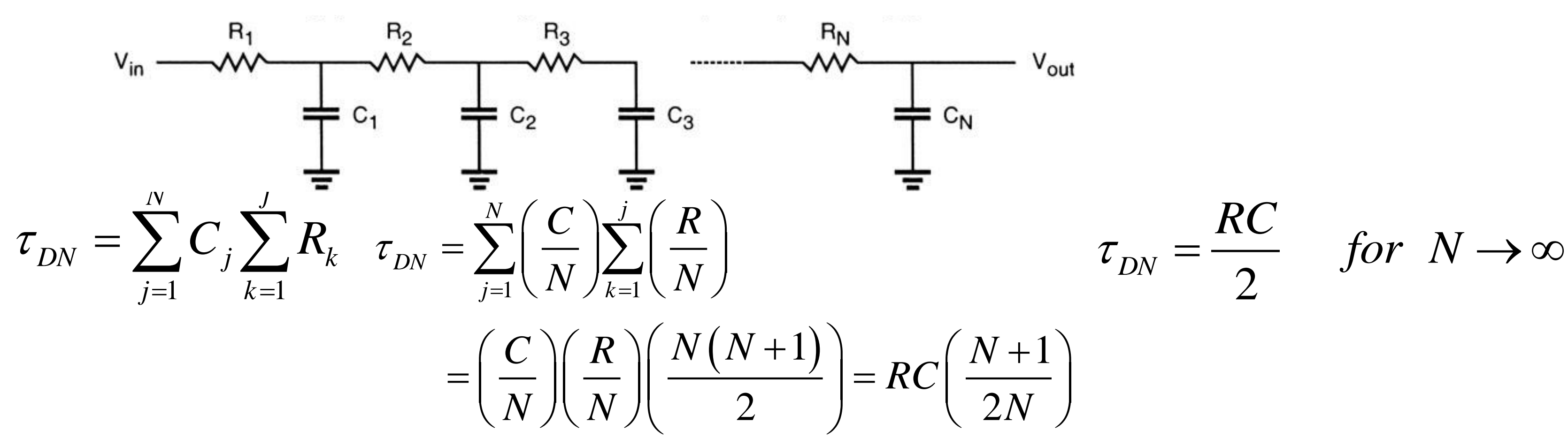
- The general topology of the RC tree network
  - Let *Pi* denote the unique path from the input node to node *i*, *i* = 1, 2, 3, ..., N.
  - Let *Pij* = *Pi* « *Pj* denote the portion of the path between the input and the node *i,* which is *common* to the path between the input and node *j*.

The Elmore Delay (2)



$$\tau_{D7} = R_1C_1 + R_1C_2 + R_1C_3 + R_1C_4 + R_1C_5 + \left(R_1 + R_6\right)C_6 + \left(R_1 + R_6 + R_7\right)C_7 + \left(R_1 + R_6 + R_7\right)C_8$$

$$\tau_{D5} = R_1C_1 + \left(R_1 + R_2\right)C_2 + \left(R_1 + R_2\right)C_3 + \left(R_1 + R_2 + R_4\right)C_4 + \left(R_1 + R_2 + R_4 + R_5\right)C_5 + R_1C_6 + R_1C_7 + R_1C_8$$



$$\tau_{DN} = \sum_{j=1}^{N} C_j \sum_{k=1}^{j} R_k \qquad \tau_{DN} = \sum_{j=1}^{N}\left(\frac{C}{N}\right)\sum_{k=1}^{j}\left(\frac{R}{N}\right) = \left(\frac{C}{N}\right)\left(\frac{R}{N}\right)\left(\frac{N\left(N+1\right)}{2}\right) = RC\left(\frac{N+1}{2N}\right)$$

$$\tau_{DN} = \frac{RC}{2} \quad for \;\; N \rightarrow \infty$$

Example 6.5 (optional)

- 1) Examine the propagation delay across a long polysilicon interconnect line (length=1000μm, width=1μm)
  - R_{sheet}= 15 Ω/square
- 1) Sol.

$$R_{lumped} = R_{sheet} \times (\#\text{ of squares})$$

$$= 15(\Omega/\text{square}) \times \left(\frac{1000\mu\text{m}}{1\mu\text{m}} \right) = 15\text{k}\Omega$$

$$C_{parallel-plate} = (\text{unit area capacitance}) \times (\text{area})$$

$$= 0.106\text{fF}/\mu\text{m}^2 \times (1000\mu\text{m} \times 1\mu\text{m}) = 106\text{fF}$$

$$C_{fringe} = (\text{unit length capacitance}) \times (\text{perimeter})$$

$$= 0.043\text{fF}/\mu\text{m} \times (1000\mu\text{m} + 1000\mu\text{m} + 1\mu\text{m} + 1\mu\text{m}) = 86\text{fF}$$

$$C_{lumped-total} = C_{parallel-plate} + C_{fringe} = 192\text{fF}$$

Example 6.5 (2)

- The simulated output voltage waveforms

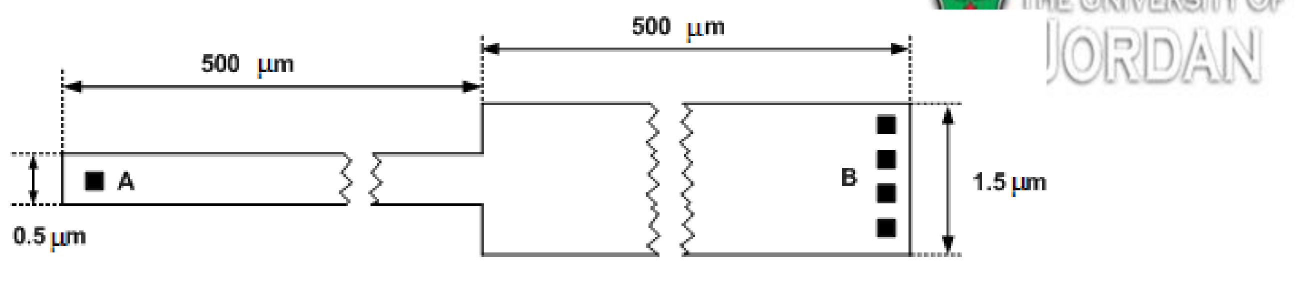
Voltage (V)
1.2
1.0
0.8
0.6
0.4
0.2
0.0

3.5×10^{-8} 4.0×10^{-8} 4.5×10^{-8} 5.0×10^{-8} 5.5×10^{-8} 6.0×10^{-8} 6.5×10^{-8} 7.0×10^{-8} 7.5×10^{-8}

Time (s)

Input
Lumped RC model
Distributed RC model

Example 6.5 (4)



- 2) Consider a polysilicon line consisting of two segments (W=1.5 µm & W=0.5 µm, each 500 µm)

- 2) Sol.

$$R_{lumped-1} = 15(\Omega/\text{square}) \times \left(\frac{500\mu\text{m}}{0.5\mu\text{m}}\right) = 15\text{k}\Omega,\ R_{lumped-2} = 15(\Omega/\text{square}) \times \left(\frac{500\mu\text{m}}{1.5\mu\text{m}}\right) = 5\text{k}\Omega$$

$$R_{lumped-total} = R_{lumped-1} + R_{lumped-2} = 20\text{k}\Omega$$

$$C_{parallel-plate-1} = 0.106\text{fF}/\mu\text{m}^2 \times (500\mu\text{m} \times 0.5\mu\text{m}) = 26.5\text{fF}$$

$$C_{parallel-plate-2} = 0.106\text{fF}/\mu\text{m}^2 \times (500\mu\text{m} \times 1.5\mu\text{m}) = 79.5\text{fF}$$

$$C_{fringe-1} \approx C_{fringe-2} = 46\text{fF}$$

$$C_{lumped-total} = C_{parallel-plate-1} + C_{parallel-plate-2} + C_{fringe-1} + C_{fringe-2} = 192\text{fF}$$

Layout, Fabrication, and Elementary Logic Design

Dr.Mohammad Abdel-Majeed

Assistant Professor

University of Jordan

Introduction

- Integrated circuits: many transistors on one chip.
  - *Very Large Scale Integration* (VLSI): very many
- *Metal Oxide Semiconductor* (MOS) transistor
  - Fast, cheap, low-power transistors
  - Complementary: mixture of n- and p-type leads to less power
- Today: How to build your own simple CMOS chip
  - CMOS transistors
  - Building logic gates from transistors
  - Transistor layout and fabrication
- Rest of the course: How to build a good CMOS chip

Silicon Lattice

- Transistors are built on a silicon substrate
- Silicon is a Group IV material
- Forms crystal lattice with bonds to four neighbors

```
    ||     ||     ||
══ Si ══ Si ══ Si ══
    ||     ||     ||
══ Si ══ Si ══ Si ══
    ||     ||     ||
══ Si ══ Si ══ Si ══
    ||     ||     ||
```

Dopants

- Silicon is a semiconductor
- Pure silicon has no free carriers and conducts poorly
- Adding dopants increases the conductivity
- Group V: extra electron (n-type)
- Group III: missing electron, called hole (p-type)

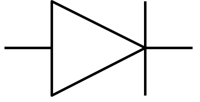
```
    ||     ||     ||              ||     ||     ||
== Si == Si == Si ==          == Si == Si == Si ==
    ||     ||-    ||              ||     |:+    ||
    ||     ||  +  ||              ||     |: -   ||
== Si == As == Si ==          == Si ==  B  == Si ==
    ||     ||     ||              ||     ||     ||
== Si == Si == Si ==          == Si == Si == Si ==
    ||     ||     ||              ||     ||     ||
```

p-n Junctions

- A junction between p-type and n-type semiconductor forms a diode.
- Current flows only in one direction

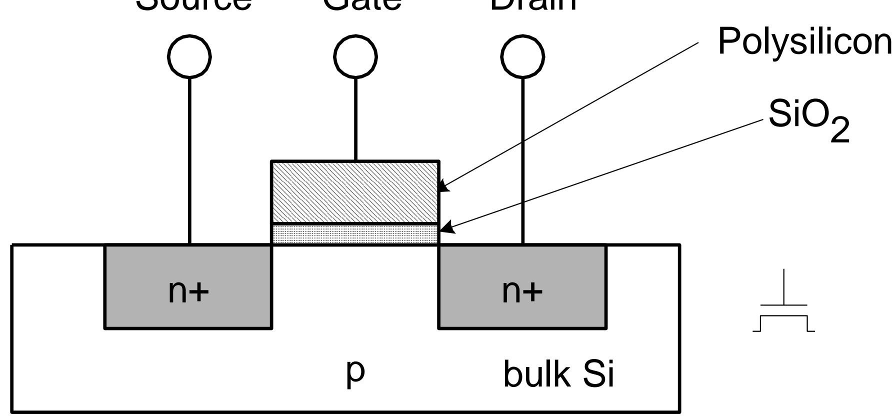
p-type	n-type

anode cathode



nMOS Transistor

- Four terminals: gate, source, drain, body
- Gate – oxide – body stack looks like a capacitor
  - Gate and body are conductors
  - SiO_2 (oxide) is a very good insulator
  - Called metal – oxide – semiconductor (MOS) capacitor
  - Even though gate is no longer made of metal



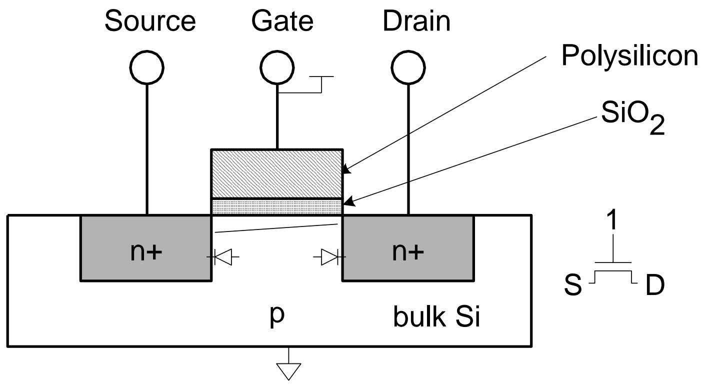
nMOS Operation

- Body is commonly tied to ground (0 V)
- When the gate is at a low voltage:
  - P-type body is at low voltage
  - Source-body and drain-body diodes are OFF
  - No current flows, transistor is OFF

Source
Gate
Drain
Polysilicon
SiO_2
n+
n+
p
bulk Si
0
S
D

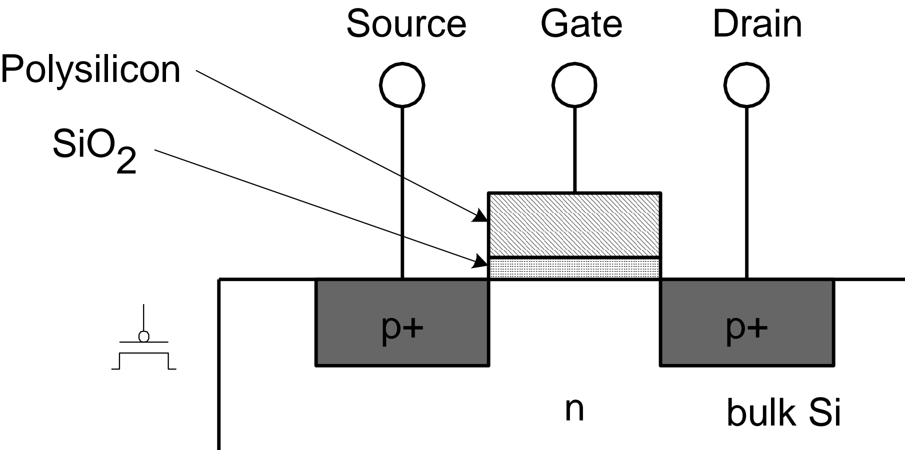
nMOS Operation

- When the gate is at a high voltage:
  - Positive charge on gate of MOS capacitor
  - Negative charge attracted to body
  - Inverts a channel under gate to n-type
  - Now current can flow through n-type silicon from source through channel to drain, transistor is ON



pMOS Transistor

- Similar, but doping and voltages reversed
  - Body tied to high voltage (V_{DD})
  - Gate low: transistor ON
  - Gate high: transistor OFF
  - Bubble indicates inverted behavior



Power Supply Voltage

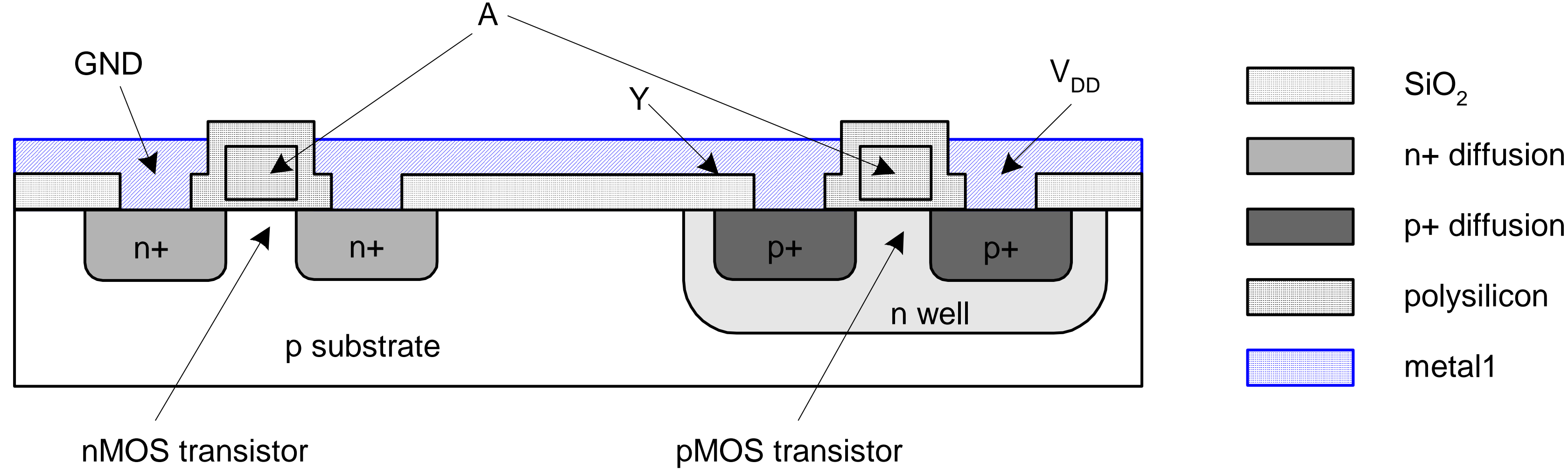
- GND = 0 V
- In 1980's, V_{DD} = 5V
- V_{DD} has decreased in modern processes
  - High V_{DD} would damage modern tiny transistors
  - Lower V_{DD} saves power
- V_{DD} = 3.3, 2.5, 1.8, 1.5, 1.2, 1.0, …

CMOS Fabrication

- CMOS transistors are fabricated on silicon wafer
- Lithography process similar to printing press
- On each step, different materials are deposited or etched
- Easiest to understand by viewing both top and cross-section of wafer in a simplified manufacturing process

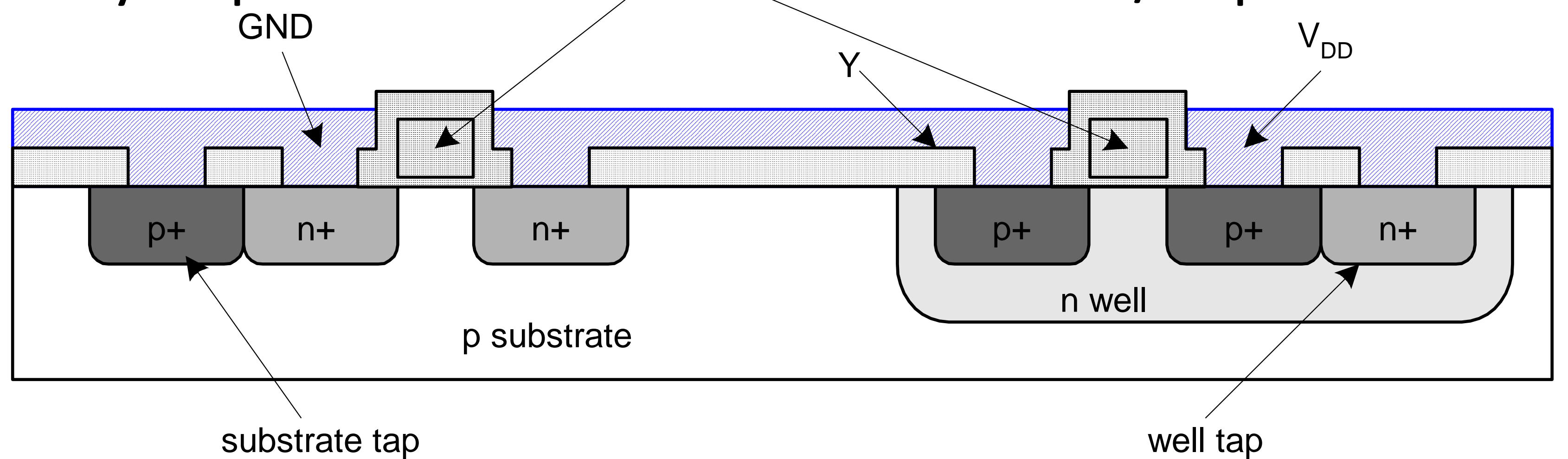
Inverter Cross-section

- Typically use p-type substrate for nMOS transistor
  - Requires n-well for body of pMOS transistors
  - Several alternatives: SOI, twin-tub, etc.



