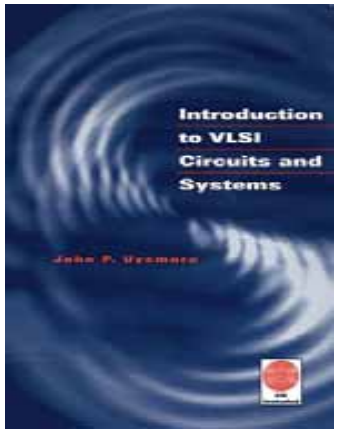
國立勤益科技大學
National Chin-Yi University of Technology

Introduction to VLSI Circuits and Systems
積體電路概論

Chapter 03
Physical Structure of CMOS Integrated Circuits



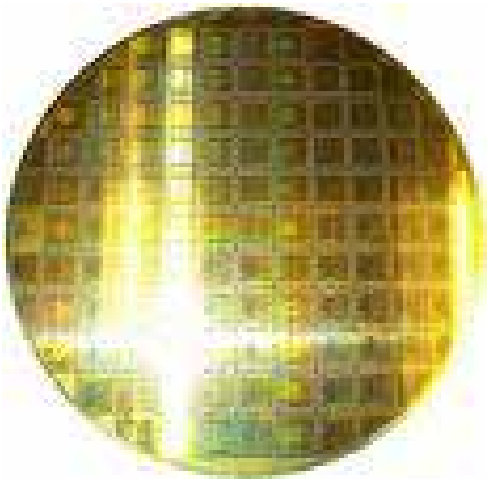
賴秉樑
Dept. of Electronic Engineering
National Chin-Yi University of Technology
Fall 2007

Outline

- An Overview CMOS Fabrication
- Integrated Circuit Layers
- MOSFETs
- CMOS Layers
- Designing FET Arrays

CMOS Fabrication

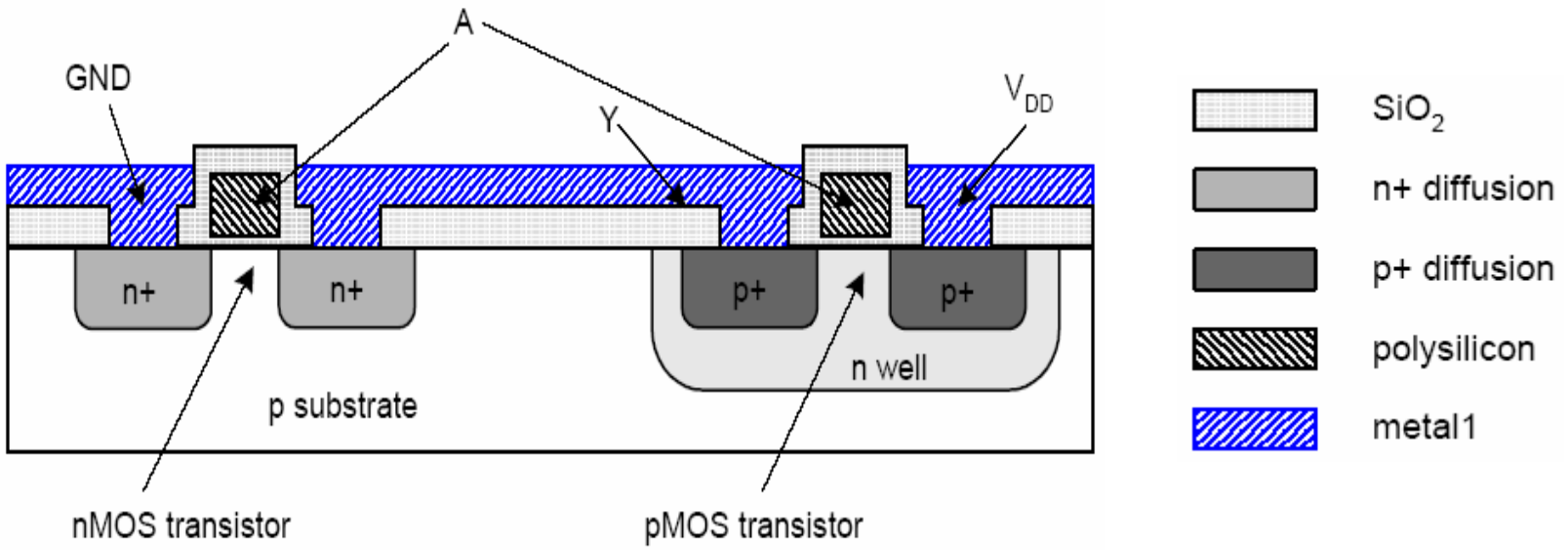
- CMOS transistors are fabricated on silicon wafer
- Lithography process similar to printing press
- On each step, different materials are deposited or etched
- Easiest to understand by viewing both top and cross-section of wafer in a simplified manufacturing process



Wafer