Well and Substrate Taps

- Substrate must be tied to GND and n-well to V_{DD}
- Metal to lightly-doped semiconductor forms poor connection called Shottky Diode
- Use heavily doped well and substrate contacts / taps

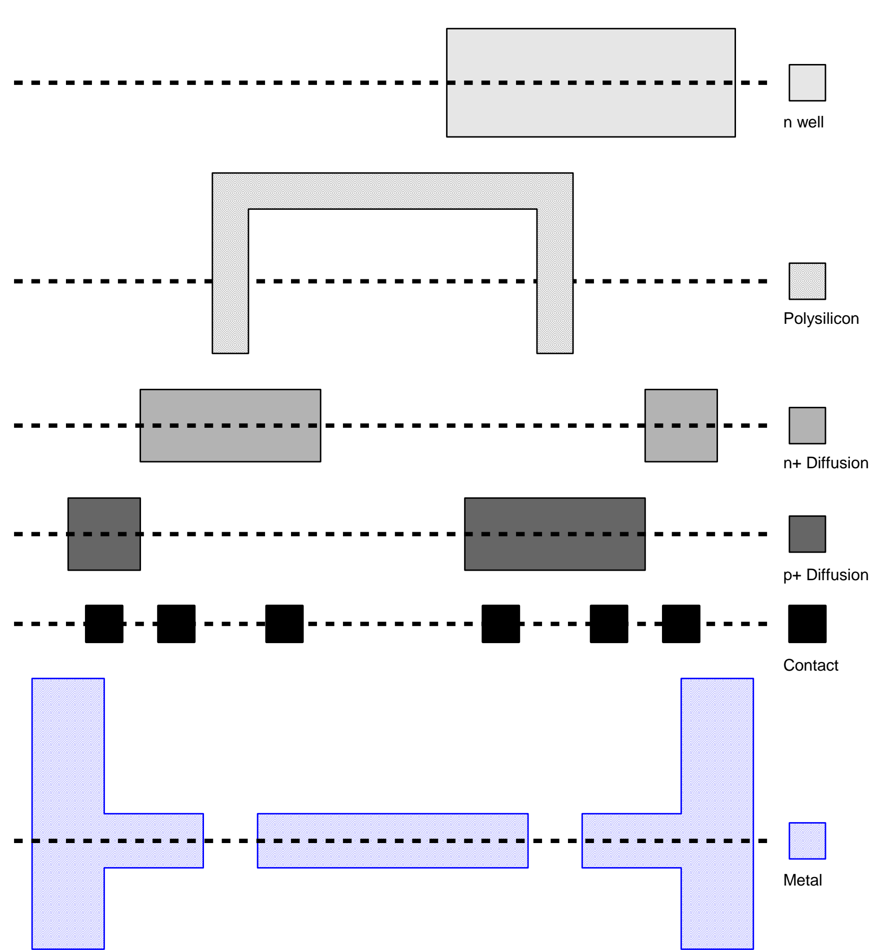


Inverter Mask Set

- Transistors and wires are defined by *masks*
- Cross-section taken along dashed line

Detailed Mask Views

- Six masks
  - n-well
  - Polysilicon
  - n+ diffusion
  - p+ diffusion
  - Contact
  - Metal

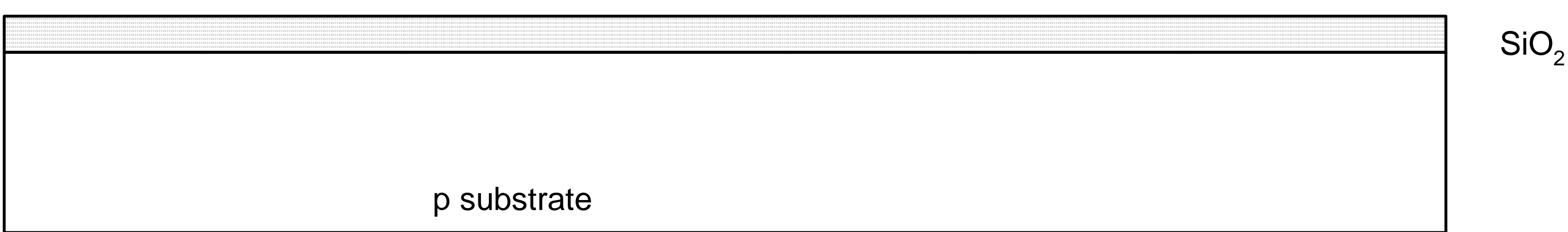


Fabrication Steps

- Start with blank wafer
- Build inverter from the bottom up
- First step will be to form the n-well
  - Cover wafer with protective layer of SiO_2 (oxide)
  - Remove layer where n-well should be built
  - Implant or diffuse n dopants into exposed wafer
  - Strip off SiO_2

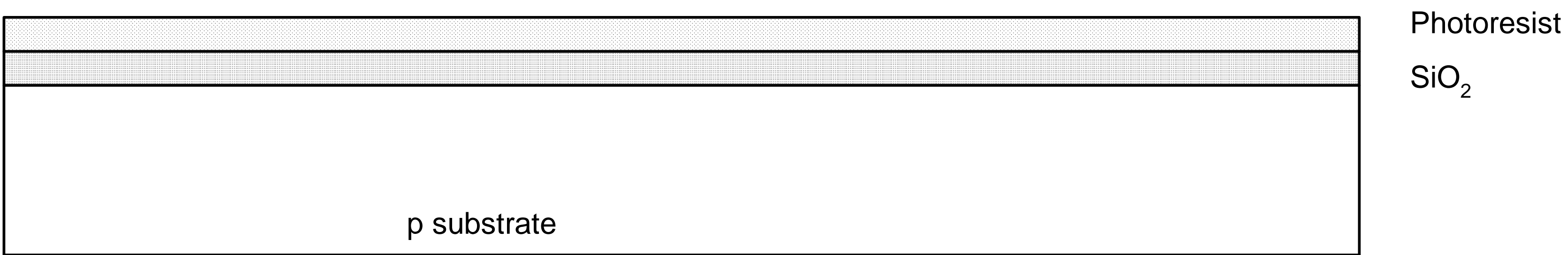
Oxidation

- Grow SiO_2 on top of Si wafer
  - 900 – 1200 C with H_2O or O_2 in oxidation furnace



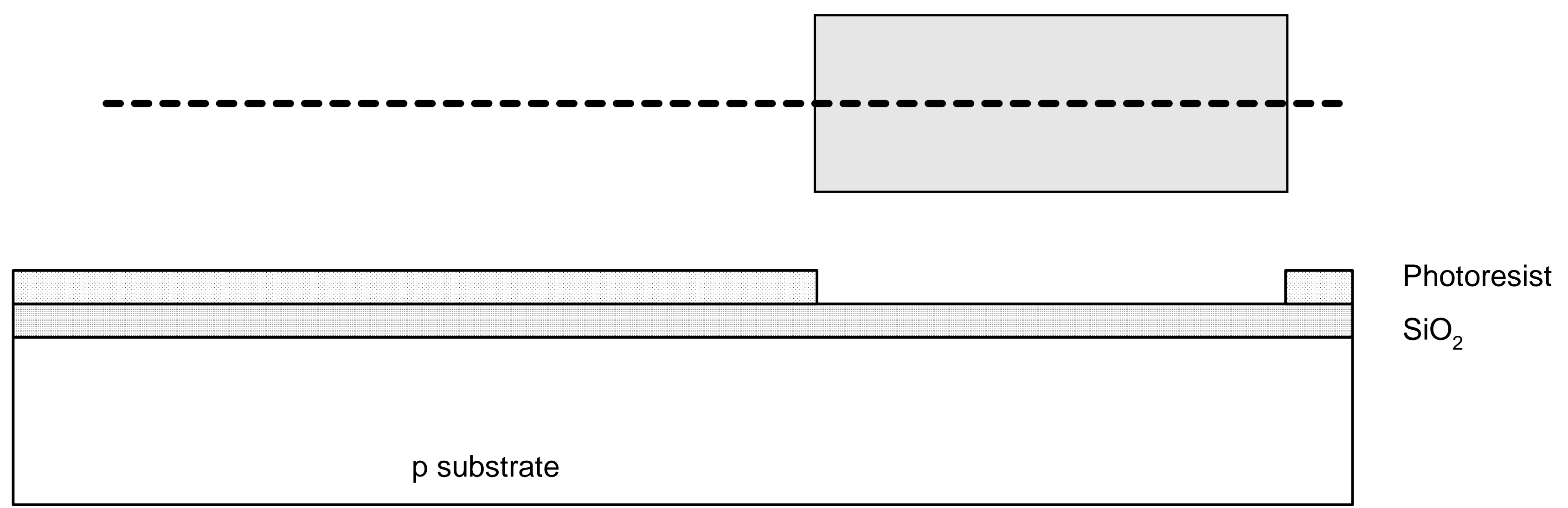
Photoresist

- Spin on photoresist
  - Photoresist is a light-sensitive organic polymer
  - Softens where exposed to light



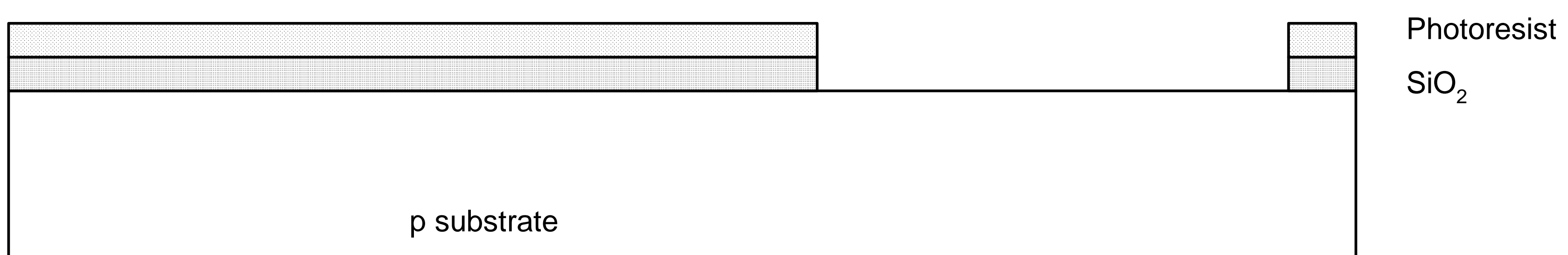
Lithography

- Expose photoresist through n-well mask
- Strip off exposed photoresist



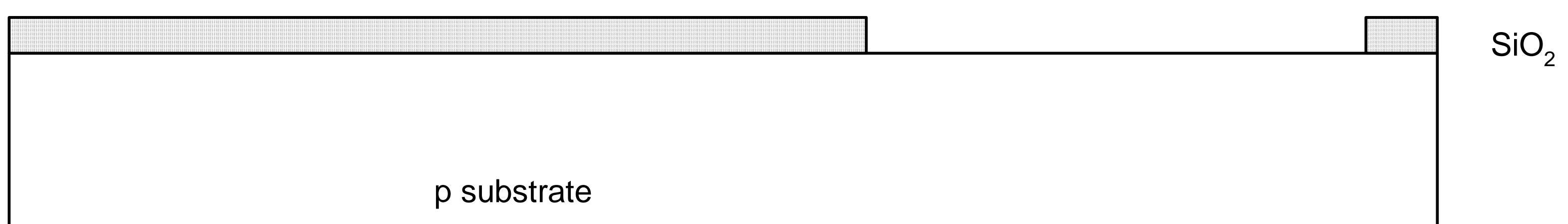
Etch

- Etch oxide with hydrofluoric acid (HF)
  - Seeps through skin and eats bone; nasty stuff!!!
- Only attacks oxide where resist has been exposed



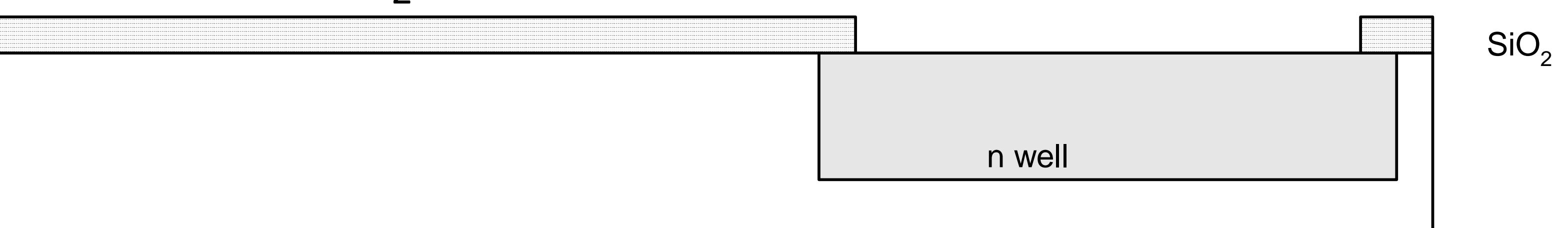
Strip Photoresist

- Strip off remaining photoresist
  - Use mixture of acids called piranah etch
- Necessary so resist doesn't melt in next step



n-well

- n-well is formed with diffusion or ion implantation
- Diffusion
  - Place wafer in furnace with arsenic gas
  - Heat until As atoms diffuse into exposed Si
- Ion Implanatation
  - Blast wafer with beam of As ions
  - Ions blocked by SiO_2, only enter exposed Si

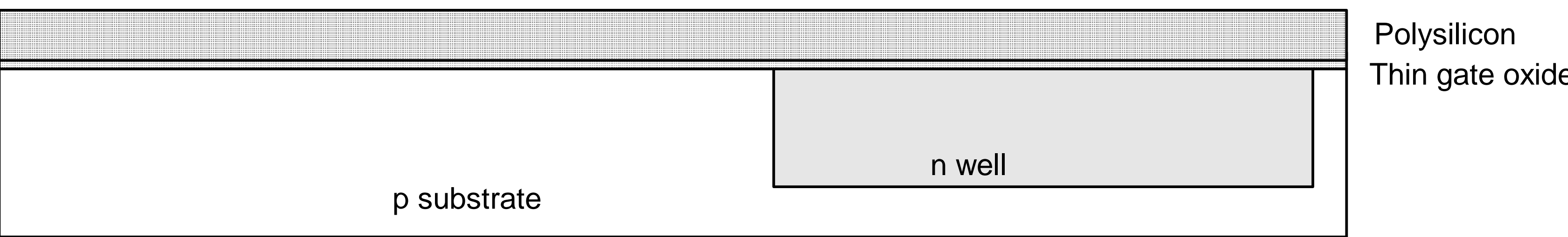


Strip Oxide

- Strip off the remaining oxide using HF
- Back to bare wafer with n-well
- Subsequent steps involve similar series of steps

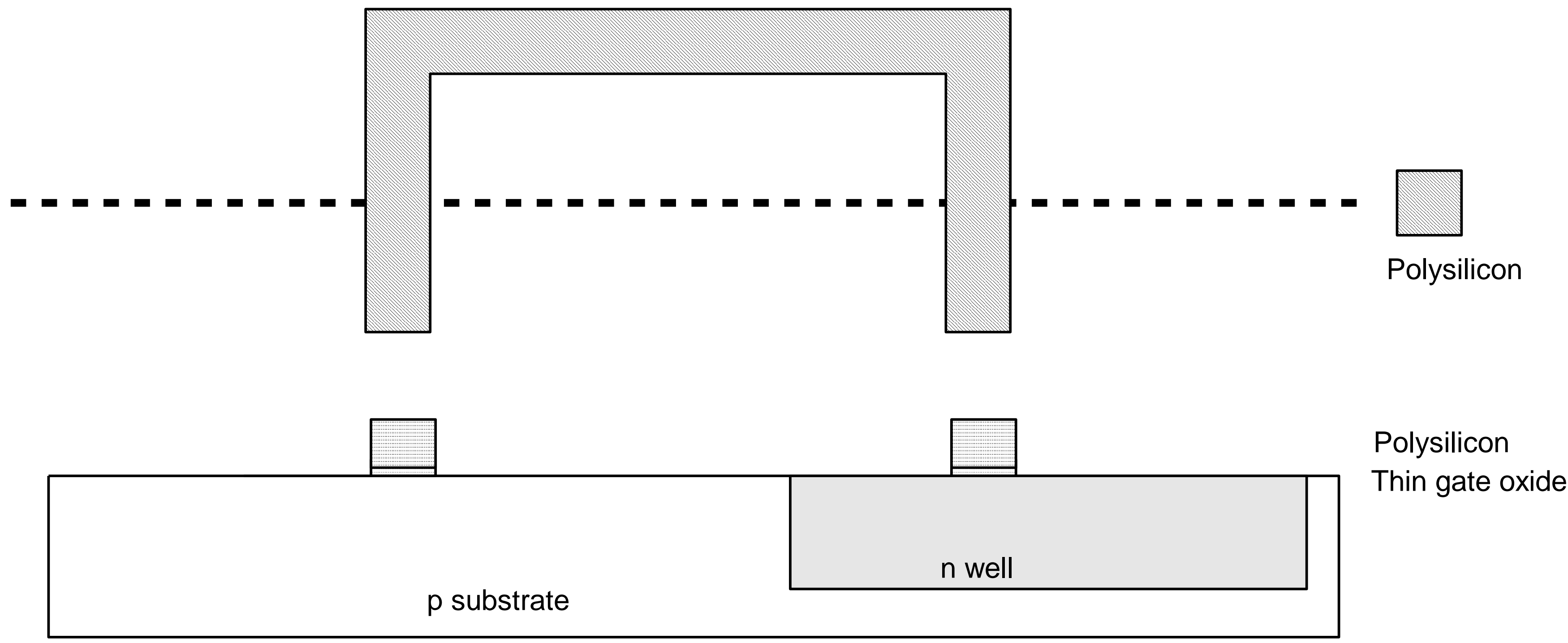
Polysilicon

- Deposit very thin layer of gate oxide
  - < 20 Å (6-7 atomic layers)
- Chemical Vapor Deposition (CVD) of silicon layer
  - Place wafer in furnace with Silane gas (SiH_4)
  - Forms many small crystals called polysilicon
  - Heavily doped to be good conductor



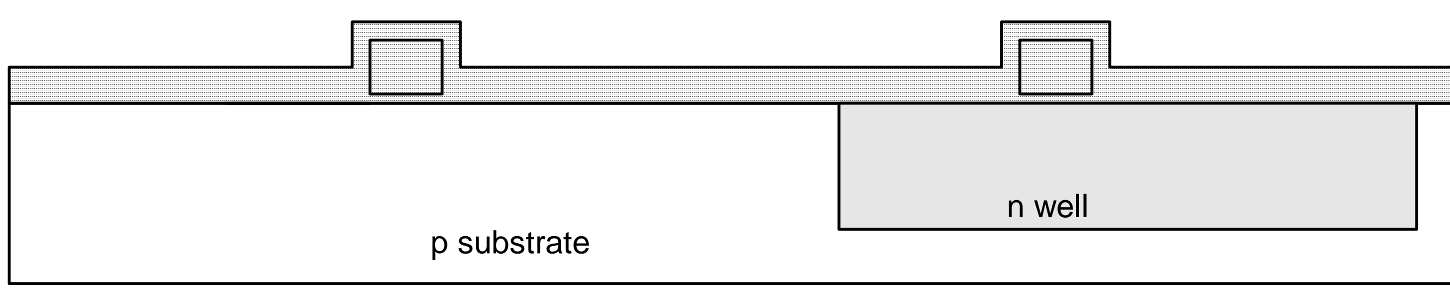
Polysilicon Patterning

- Use same lithography process to pattern polysilicon



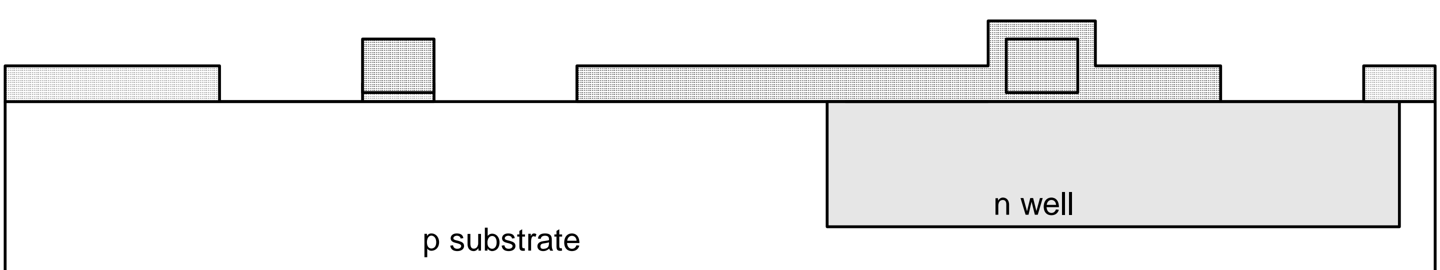
Self-Aligned Process

- Use oxide and masking to expose where n+ dopants should be diffused or implanted
- N-diffusion forms nMOS source, drain, and n-well contact



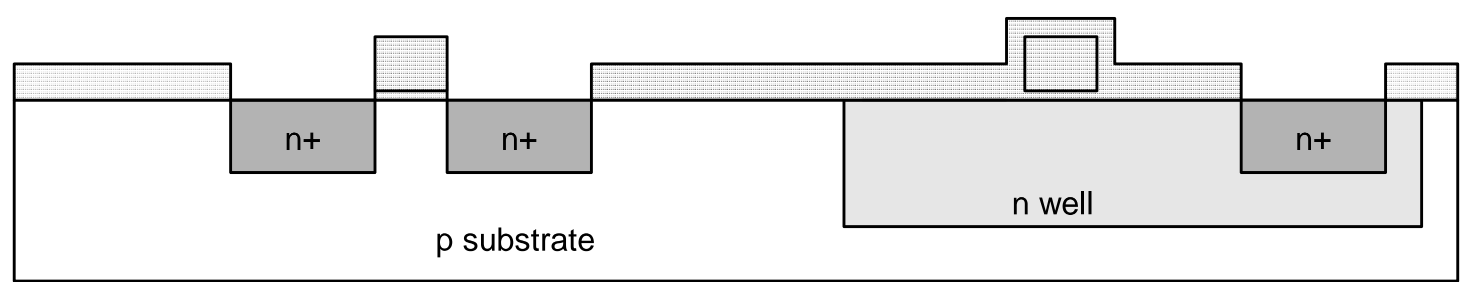
N-diffusion

- Pattern oxide and form n+ regions
- *Self-aligned process* where gate blocks diffusion
- Polysilicon is better than metal for self-aligned gates because it doesn’t melt d processing



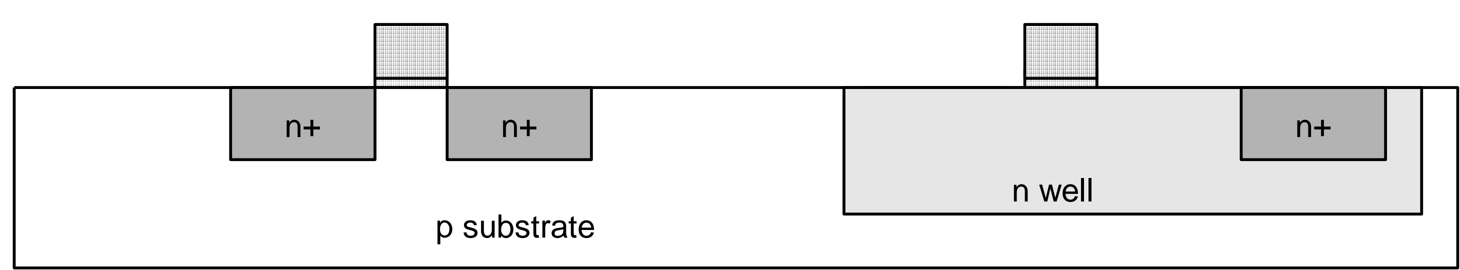
N-diffusion

- Historically dopants were diffused
- Usually ion implantation today
- But regions are still called diffusion



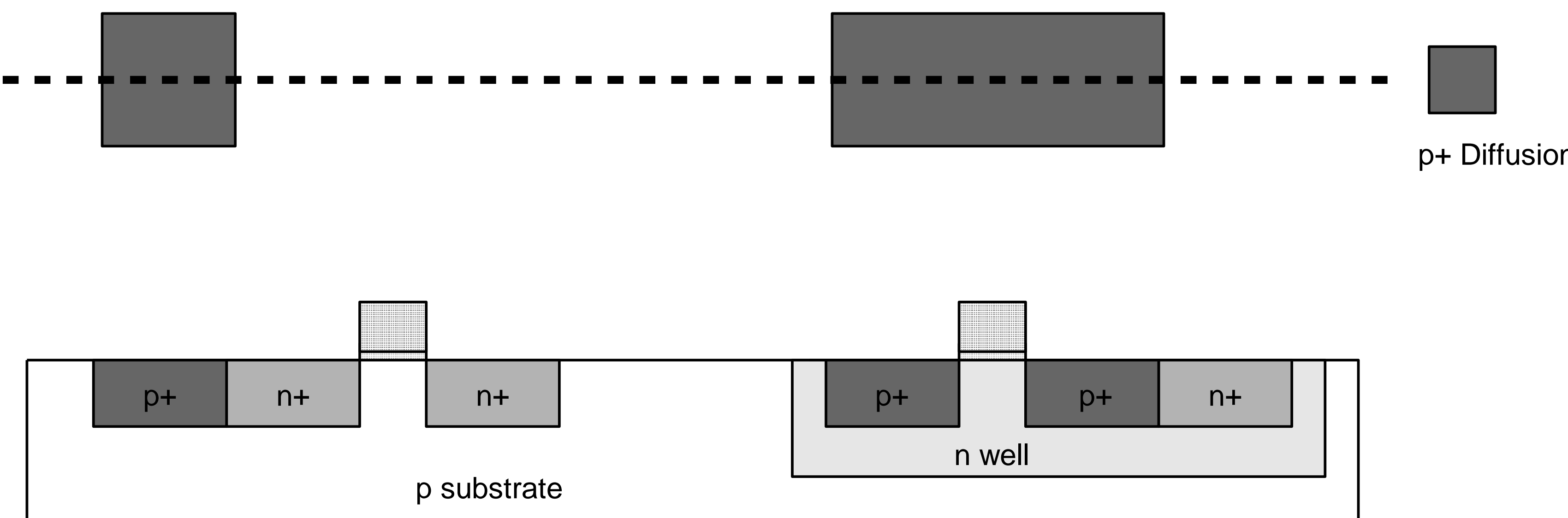
N-diffusion

- Strip off oxide to complete patterning step



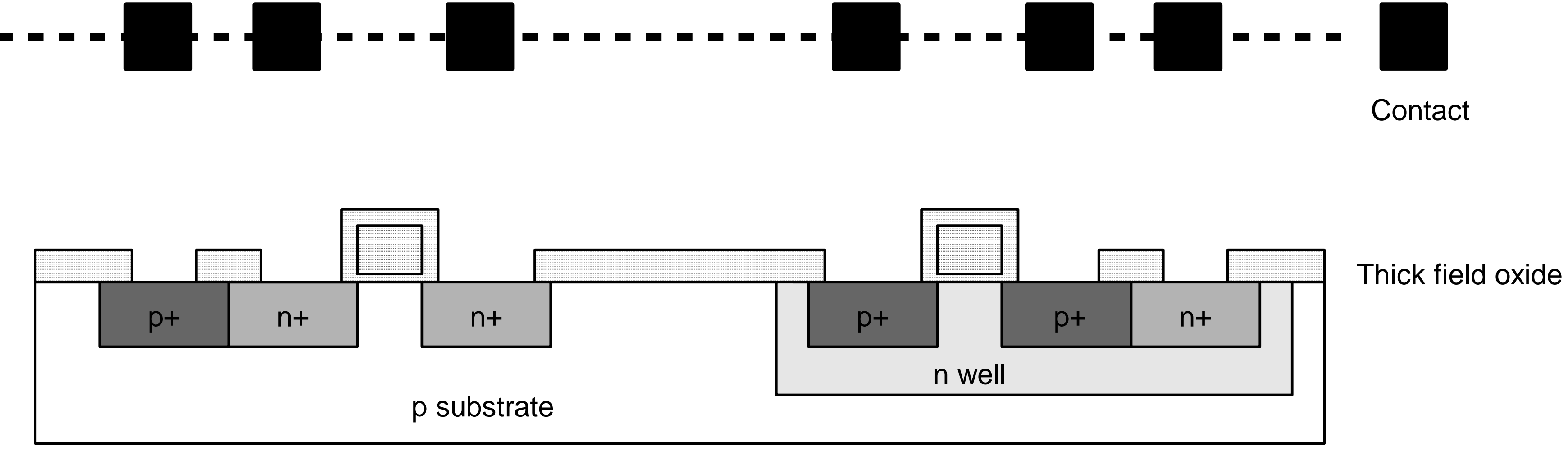
P-Diffusion

- Similar set of steps form p+ diffusion regions for pMOS source and drain and substrate contact



Contacts

- Now we need to wire together the devices
- Cover chip with thick field oxide
- Etch oxide where contact cuts are needed



Metallization

- Sputter on aluminum over whole wafer
- Patter remove excess metal, leaving wires

Metal

Metal

Thick field oxide

p+ n+ n+ p+ p+ n+

n well

p substrate

Layout

- Chips are specified with set of masks
- Minimum dimensions of masks determine transistor size (and hence speed, cost, and power)
- Feature size f = distance between source and drain
  - Set by minimum width of polysilicon
- Feature size improves 30% every 3 years or so
- Normalize for feature size when describing design rules
- Express rules in terms of $\lambda = f/2$
  - E.g. $\lambda = 0.3\ \mu m$ in $0.6\ \mu m$ process

Simplified Design Rules

- Conser

Metal1
4 λ 4 λ
spacing width

Metal2
4 λ 4 λ

Diffusion
4 λ 4 λ

Polysilicon
2 λ
3 λ

n-well
2 λ
6 λ
6 λ

1 λ
2 λ
3 λ

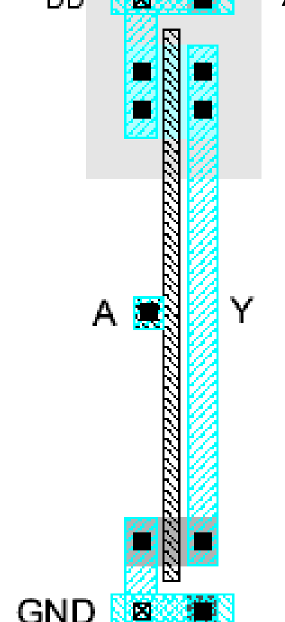
Metal1-
Diffusion
Contact

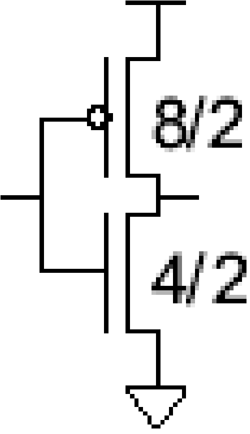
Metal1-
Polysilicon
Contact

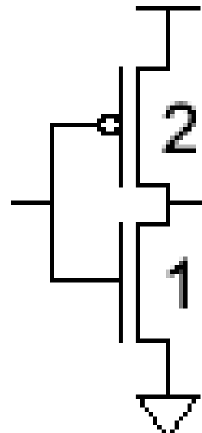
Metal1-
Metal2
Vias

Inverter Layout

- Transistor dimensions specified as Width / Length
  - Minimum size is 4λ / 2λ, sometimes called 1 unit
  - For 0.6 μm process, W=1.2 μm, L=0.6 μm







Summary

- MOS Transistors are stack of gate, oxide, silicon
- Can be viewed as electrically controlled switches
- Build logic gates out of switches
- Draw masks to specify layout of transistors

- Now you know everything necessary to start designing schematics and layout for a simple chip!

Combinational Logic

Dr.Mohammad Abdel-Majeed

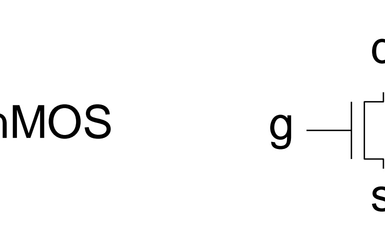
Assistant Professor

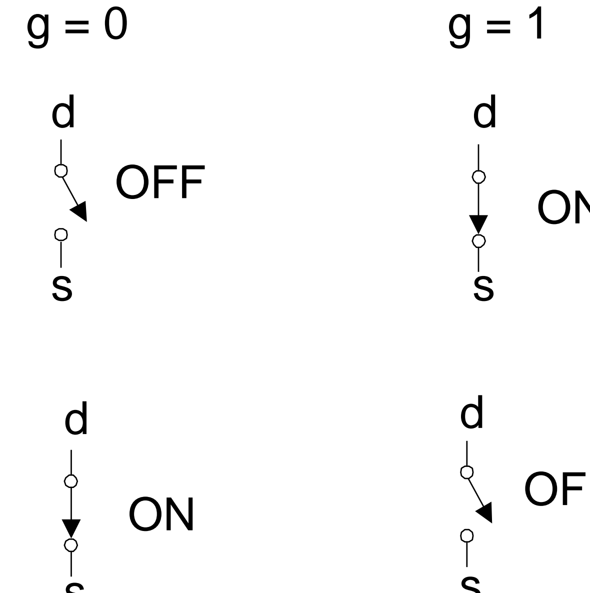
University of Jordan

Transistors as Switches

- We can view MOS transistors as electrically controlled switches
- Voltage at gate controls path from source to drain

nMOS



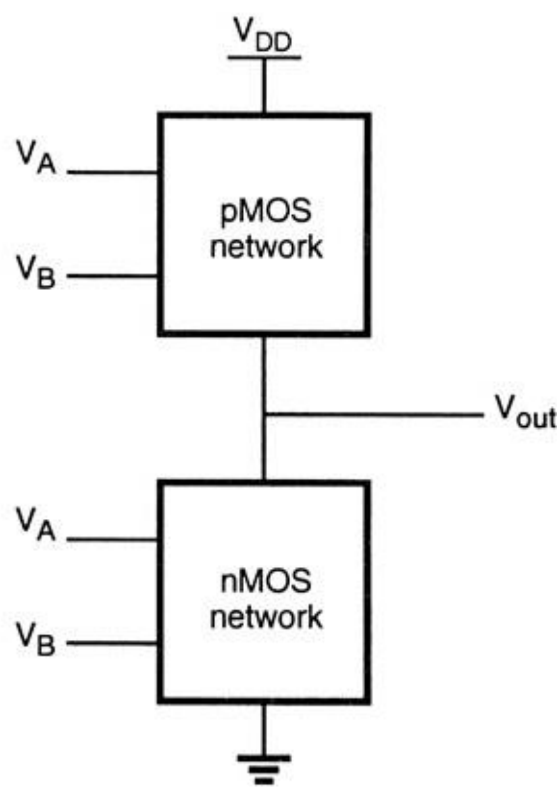


pMOS

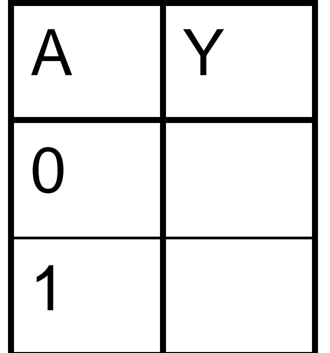
d
g
s

CMOS Circuits

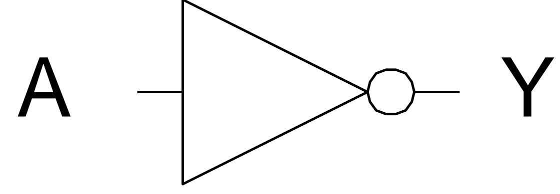
- For the gate to output a ‘1’
  - Some path of PMOS transistors from VDD to outp
    on
  - We call the PMOS transistors the Pull-Up Netwo
- For the gate to output a ‘0
  - Some path of NMOS transistors from GND to ou
    on
  - We call the NMOS transistors the Pull-Down Net

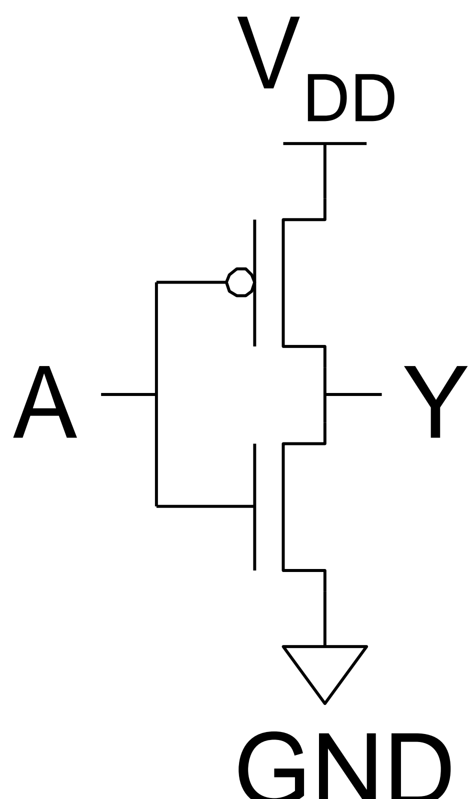


CMOS Inverter

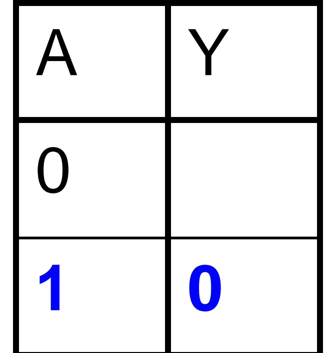


A	Y
0	
1	

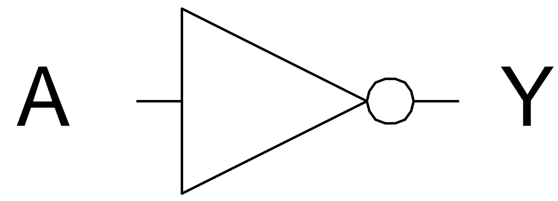


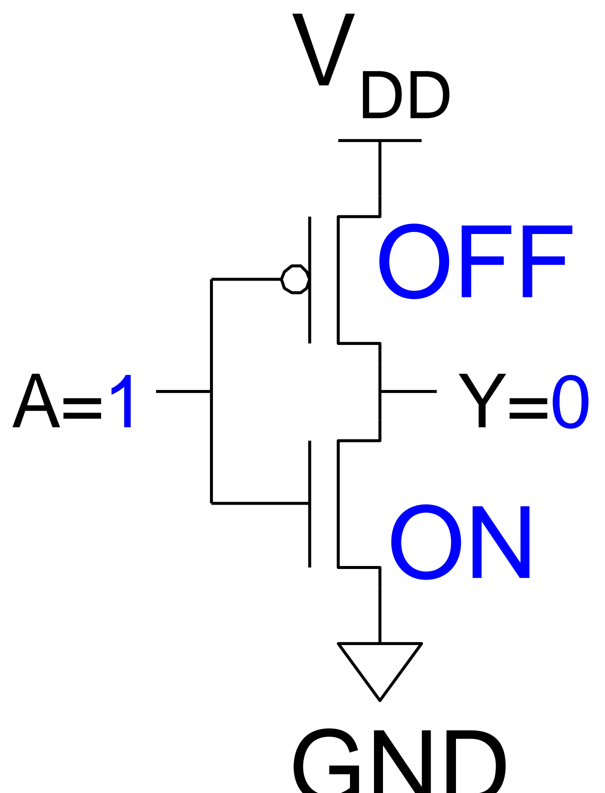


CMOS Inverter

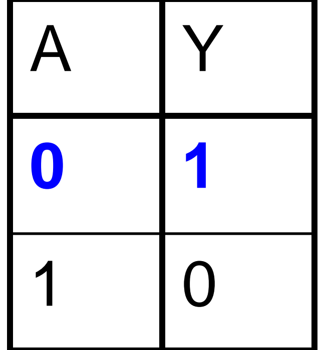


A	Y
0	
1	**0**

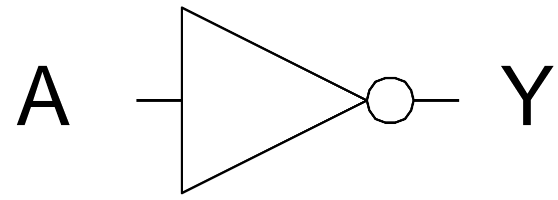


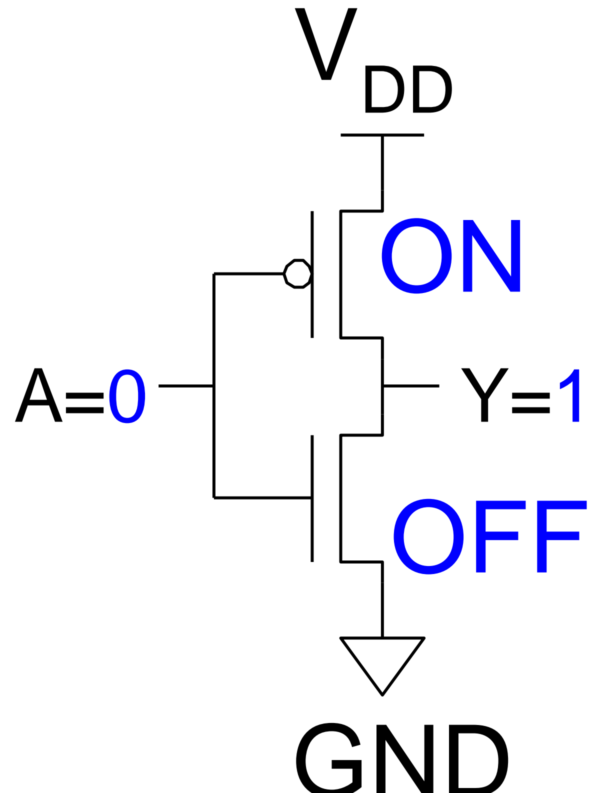


CMOS Inverter

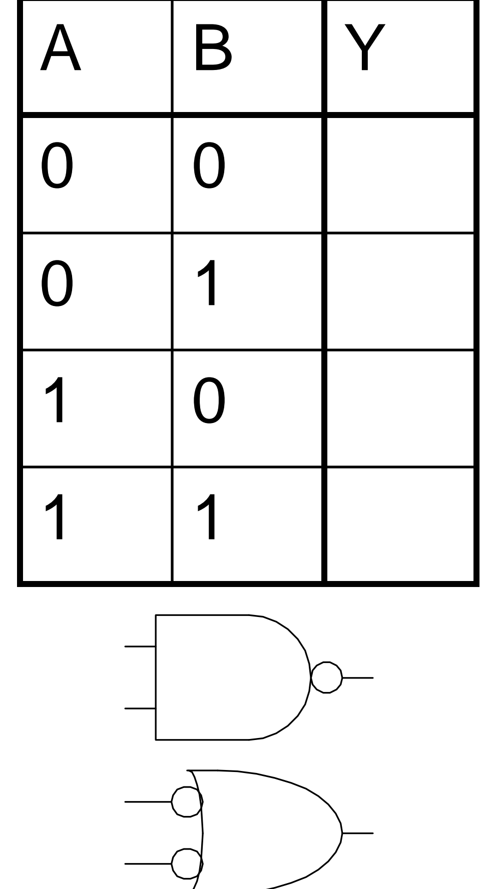


A	Y
0	**1**
1	0

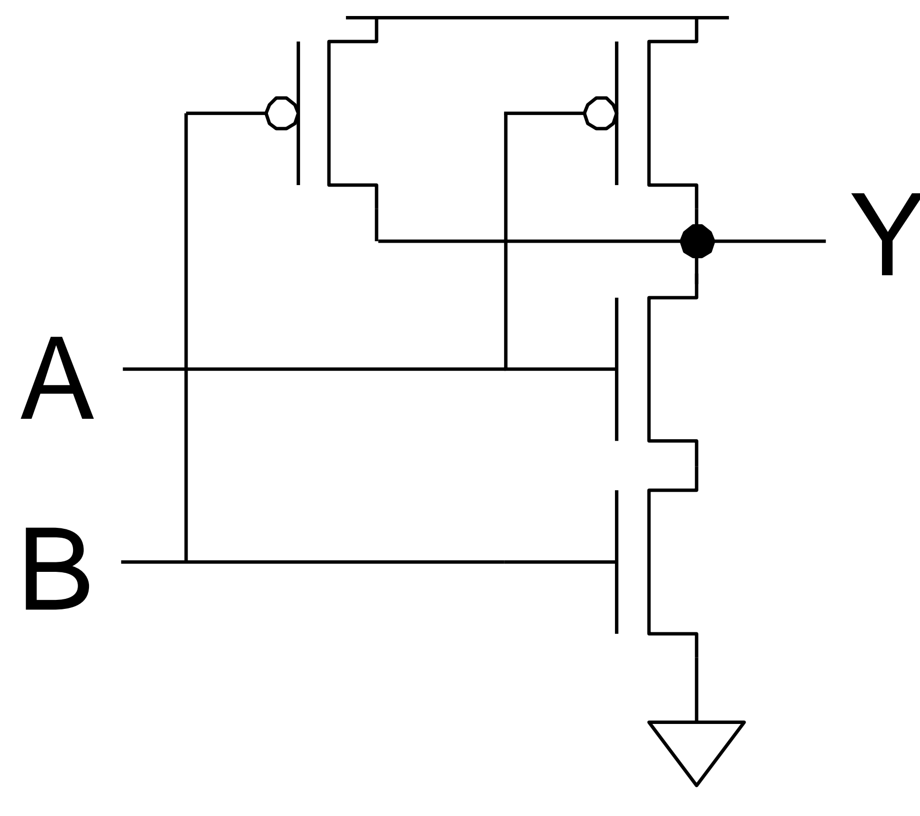




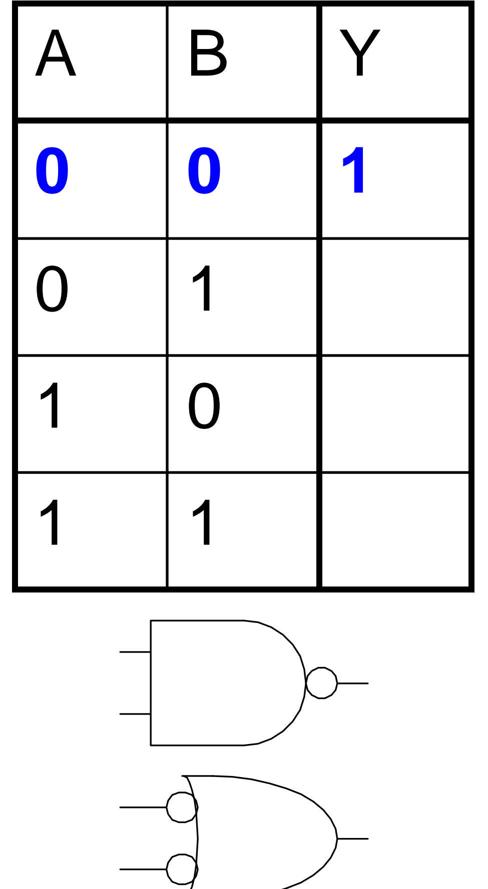
CMOS NAND Gate



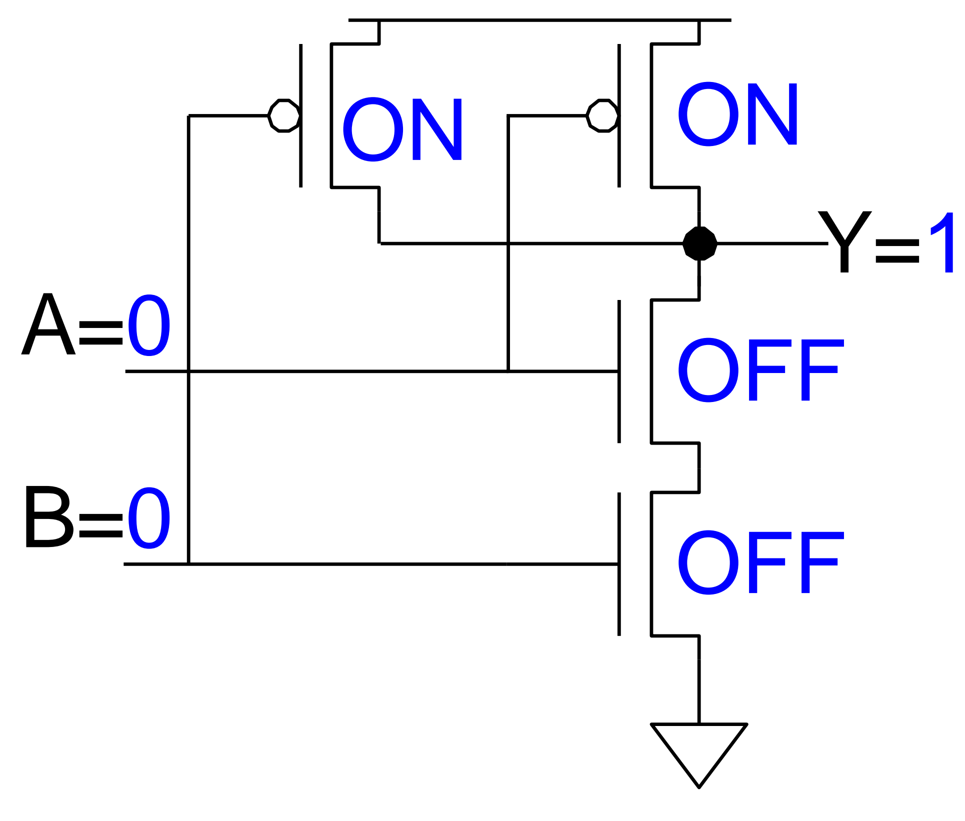
A	B	Y
0	0	
0	1	
1	0	
1	1	



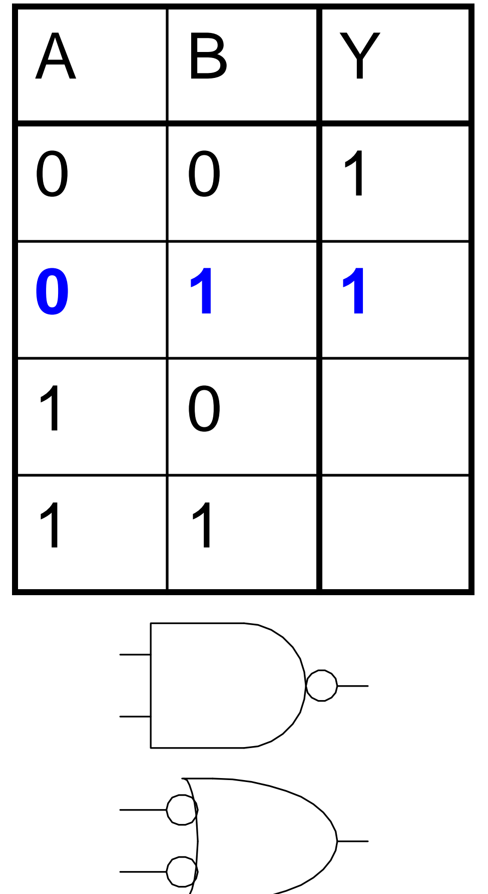
CMOS NAND Gate



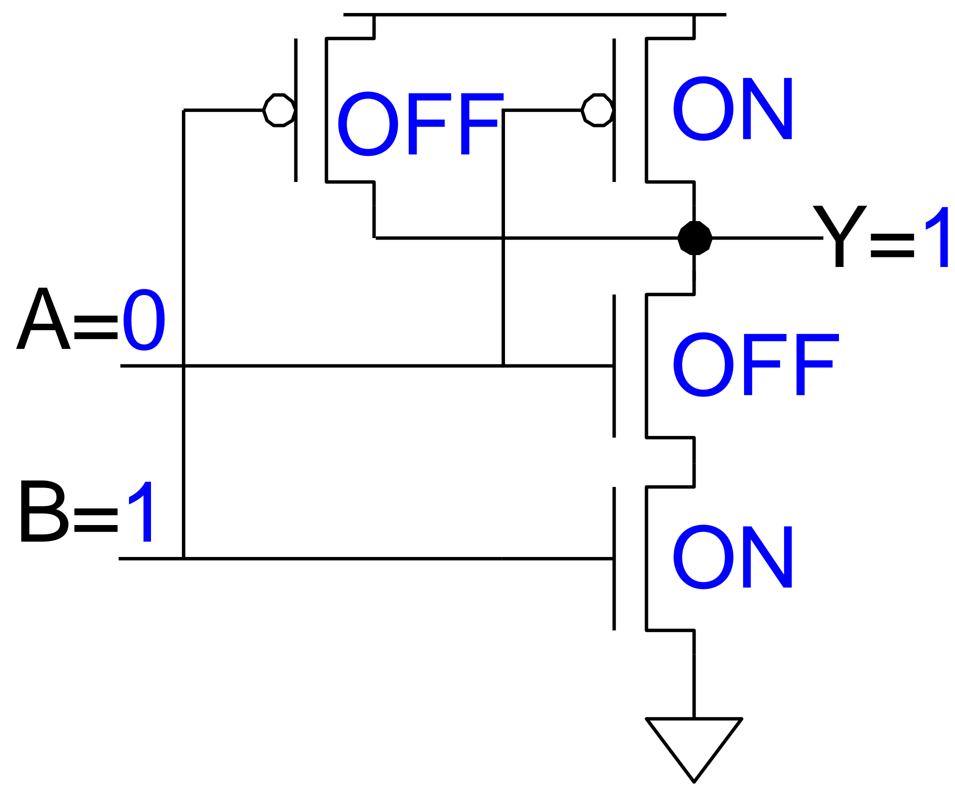
A	B	Y
0	**0**	**1**
0	1	
1	0	
1	1	



CMOS NAND Gate

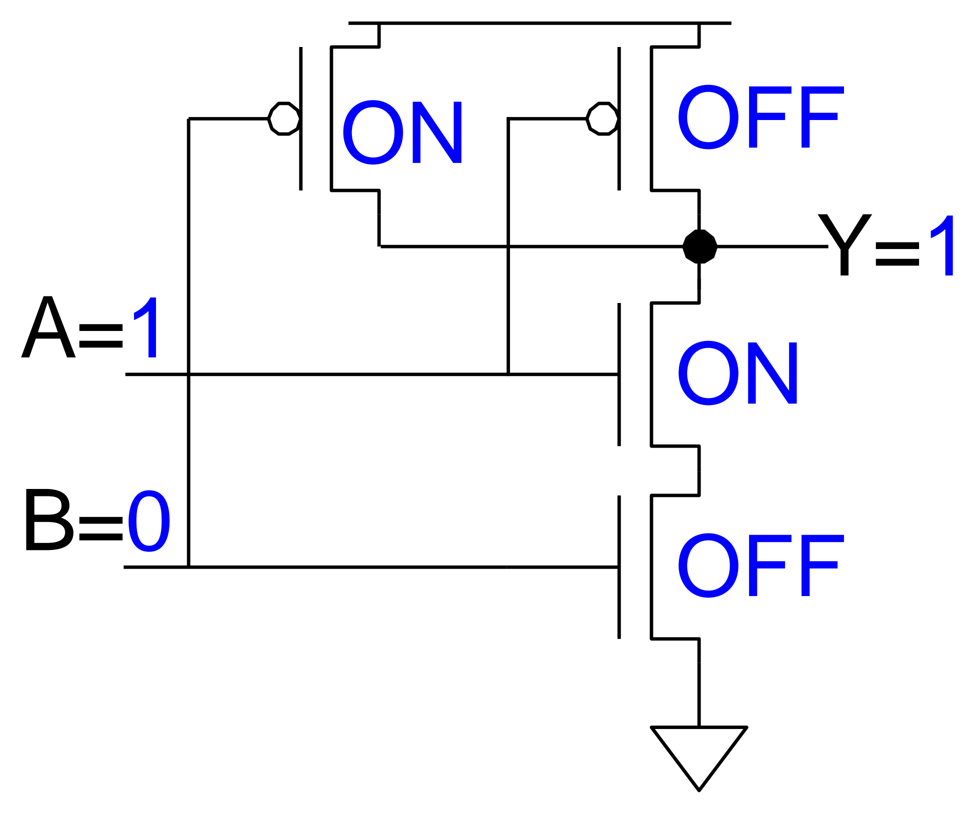


A	B	Y
0	0	1
0	**1**	**1**
1	0	
1	1	



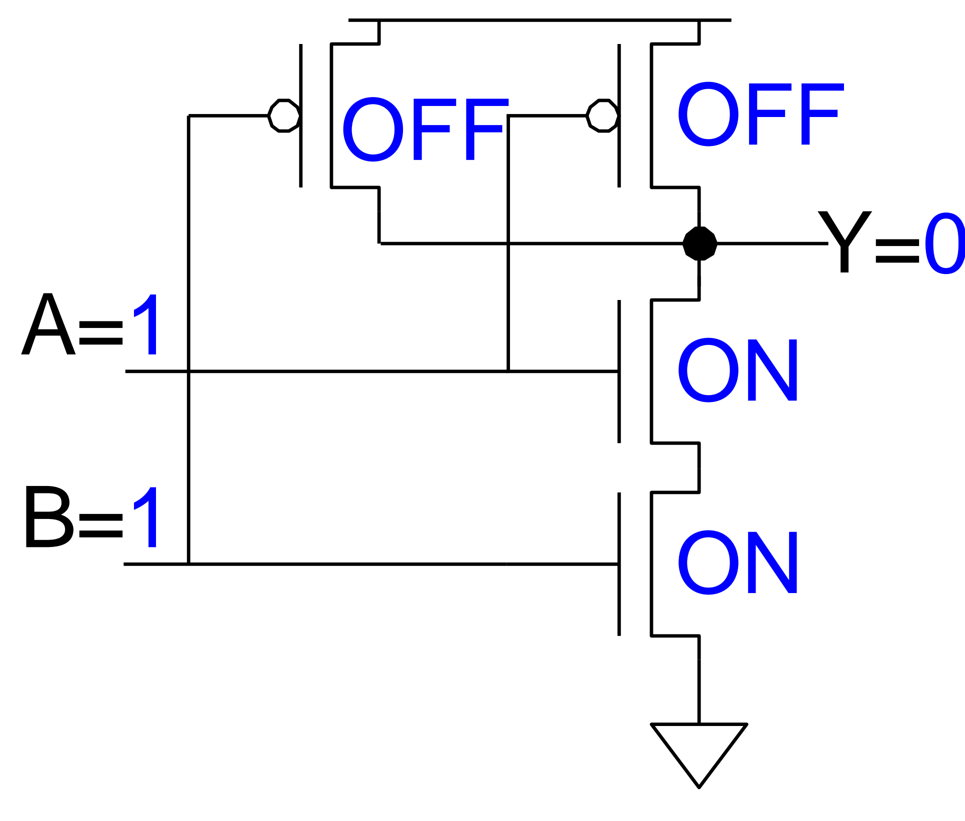
CMOS NAND Gate

A	B	Y
0	0	1
0	1	1
1	**0**	**1**
1	1	



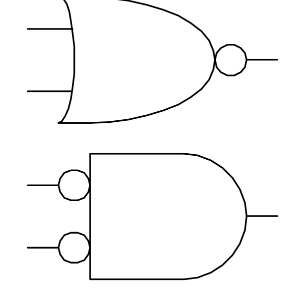
CMOS NAND Gate

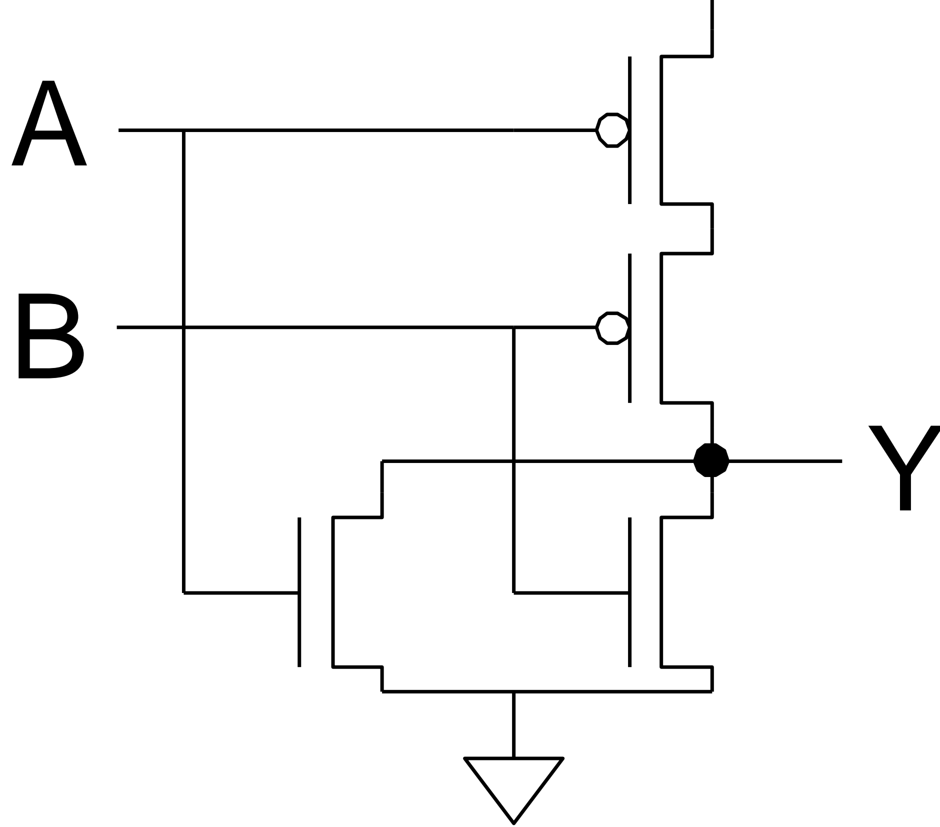
A	B	Y
0	0	1
0	1	1
1	0	1
1	**1**	**0**



CMOS NOR Gate

A	B	Y
0	0	1
0	1	0
1	0	0
1	1	0



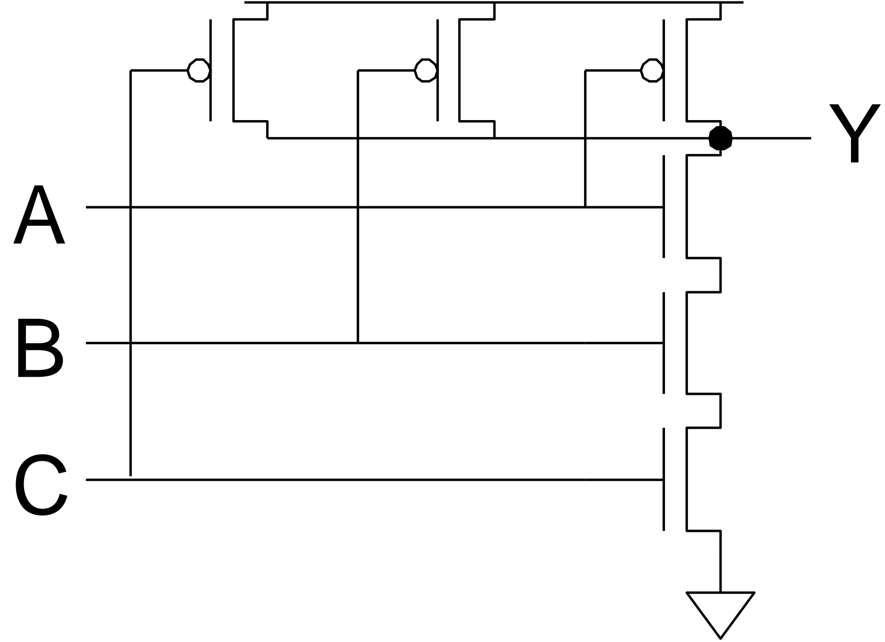


3-input NAND Gate

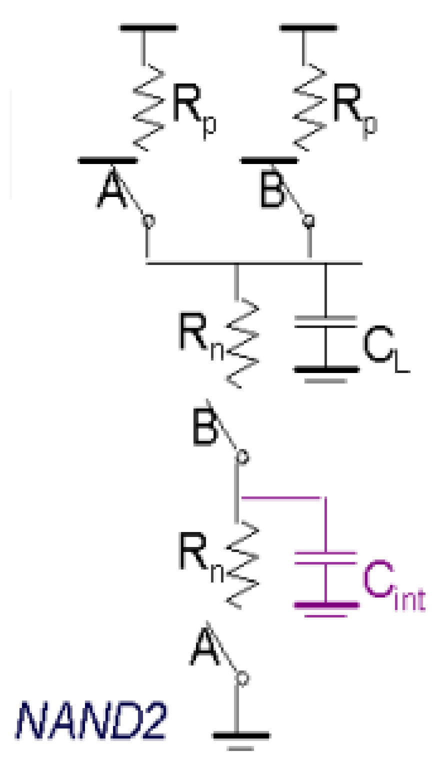
- Y pulls low if ALL inputs are 1
- Y pulls high if ANY input is 0

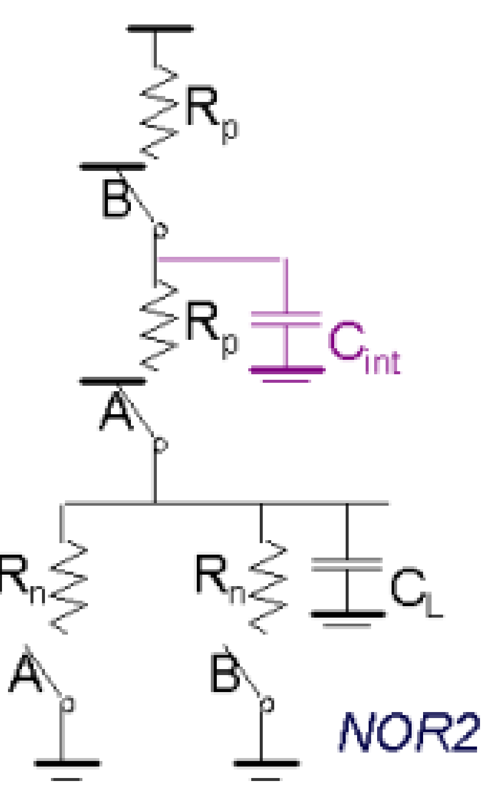
3-input NAND Gate

- Y pulls low if ALL inputs are 1
- Y pulls high if ANY input is 0



Sizing



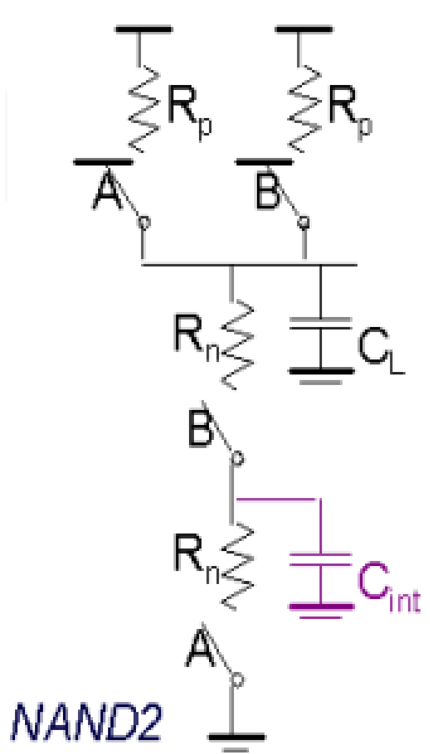


Sizing

- Delay is dependent on the pattern of inputs
- The ratio of the {W/L}PUN/ {W/L}PDN should be two (or higher) to make up for slow PMOS

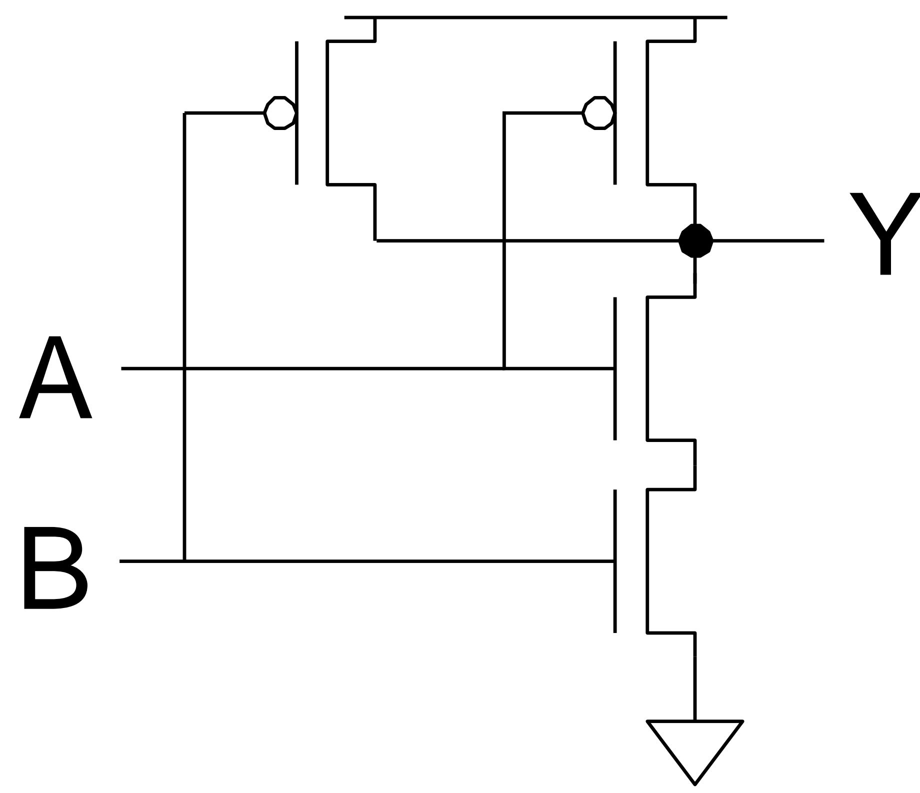
Sizing

- Delay is dependent on the pattern of inputs
- Low to high transition
  - both inputs go low
  - delay is $(R_p/2)*C_L$
- one input goes low
  - delay is $(R_p)*C_L$
    → Worst case: R_p*C_L
- High to Low transition
  - both inputs go high
  - delay is $(2*R_n)*C_L$

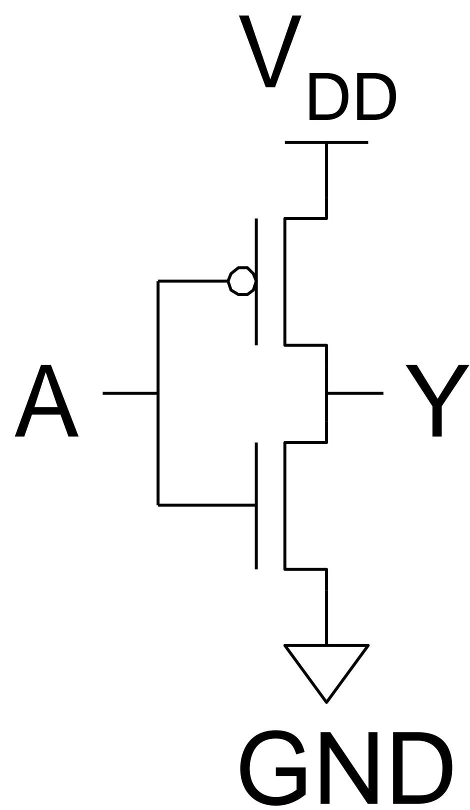


R_p=2R_n → Wp= Wn

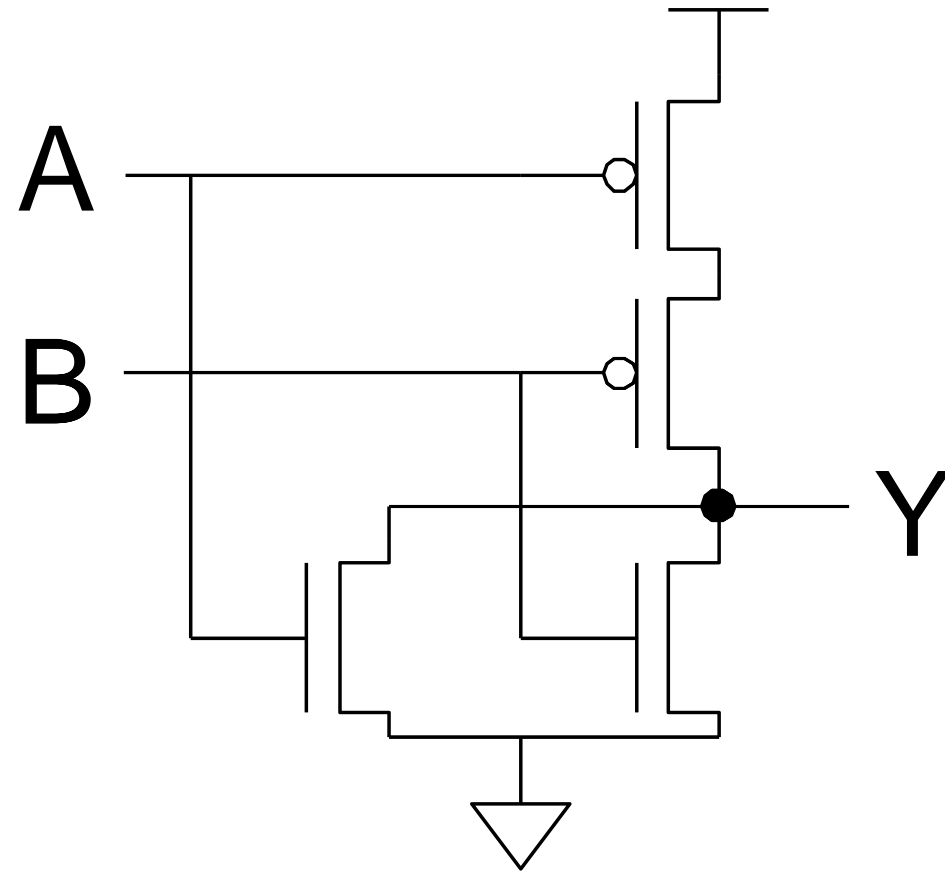
NAND Sizing



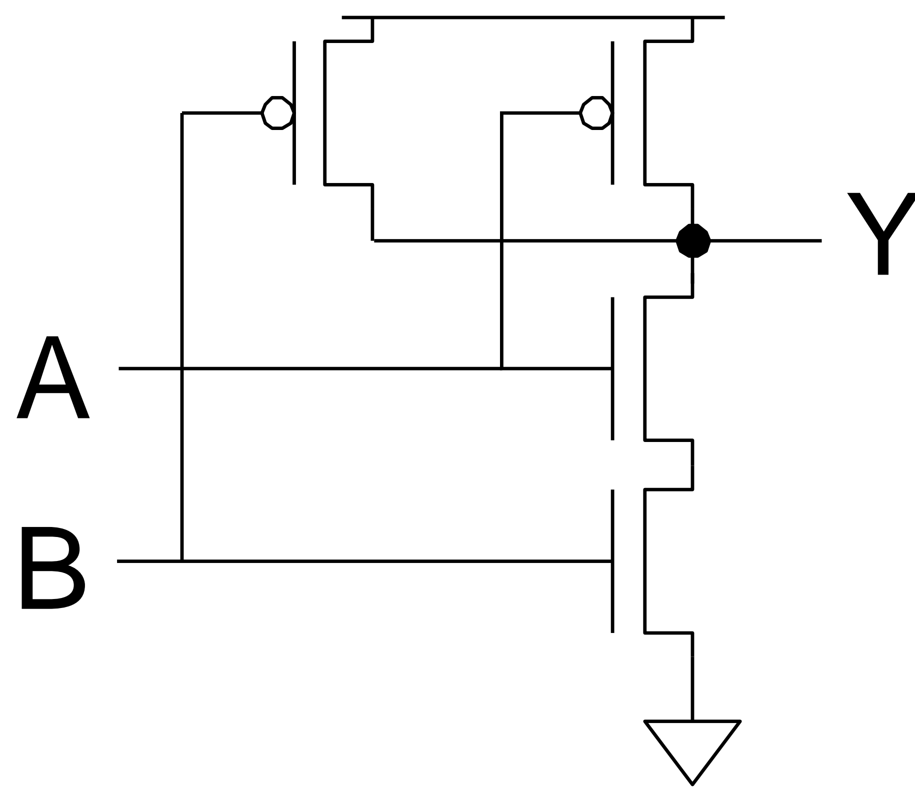
Inverter Sizing



NOR Sizing



NAND capacitance



Complex Gates

- $A = \overline{D + A.(B + C)}$

Complex Gates Sizing

- $A = \overline{D + A.(B + C)}$

Stick Diagrams

- VLSI design aims to translate circuit concepts onto silicon.
- stick diagrams are a means of capturing topography and layer information using simple diagrams.
- Stick diagrams convey layer information through colour codes (or monochrome encoding).
- Acts as an interface between symbolic circuit and the actual layout.

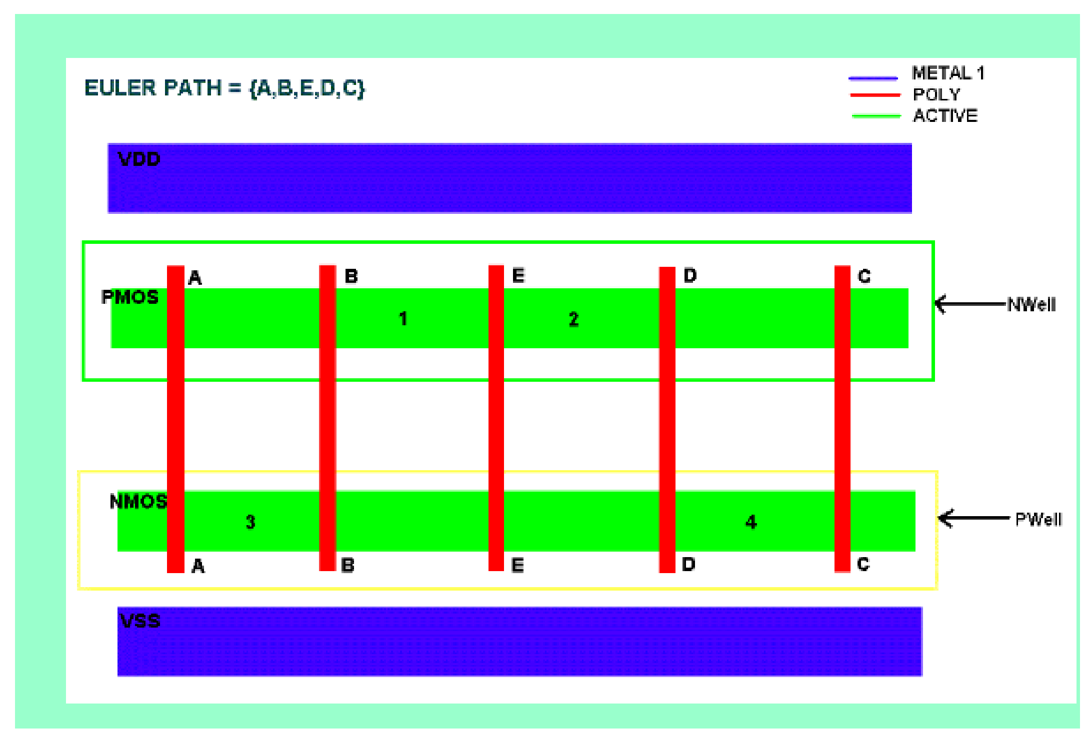
Stick Diagrams

- Does show all components/vias.
- It shows relative placement of components.
- Goes one step closer to the layout
- Helps plan the layout and routing

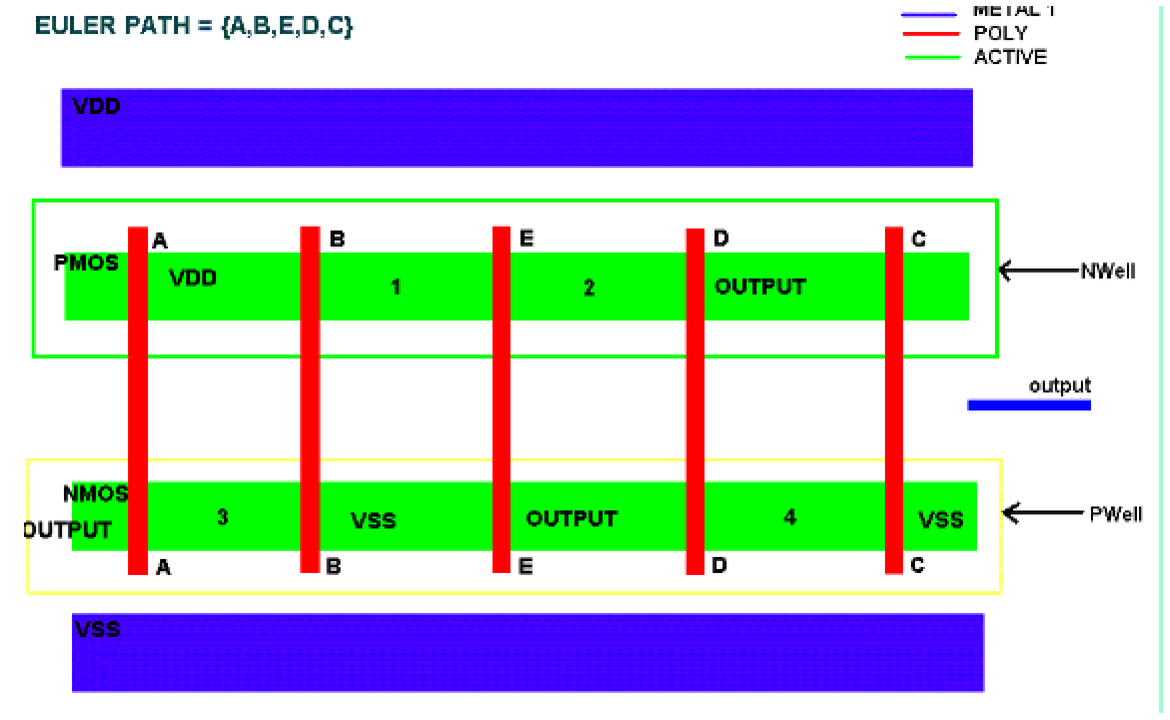
Stick Diagrams

- Does ***not*** show
  - Exact placement of components
  - Transistor sizes
  - Wire lengths, wire widths, tub boundaries.
  - Any other low level details such as parasitics..

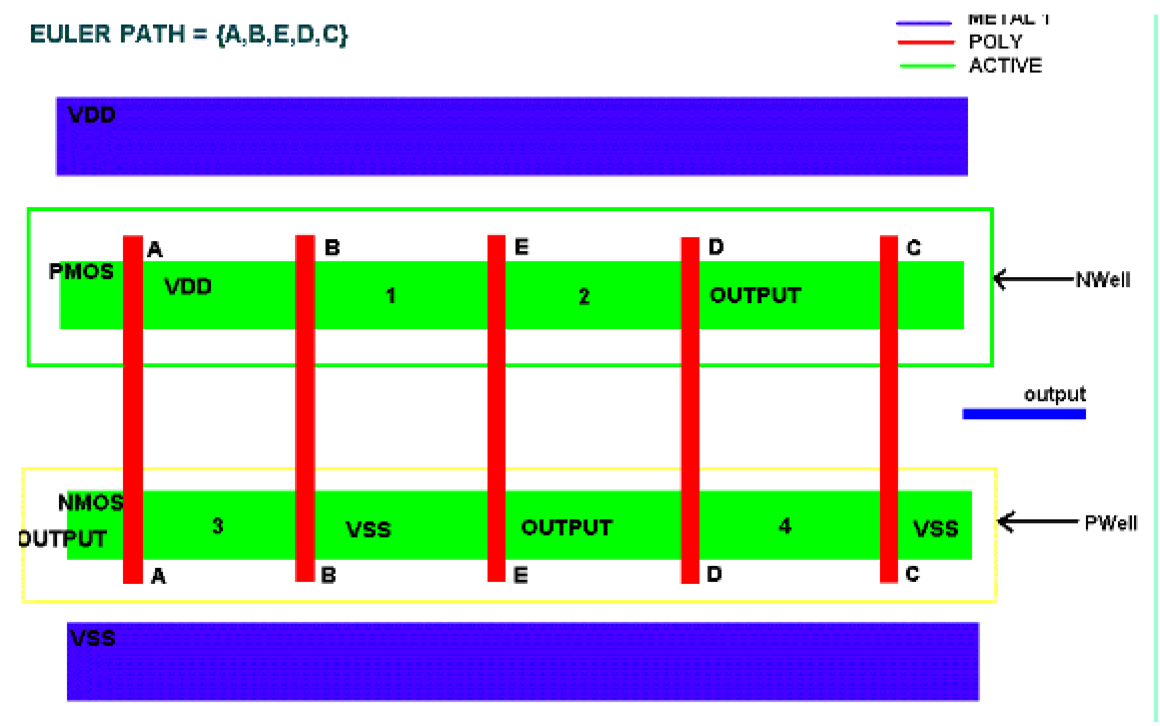
Stick Diagram



THE UNIVERSITY OF
JORDAN


EULER PATH = {A,B,E,D,C}
POLY
ACTIVE
VDD
PMOS
A
B
E
D
C
VDD
1
2
OUTPUT
NWell
output
NMOS
OUTPUT
3
VSS
OUTPUT
4
VSS
PWell
A
B
E
D
C
VSS

Stick Diagram



Example

- $\mathrm{F} = \overline{A + B}$

Example

- $F = \overline{A.B}$

Example

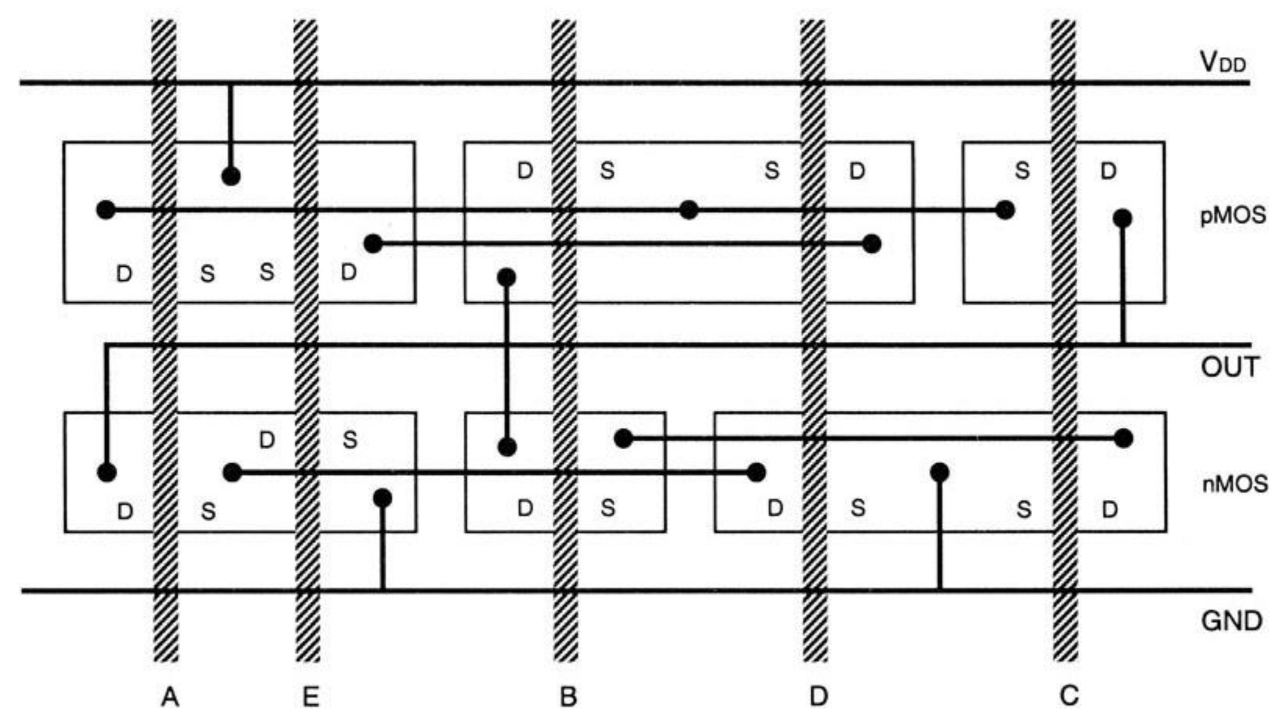
- $F = \overline{A + BC}$

Example

- $\mathrm{F} = \overline{A(D+E)+BC}$

THE UNIVERSITY OF
JORDAN

THE UNIVERSITY OF
JORDAN


VDD
pMOS
OUT
nMOS
GND
D
S
A
E
B
D
C

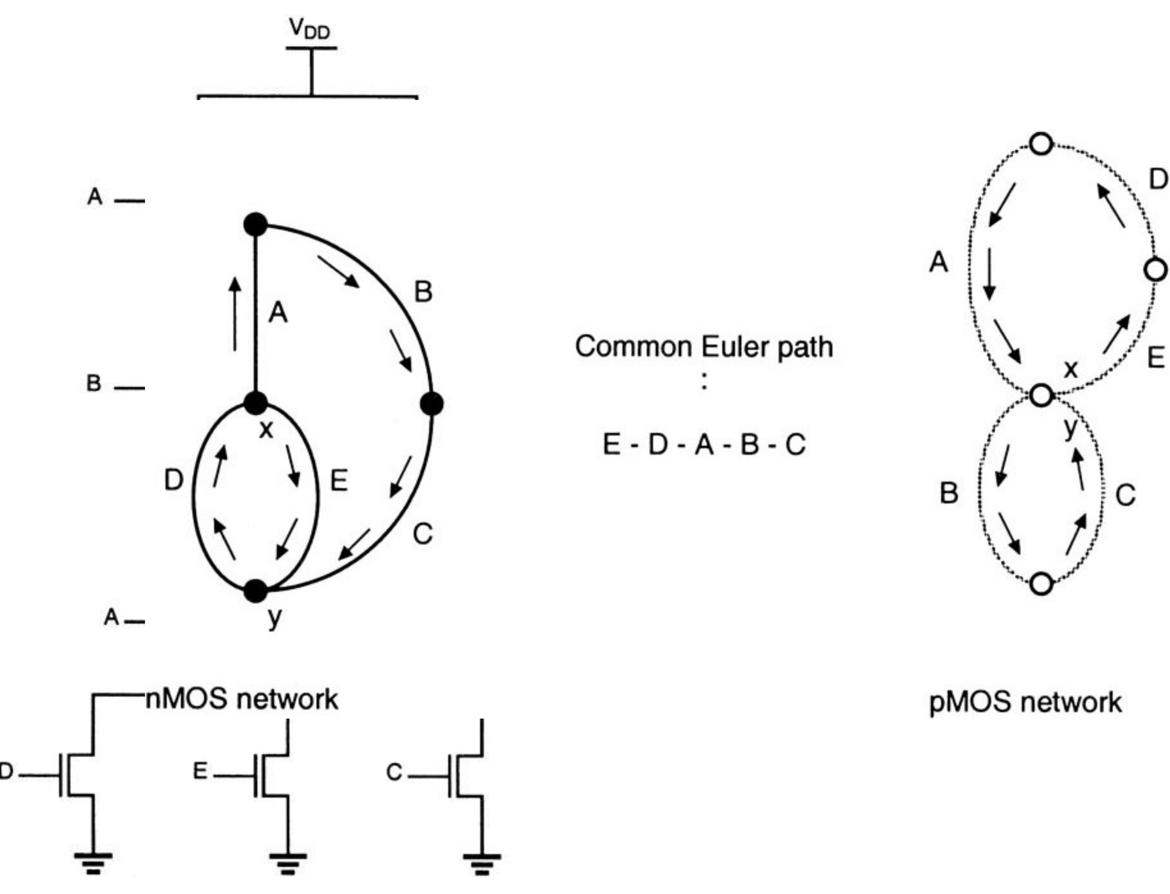
Euler Path

- The number of diffusion breaks can be minimized by changing the ordering of the polysilicon columns
- A simple method for finding the optimum gate ordering is the **Euler-path** approach
- Find a common Euler path for both pull-down and pull-up graphs
- The polysilicon columns can be arranged according to the sequence (in Euler-path)
- Diffusion will be unbroken if identically labeled Euler paths can be found for the p and n trees

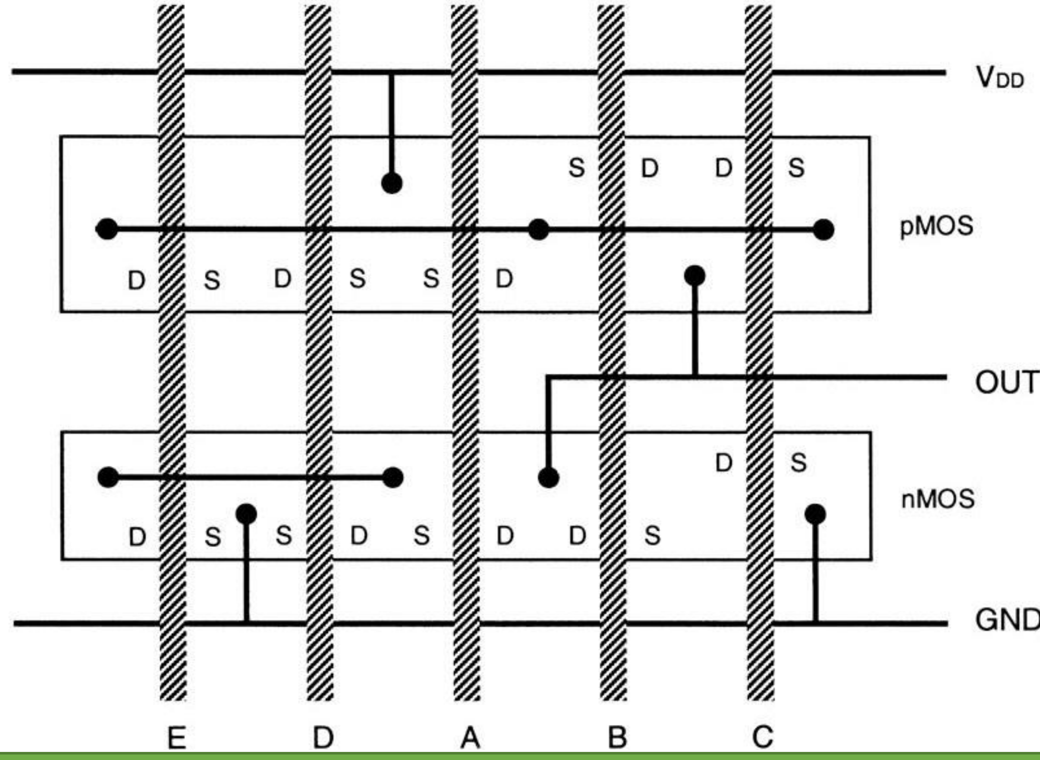
Euler Path

- construct one Euler path for both the Pull up and Pull down network
  - Path the traverses each node in the path, such that each edge is visited only once.
  - If the path traverses transistor A then B then C. Then the path name is {A, B, C}
  - The Euler path of the Pull up network must be the same as the path of the Pull down network.
  - Euler paths are not necessarily unique.
  - It may be necessary to redefine the function to find a Euler path.

Euler Path



- The advantages of this new layout are more compact layout area, simple routing of signals, and consequently, lower parasitic capacitance



Euler Path

- $A = \overline{D + AB + CE}$

Example

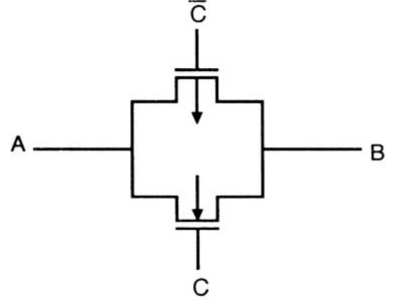
- $A = \overline{D + AB + CE}$

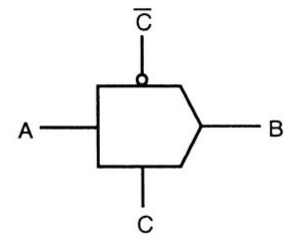
Example

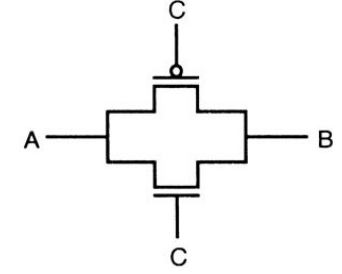
- f= $\overline{A(D+E)+BC}$

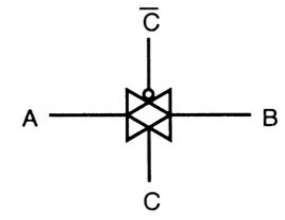
CMOS Transmission Gates (Pass Gates)

- Representations of the CMOS transmission gate (TG)



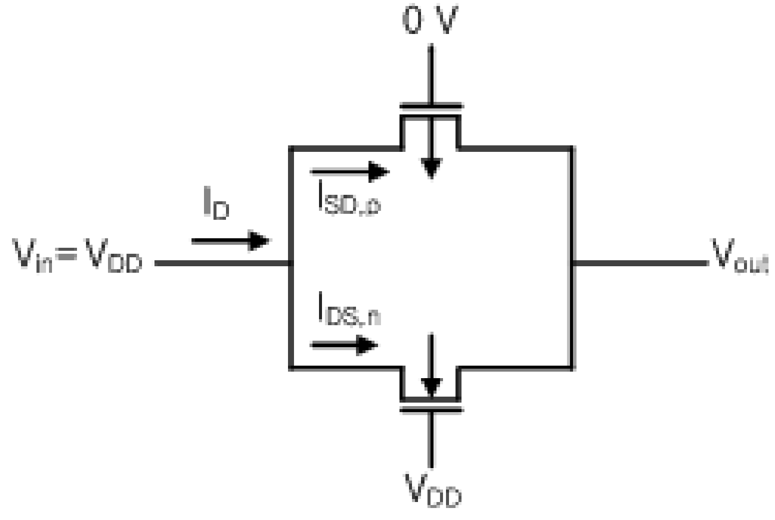


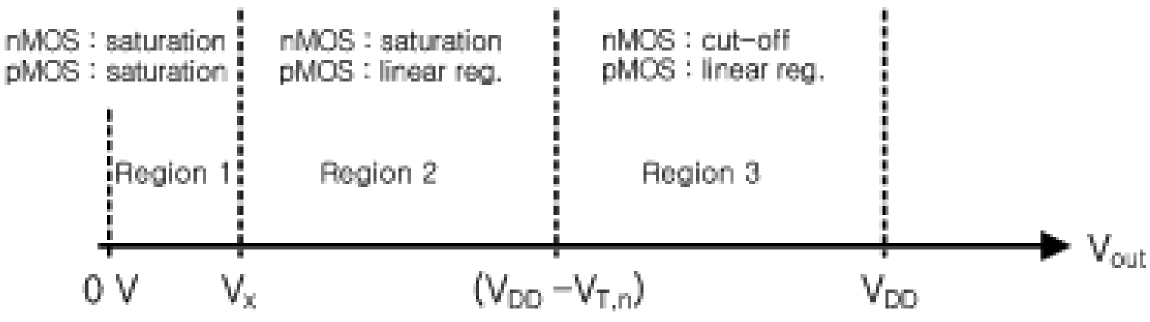




- One nMOS + one pMOS
- Bidirectional switch
- If C=V_{DD}, TG is turned on (low-resistance path)
- If C=GND, TG is turned off (high-impedance state)

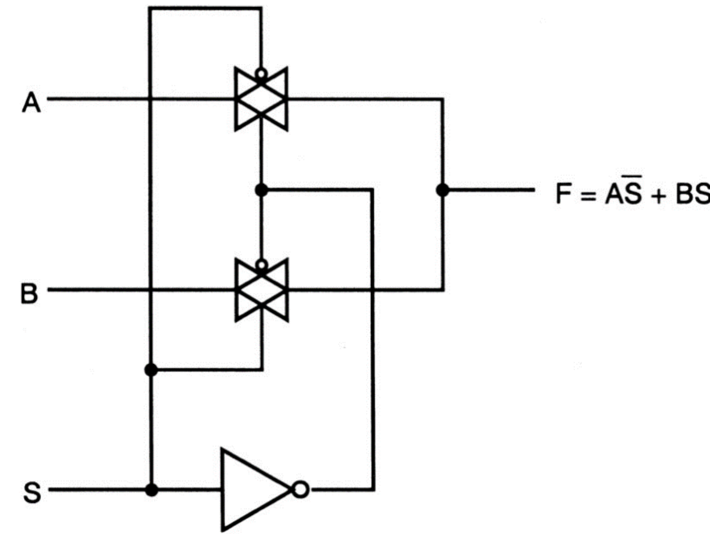
DC Analysis of Transmission Gate (1)



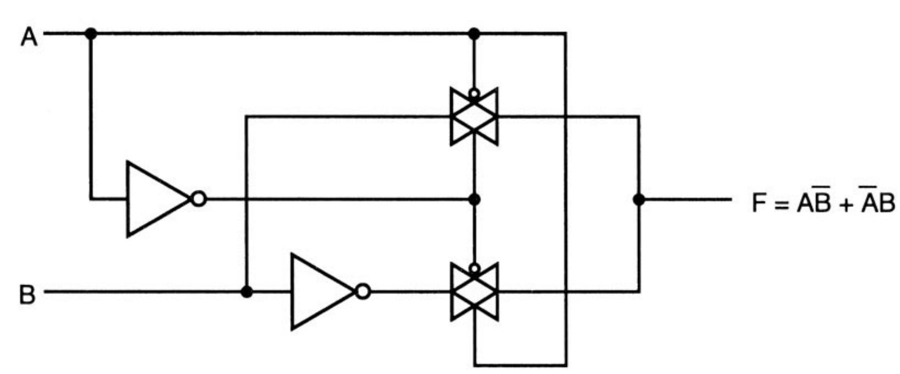


CMOS TG Implementations (1)

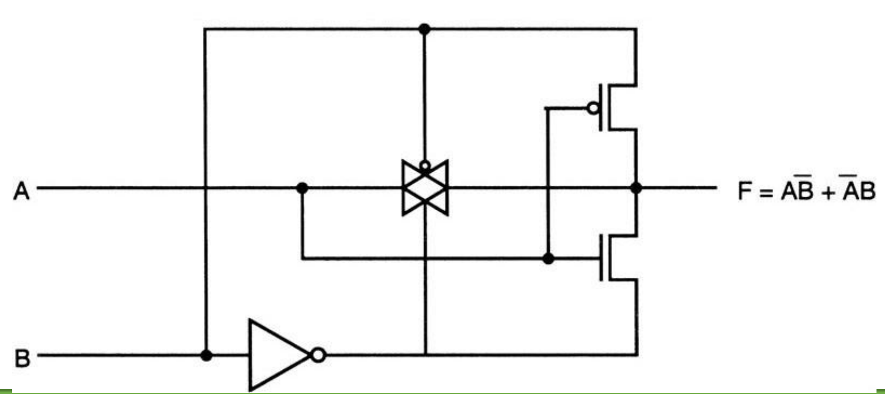
- Two-input multiplexor



- 8-TRs XOR function

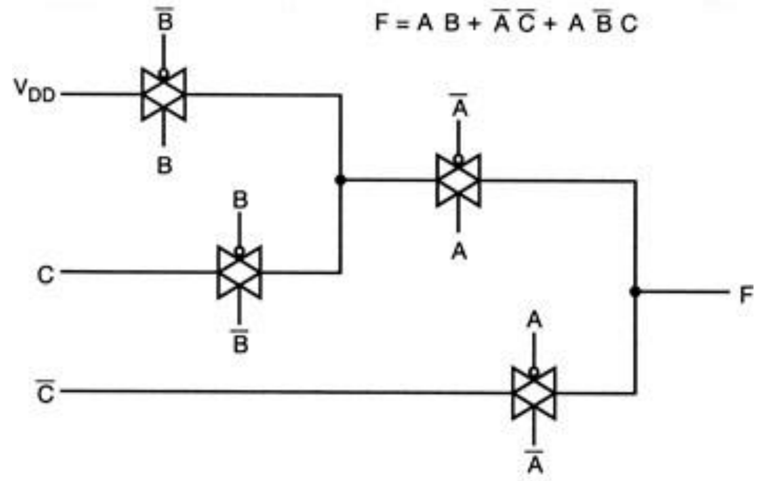


- 6-TRs XOR function



CMOS TG Implementations (2)

- Boolean function realization



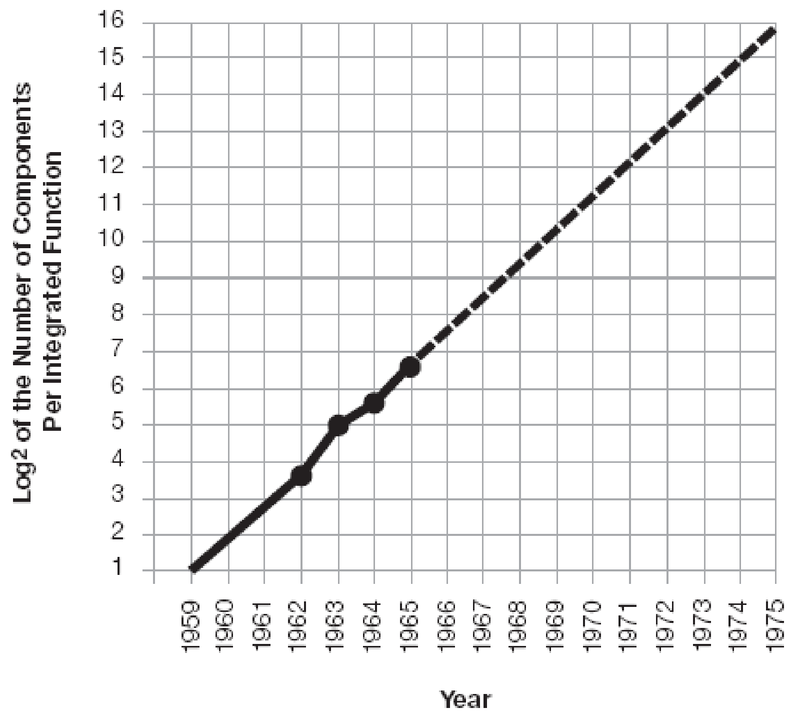
Transitors scaling

Outline

- Scaling
  - Transistors
  - Interconnect
  - Future Challenges

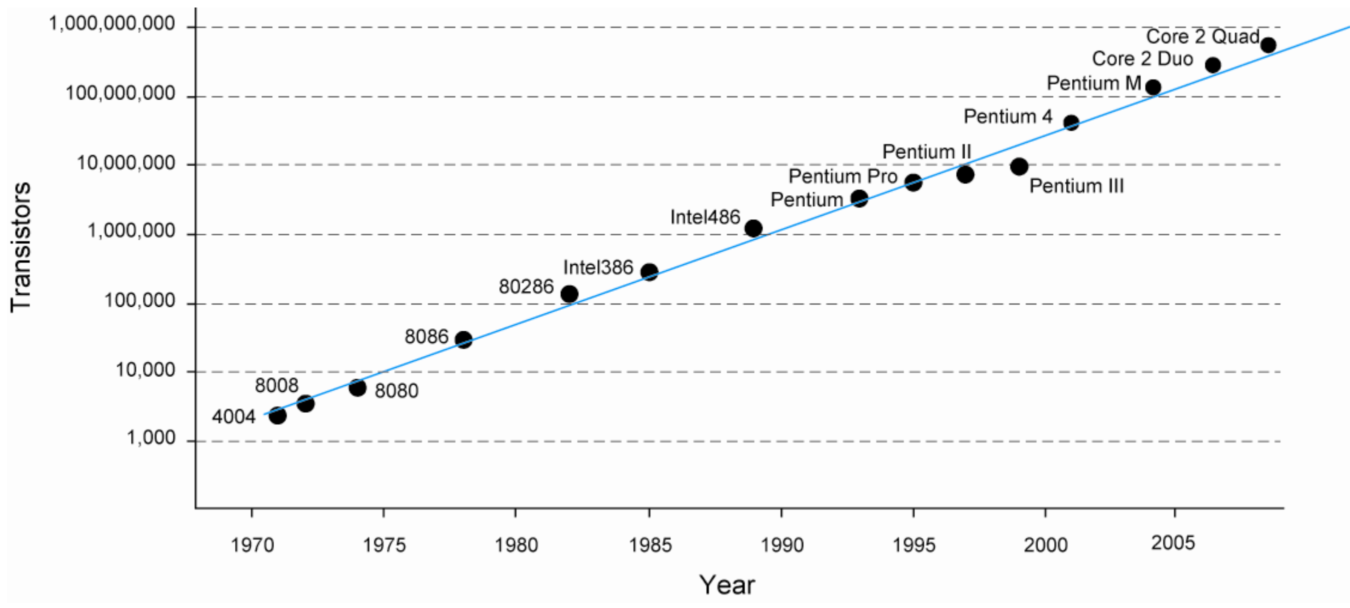
Moore’s Law

- Recall that Moore’s Law has been driving CMOS



Moore’s Law

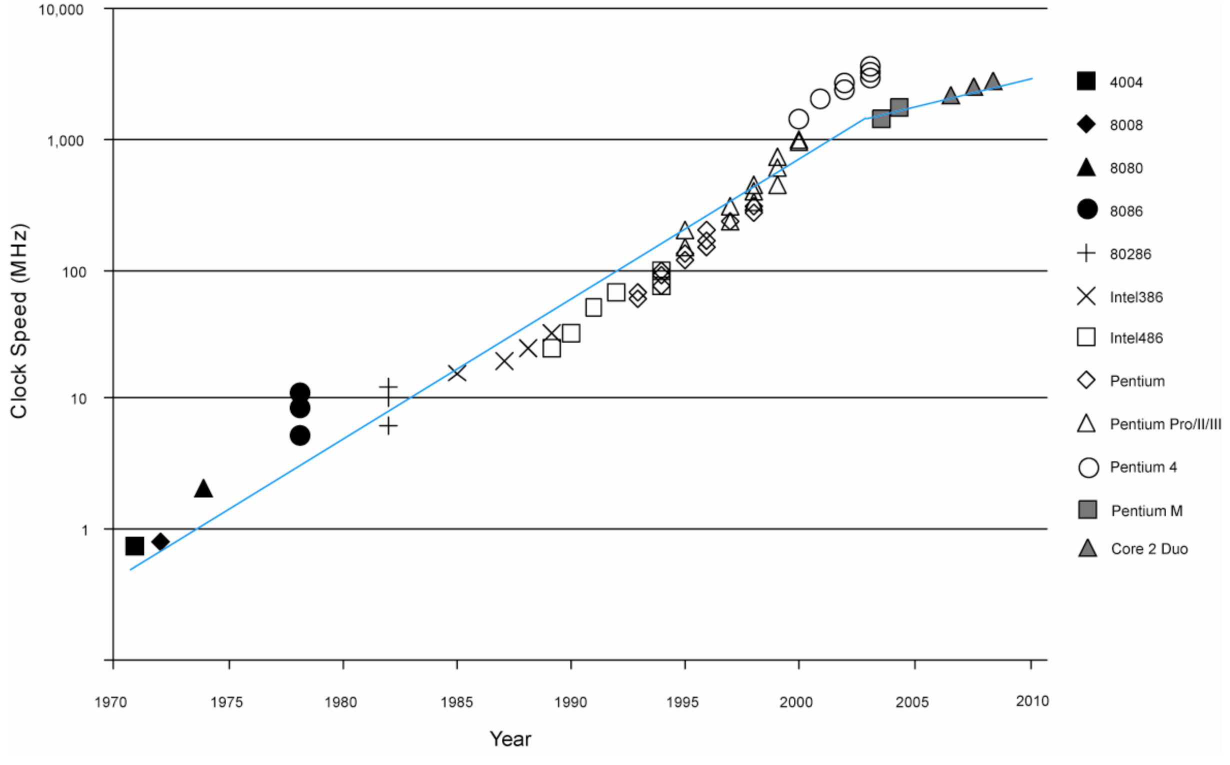
- Recall that Moore’s Law has been driving CMOS



Moore’s Law today

Moore’s Law

- Recall that Moore’s Law has been driving CMOS



Corollary: clock speeds have improved

Why?

- Why more transistors per IC?
  - Smaller transistors
  - Larger dice
- Why faster computers?
  - Smaller, faster transistors
  - Better microarchitecture (more IPC)
  - Fewer gate delays per cycle

Scaling

- The only constant in VLSI is constant change
- Feature size shrinks by 30% every 2-3 years
  - Transistors become cheaper
  - Transistors become faster and lower power
  - Wires do not improve
    (and may get worse)
- Scale factor S
  - Typically $S = \sqrt{2}$
  - Technology nodes

Dennard Scaling

- Proposed by Dennard in 1974
- Also known as *constant field* scaling
  - Electric fields remain the same as features scale
- Scaling assumptions
  - All dimensions (x, y, z => W, L, t_{ox})
  - Voltage (V_{DD})
  - Doping levels

Device Scaling

Parameter	Sensitivity	Dennard Scaling
L: Length		1/S
W: Width		1/S
t_{ox}: gate oxide thickness		1/S
V_{DD}: supply voltage		1/S
V_t: threshold voltage		1/S
NA: substrate doping		S
β	$W/(Lt_{ox})$	S
I_{on}: ON current	$\beta(V_{DD}-V_t)^2$	1/S
R: effective resistance	V_{DD}/I_{on}	1
C: gate capacitance	WL/t_{ox}	1/S
τ: gate delay	RC	1/S
f: clock frequency	$1/\tau$	S
E: switching energy / gate	CV_{DD}^2	$1/S^3$
P: switching power / gate	Ef	$1/S^2$
A: area per gate	WL	$1/S^2$
Switching power density	P/A	1
Switching current density	I_{on}/A	S

Observations

- Gate capacitance per micron is nearly independent of process
- Gates get faster with scaling (good)
- Dynamic power goes down with scaling (good)
- Current density goes up with scaling (bad)

Real Scaling

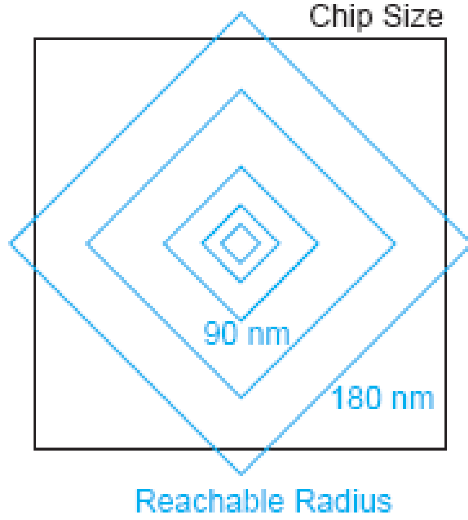
- t_{ox} scaling has slowed since 65 nm
  - Limited by gate tunneling current
  - Gates are only about 4 atomic layers thick!
  - High-k dielectrics have helped continued scaling of effective oxide thickness
- V_{DD} scaling has slowed since 65 nm
  - SRAM cell stability at low voltage is challenging
- Dennard scaling predicts cost, speed, power all improve
  - Below 65 nm, some designers find they must choose just two of the three

Interconnect Scaling

Parameter	Sensitivity	Scale Factor
w: width		1/S
s: spacing		1/S
t: thickness		1/S
h: height		1/S
D_c: die size		D_c
R_w: wire resistance/unit length	1/wt	S^2
C_{wf}: fringing capacitance / unit length	t/s	1
C_{wp}: parallel plate capacitance / unit length	w/h	1
C_w: total wire capacitance / unit length	$C_{wf} + C_{wp}$	1
t_{wu}: unrepeated RC delay / unit length	R_wC_w	S^2
t_{wr}: repeated RC delay / unit length	$\mathrm{sqrt}(RCR_wC_w)$	sqrt(S)
Crosstalk noise	w/h	1
E_w: energy per bit / unit length	$C_wV_{DD}^2$	$1/S^2$

Reachable Radius

- We can't send a signal across a large fast chip in one cycle anymore
- But the microarchitect can plan around this
  - Just as off-chip memory latencies were tolerated



Observations

- Local wires are getting faster
  - Not quite tracking transistor improvement
  - But not a major problem
- Global wires are getting slower
  - No longer possible to cross chip in one cycle

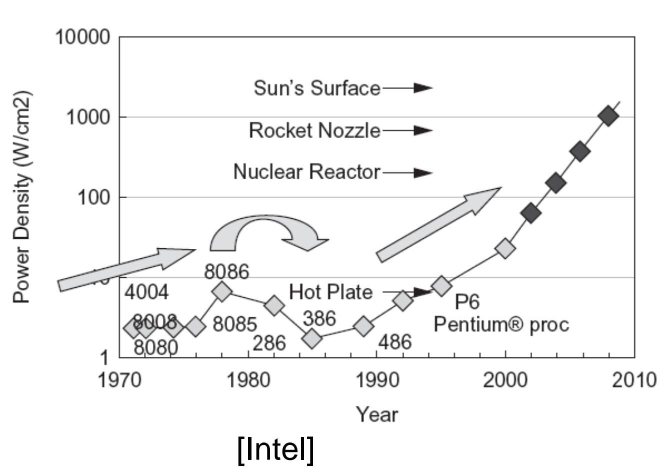
ITRS

- Semiconductor Industry Association forecast
  - Intl. Technology Roadmap for Semiconductors

Year	2009	2012	2015	2018	2021
Feature size (nm)	34	24	17	12	8.4
L_{gate} (nm)	20	14	10	7	5
V_{DD} (V)	1.0	0.9	0.8	0.7	0.65
Billions of transistors/die	1.5	3.1	6.2	12.4	24.7
Wiring levels	12	12	13	14	15
Maximum power (W)	198	198	198	198	198
DRAM capacity (Gb)	2	4	8	16	32
Flash capacity (Gb)	16	32	64	128	256

Dynamic Power

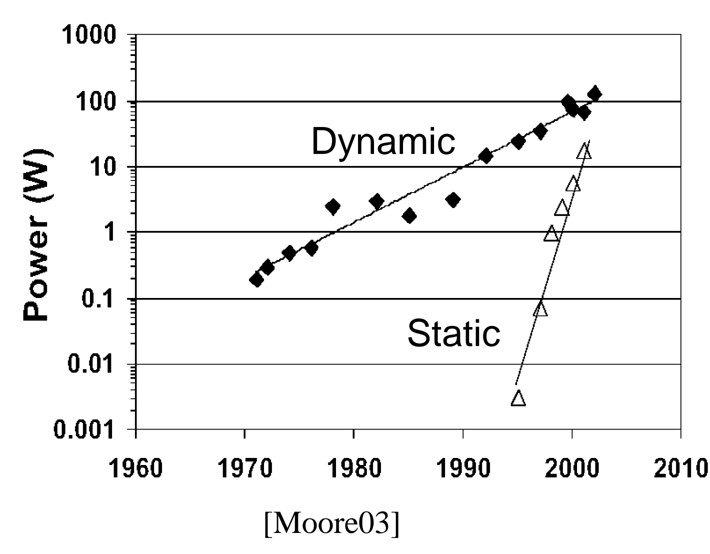
- Intel VP Patrick Gelsinger (ISSCC 2001)
  - If scaling continues at present pace, by 2005, high speed processors would have power density of nuclear reactor, by 2010, a rocket nozzle, and by 2015, surface of sun.
  - “Business as usual will not work in the future.”
- Attention to power is increasing



[Intel]

Static Power

- V_{DD} decreases
  - Save dynamic power
  - Protect thin gate oxides and short channels
  - No point in high value because of velocity sat.
- V_t must decrease to maintain device performance
- But this causes exponential increase in OFF leakage
- Major future challenge



[Moore03]

Productivity

- Transistor count is increasing faster than designer productivity (gates / week)
  - Bigger design teams
    - Up to 500 for a high-end microprocessor
  - More expensive design cost
  - Pressure to raise productivity
    - Rely on synthesis, IP blocks
  - Need for good engineering managers

Physical Limits

- Will Moore’s Law run out of steam?
  - Can’t build transistors smaller than an atom…
- Many reasons have been predicted for end of scaling
  - Dynamic power
  - Subthreshold leakage, tunneling
  - Short channel effects
  - Fabrication costs
  - Electromigration
  - Interconnect delay

Sequential Logic

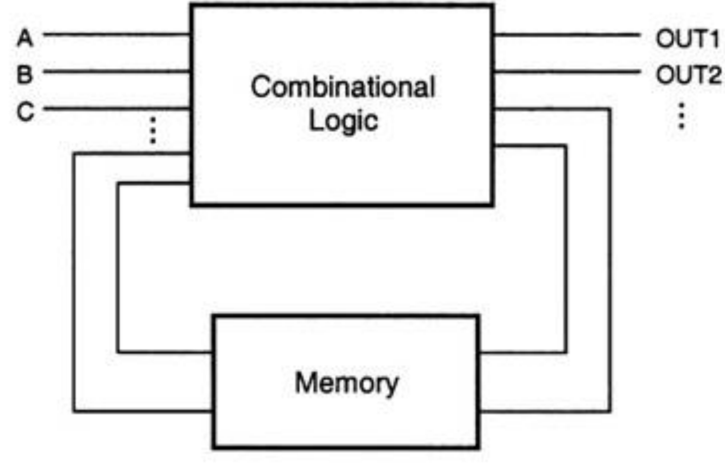
Dr.Mohammad Abdel-Majeed

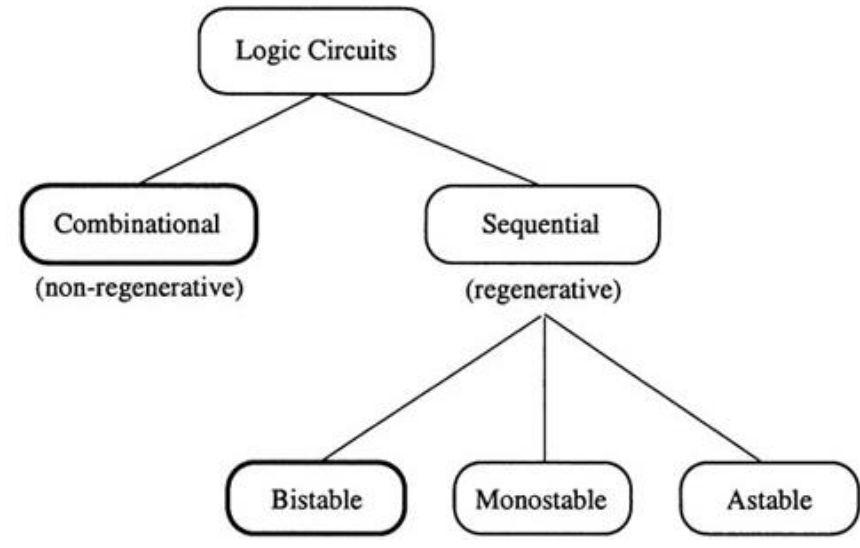
Assistant Professor

University of Jordan

Introduction

- Sequential Circuit

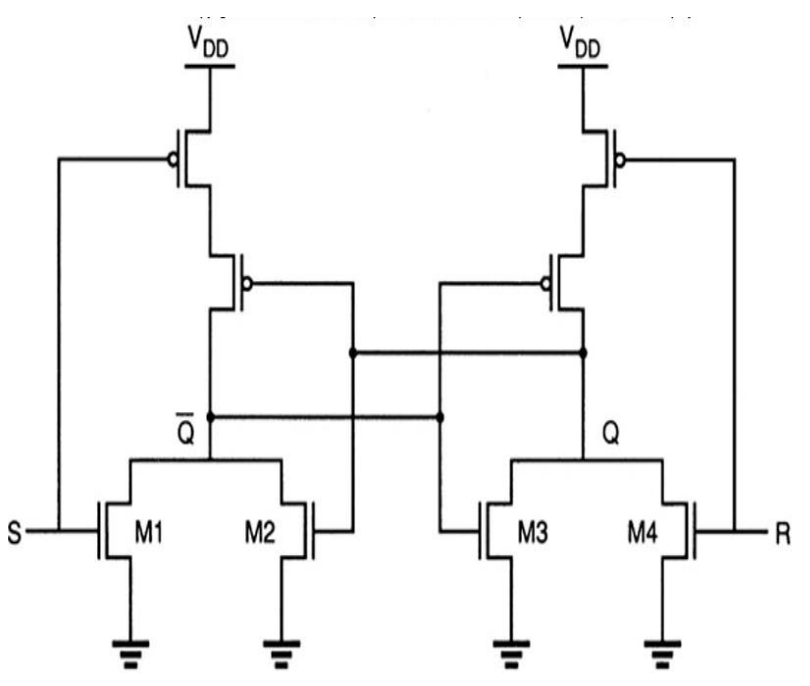




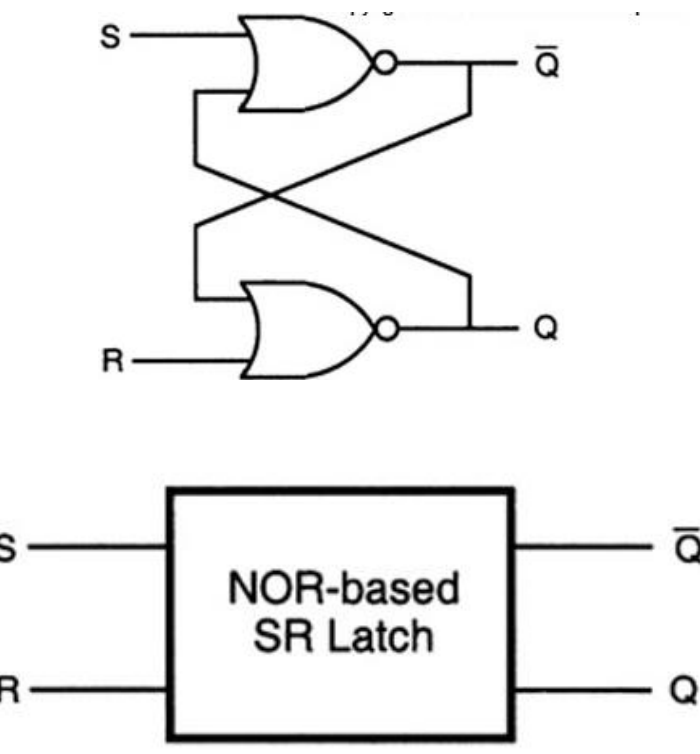
Combinational block
+memory block

SR Latch Circuit

- SR latch circuit based on NOR2 gates



- Gate level schematic and block diagram



Truth Table and Operation Mode

- Truth table of the NOR based SR latch circuit

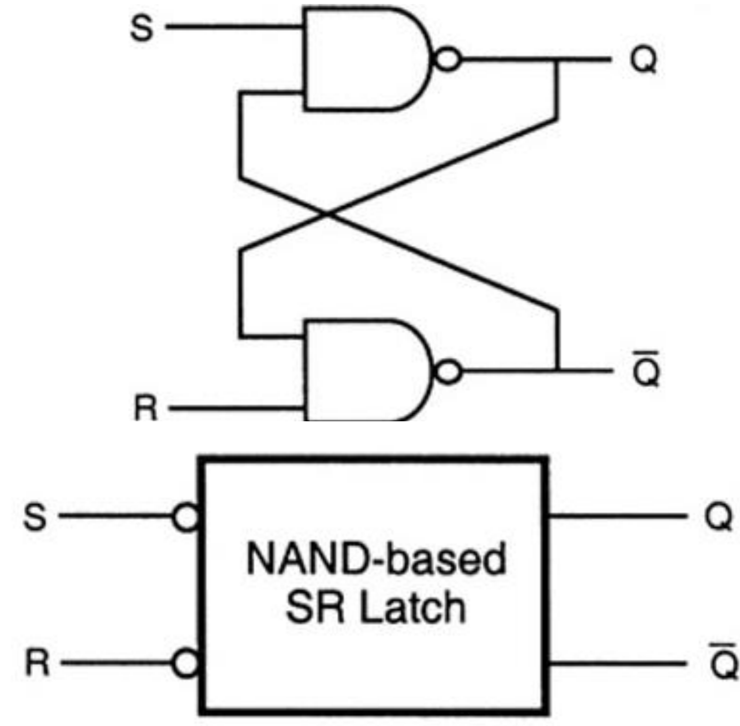
S	R	Q_{n+1}	$\overline{Q_{n+1}}$	Operation
0	0	Q_n	$\overline{Q_n}$	hold
1	0	1	0	set
0	1	0	1	reset
1	1	0	0	not allowed

- Operation mode of the NOR based SR latch circuit

S	R	Q_{n+1}	$\overline{Q_{n+1}}$	Operation
V_{OH}	V_{OL}	V_{OH}	V_{OL}	M1 and M2 on, M3 and M4 off
V_{OL}	V_{OH}	V_{OL}	V_{OH}	M1 and M2 off, M3 and M4 on
V_{OL}	V_{OL}	V_{OH}	V_{OL}	M1 and M4 off, M2 on, or
V_{OL}	V_{OL}	V_{OL}	V_{OH}	M1 and M4 off, M3 on

NAND Based SR Latch

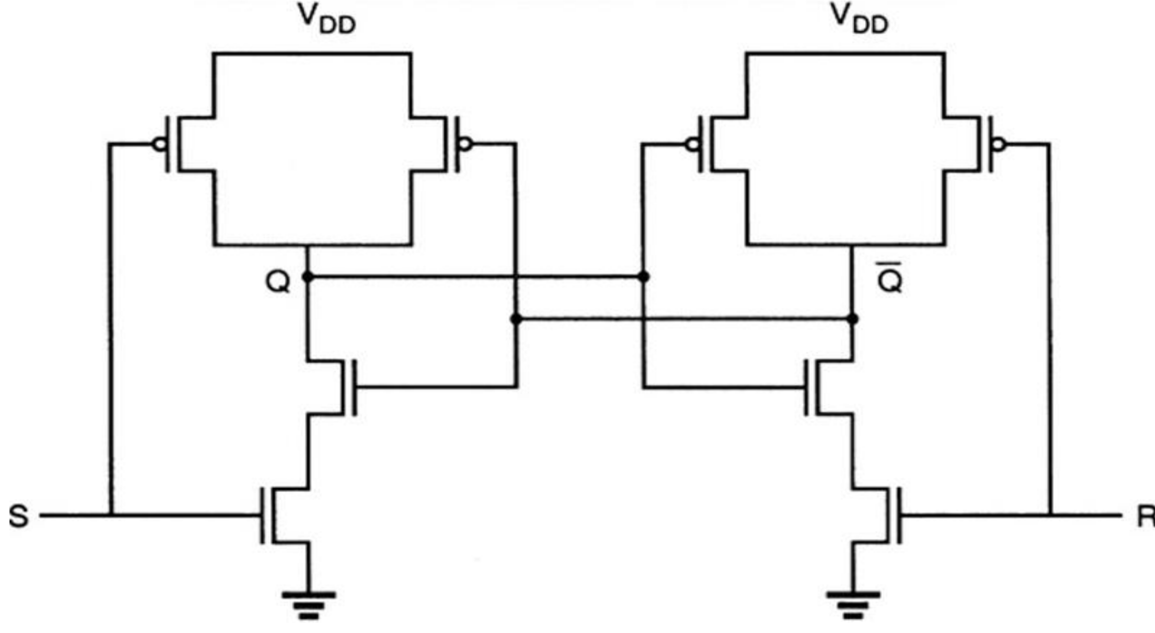
- Gate level schematic & Block diagram



S	R	Q_{n+1}	$\overline{Q_{n+1}}$	Operation
0	0	1	1	not allowed
0	1	1	0	set
1	0	0	1	reset
1	1	Q_n	$\overline{Q_n}$	hold

CMOS SR Latch : Another Type

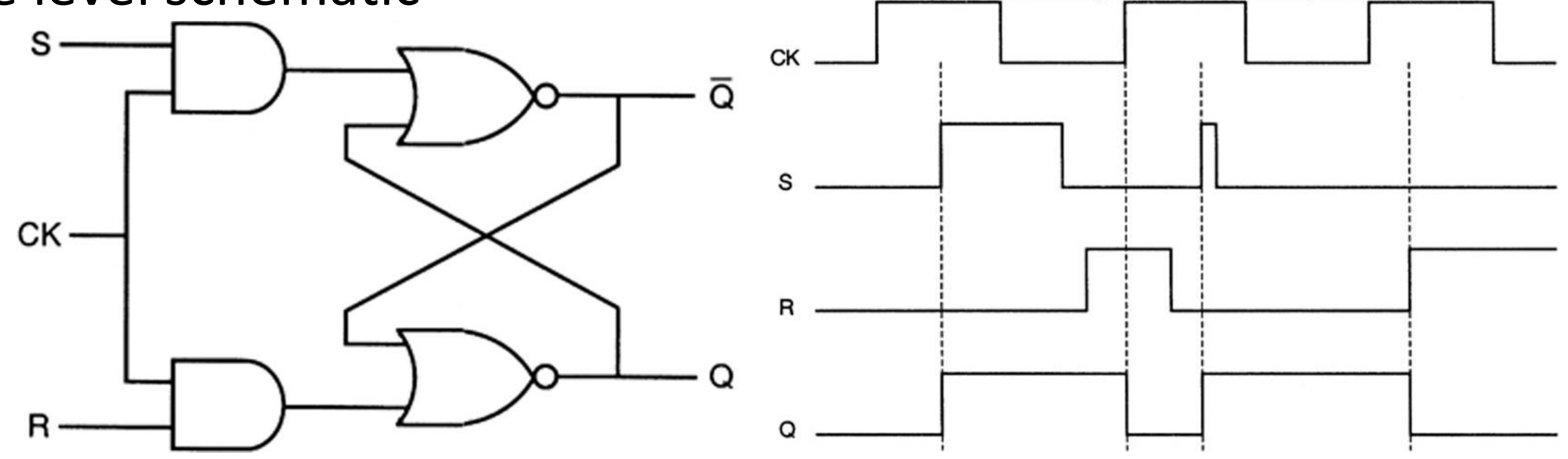
- SR latch based on NAND2 gates



Clocked SR Latch

- Input and output waveform

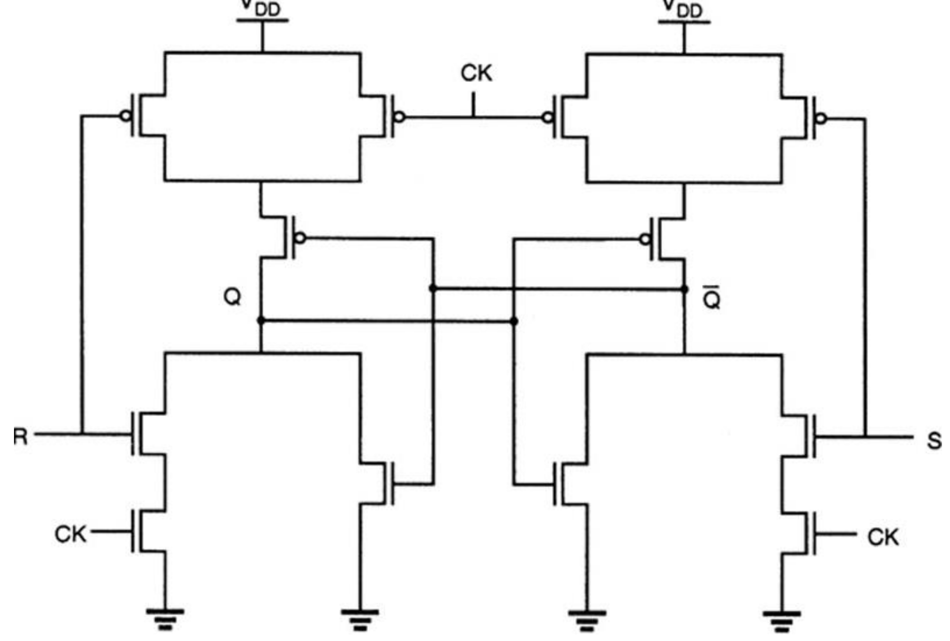
- Gate level schematic



Level sensitive circuit

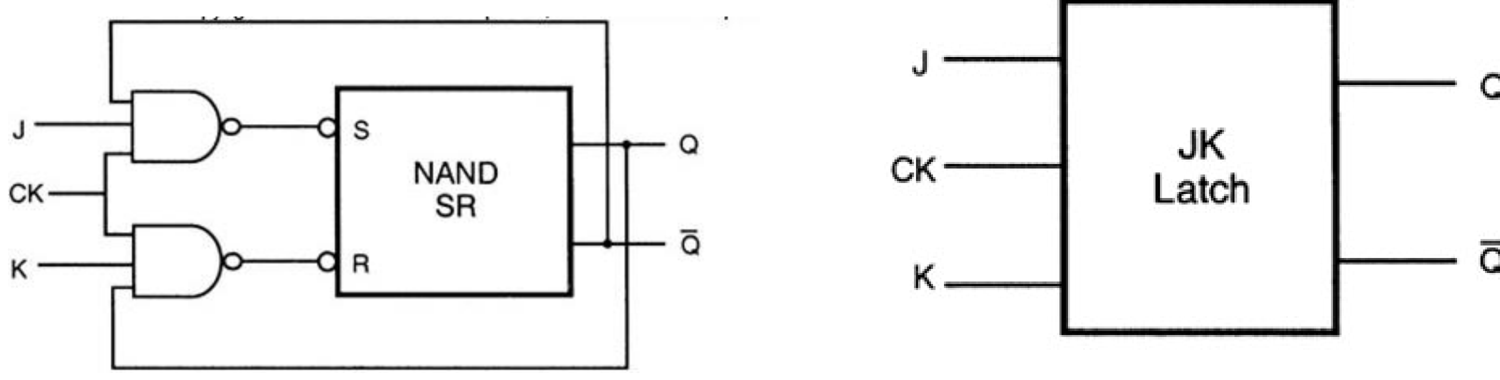
Clocked NOR Based SR Latch : AOI

- AOI-based implementation of the clocked NOR-based SR-latch Circuit



Clocked JK Latch

- Gate level schematic



- SR-latch : indeterminate when both inputs S and R are activated
- JK latch : adding two feedback lines from the outputs to the inputs

Clocked NOR-based JK latch

Clocked JK Latch : Truth Table

J	K	Q_n	$\overline{Q_n}$	S	R	Q_{n+1}	$\overline{Q_{n+1}}$	Operation
0	0	0	1	1	1	0	1	hold
		1	0	1	1	1	0	
0	1	0	1	1	1	0	1	reset
		1	0	1	0	0	1	
1	0	0	1	0	1	1	0	set
		1	0	1	1	1	0	
1	1	0	1	0	1	1	0	toggle
		1	0	1	0	0	1	

Master-Slave Flip-Flop

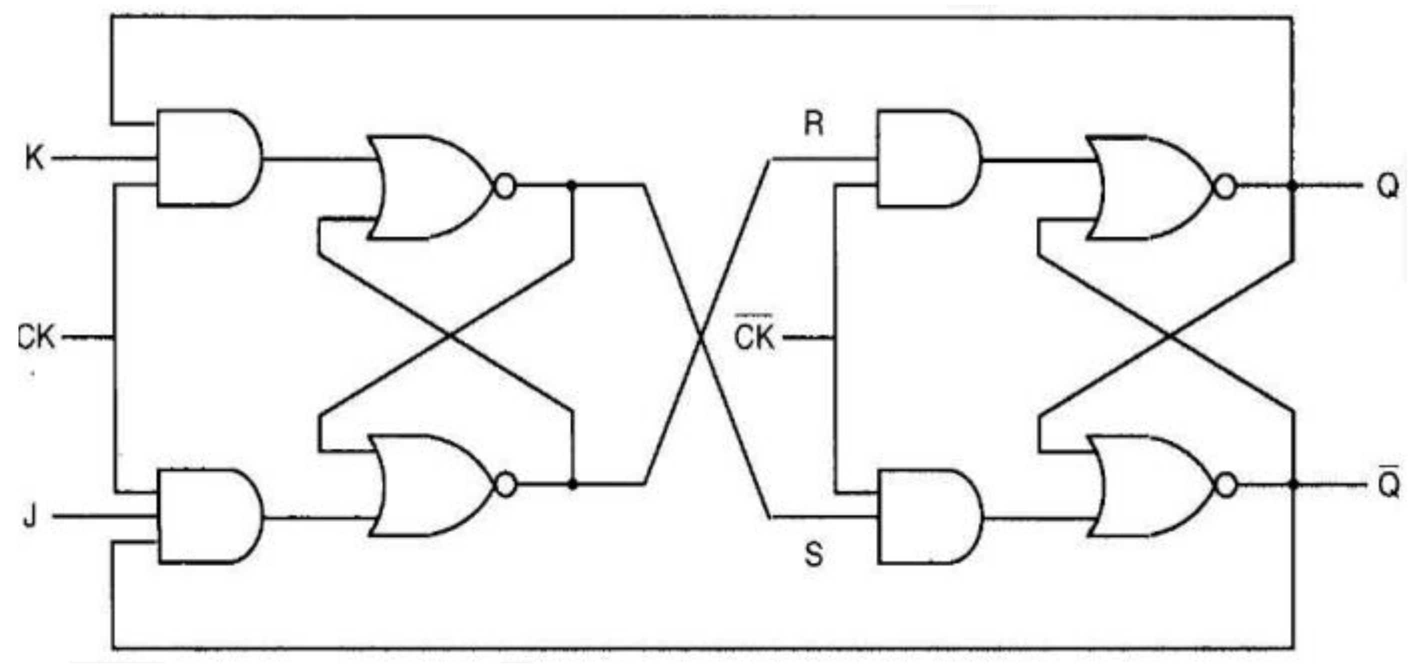
- Master-Slave Flip-Flop
  - Two c

J
CK
K
S
R
NAND SR
Q_m
$\overline{Q_m}$
S
R
NAND SR
Q_s
$\overline{Q_s}$
$\overline{CK}$
Master
Slave

Master Slave FF

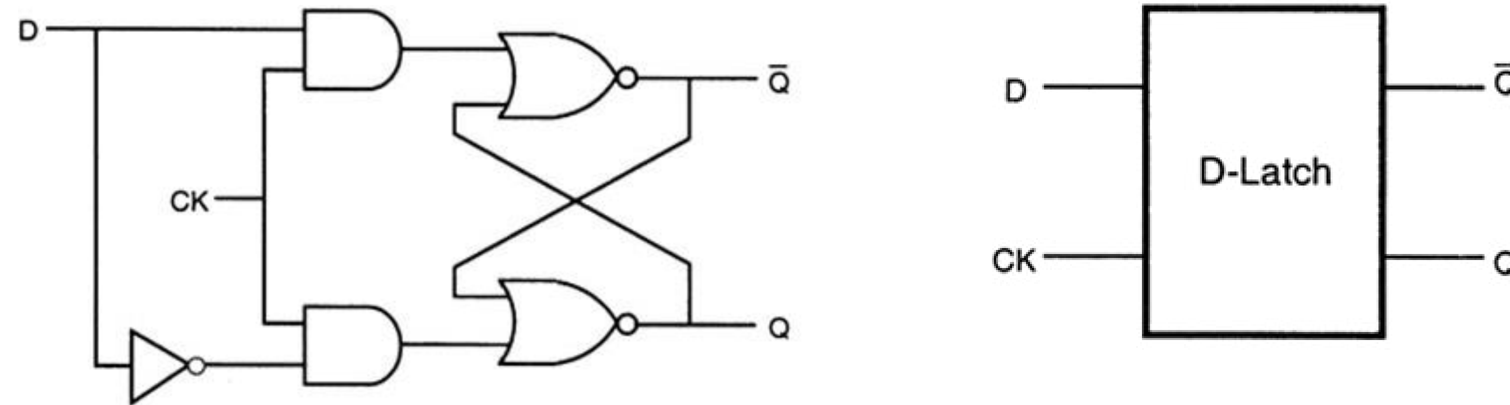
- The master latch is activated when CK="1," during this period, the primary inputs, J and K, allow data to be entered into the flip-flop, and the first stage outputs are set according to the primary inputs
- When CK="0," the master latch becomes inactive, and the slave latch becomes active
- Note that, the output levels of the flip-flop are determined during this phase (CK="0,") based on the master-stage outputs set in the previous phase(CK="1")

NOR based JK master slave FF



CMOS D-Latch

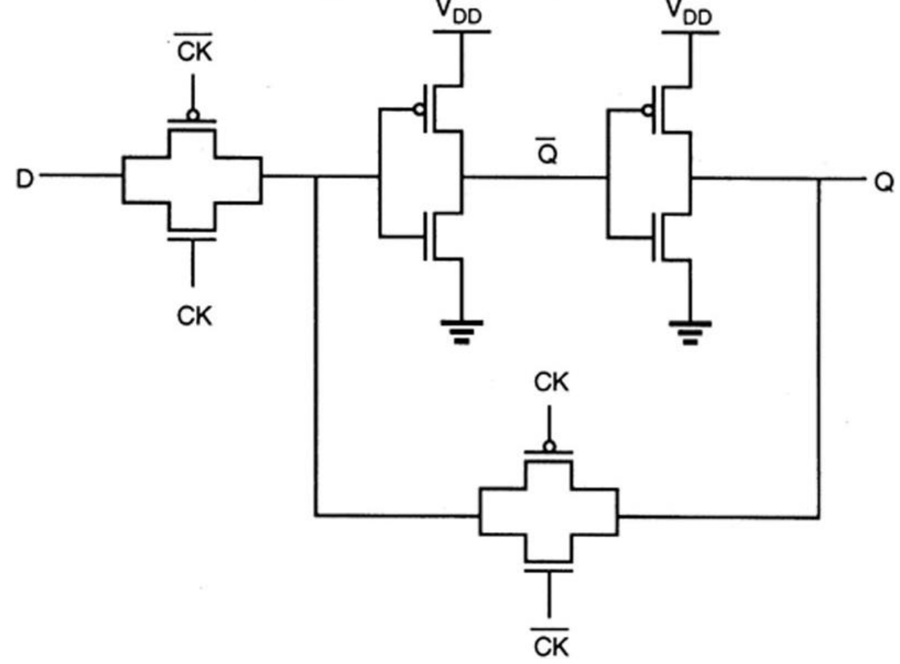
- Gate-level schematic



- CK : 1 → Q assumes the value of the input D
- CK : 0 → Q preserve its state

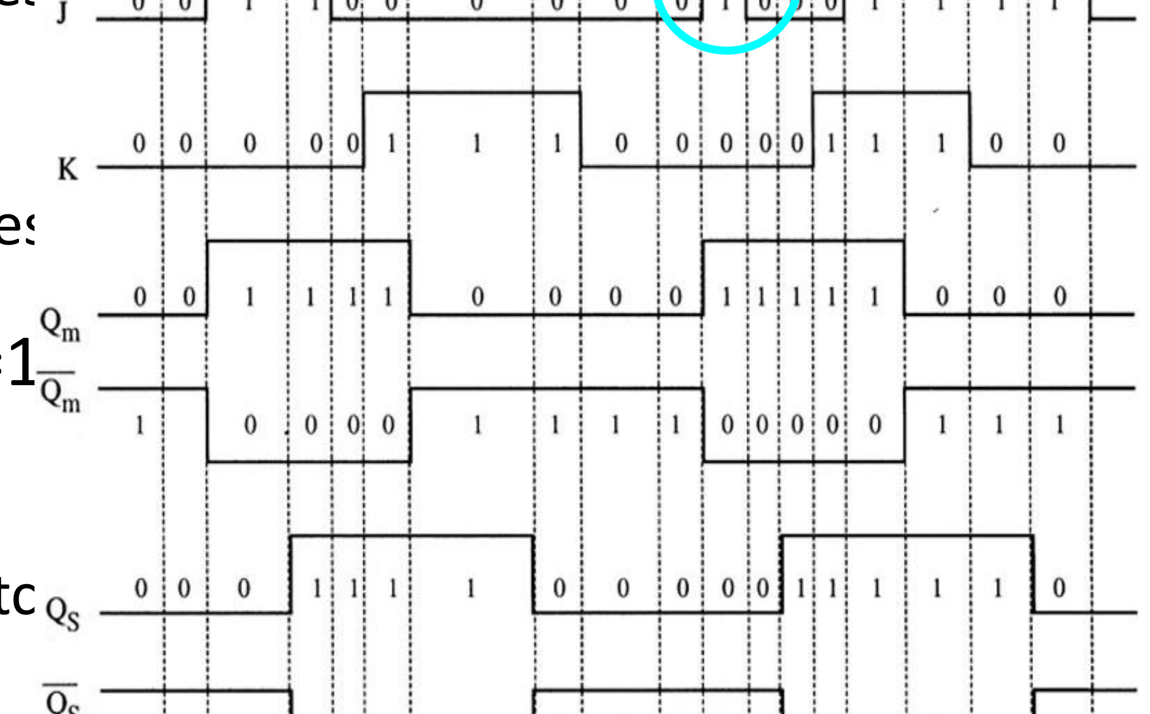
CMOS D-Latch Using Transmission gates

- Constructed by Two inverter loop + Two CMOS TG
- CK:1 → TG at input is activated
- CK:0 → TG at inverter loop is activated



Master-Slave Flip-Flop : Operation

- Clock pulse → 0
  - Master latch inactive (slave becomes
- Clock pulse → 1
  - Slave latch inactive (master becomes
- No uncontrolled oscillation : J=K=1
- Ones catching problem
  - Unwanted o/p transition due to glitc
  - Sol. : edge-triggered



Master-Slave Flip-Flop Problem

- S and/or R are permitted to change while C = 1
  - Chances of 0s or 1s catching

C

S

R

Y
Master out

Q
Slave out

Master
active

Slave
active

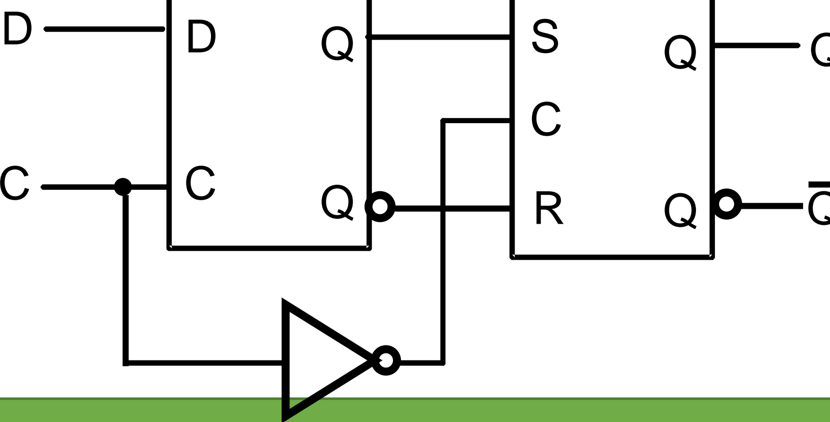
1s catching

Flip-Flop Solution

- Use *edge-triggering* instead of master-slave
- An *edge-triggered* flip-flop ignores the pulse while it is at a constant level and triggers only during a *transition* of the clock signal
- Edge-triggered flip-flops can be built directly at the electronic circuit level, or
- A *master-slave D flip-flop* which also exhibits *edge-triggered behavior* can be used

Edge-Triggered D Flip-Flop

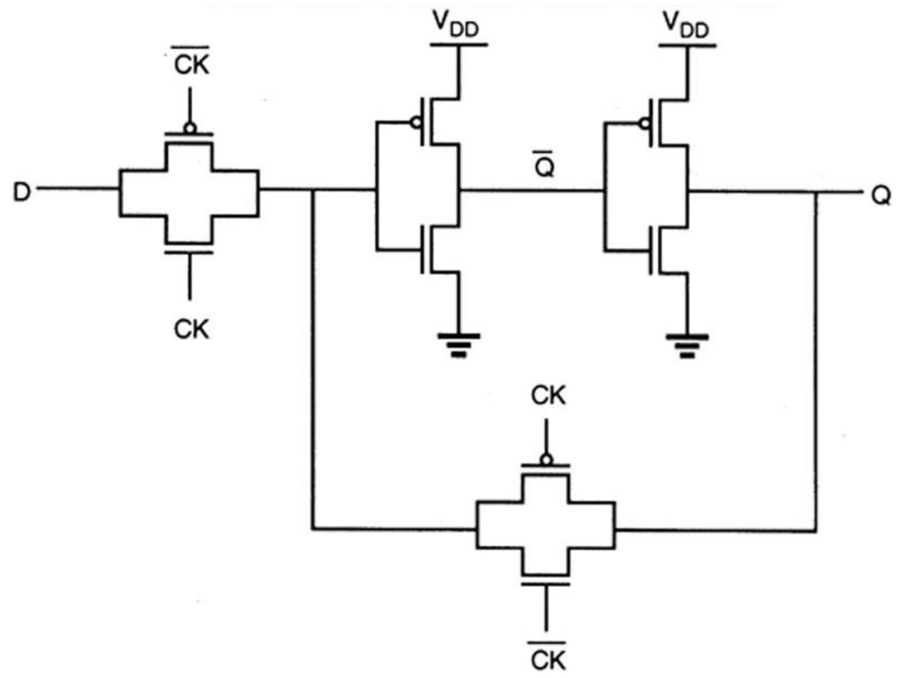
- *The edge-triggered D flip-flop is the same as the master-slave D flip-flop*
- It can be formed by:
    - Replacing the first clocked SR latch with a clocked D latch or
    - Adding a D input and inverter to a master-slave SR flip-flop
- The 1s and 0s catching behaviors are not present with D replacing S and R inputs
- The change of the D flip-flop output is associated with the negative edge at the end of the pulse
- *It is called a negative-edge triggered flip-flop*



No 1s catching in the edge-triggered D Flip-Flops

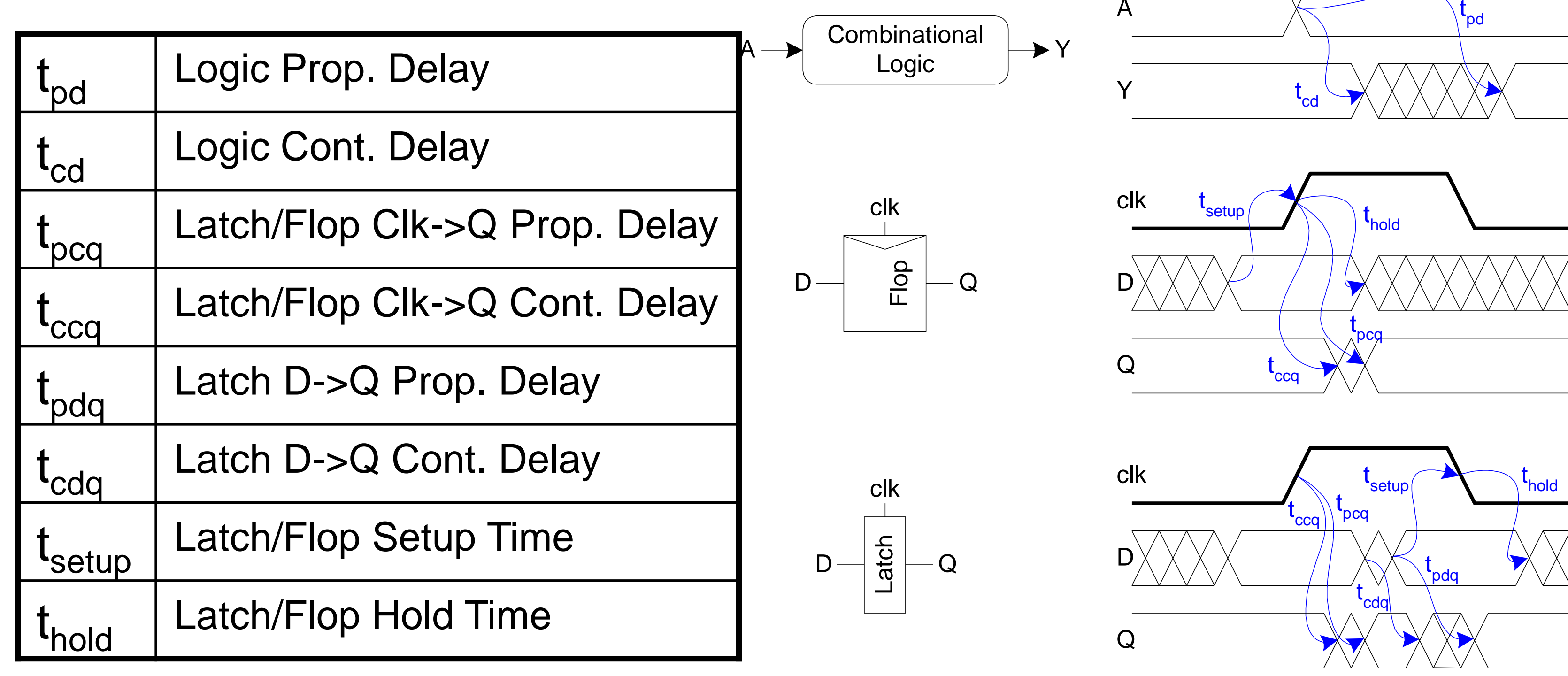
D
C
D
C
Q
Q
Y
S
C
R
Q
Q
Q
$\overline{Q}$
C
D
Y
Master out
Q
Slave out
Master active
Slave active
no 1s catching
correct output

Setup and Hold time



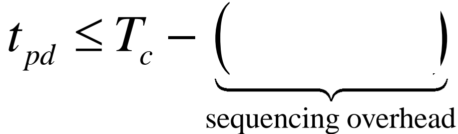
Timing Diagrams

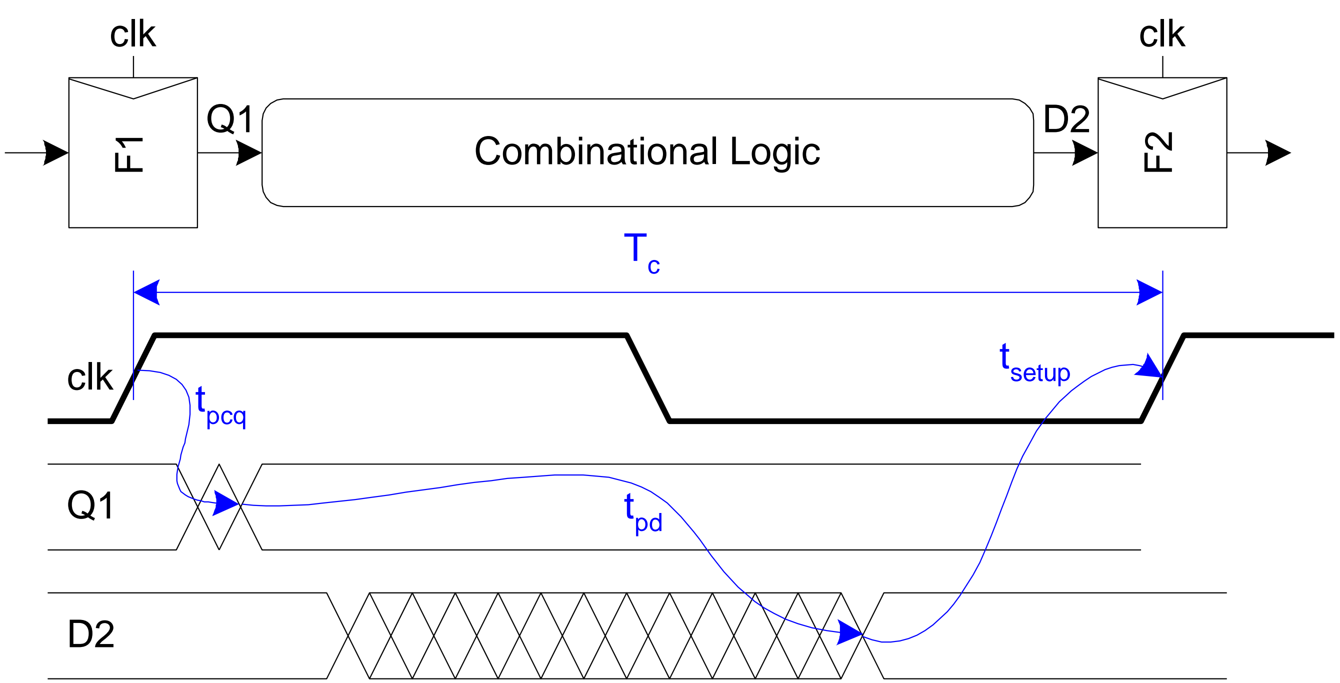
Contamination and
Propagation Delays



t_{pd}	Logic Prop. Delay
t_{cd}	Logic Cont. Delay
t_{pcq}	Latch/Flop Clk->Q Prop. Delay
t_{ccq}	Latch/Flop Clk->Q Cont. Delay
t_{pdq}	Latch D->Q Prop. Delay
t_{cdq}	Latch D->Q Cont. Delay
t_{setup}	Latch/Flop Setup Time
t_{hold}	Latch/Flop Hold Time

Max-Delay: Flip-Flops





Min-Delay: Flip-Flops

$t_{cd} \geq$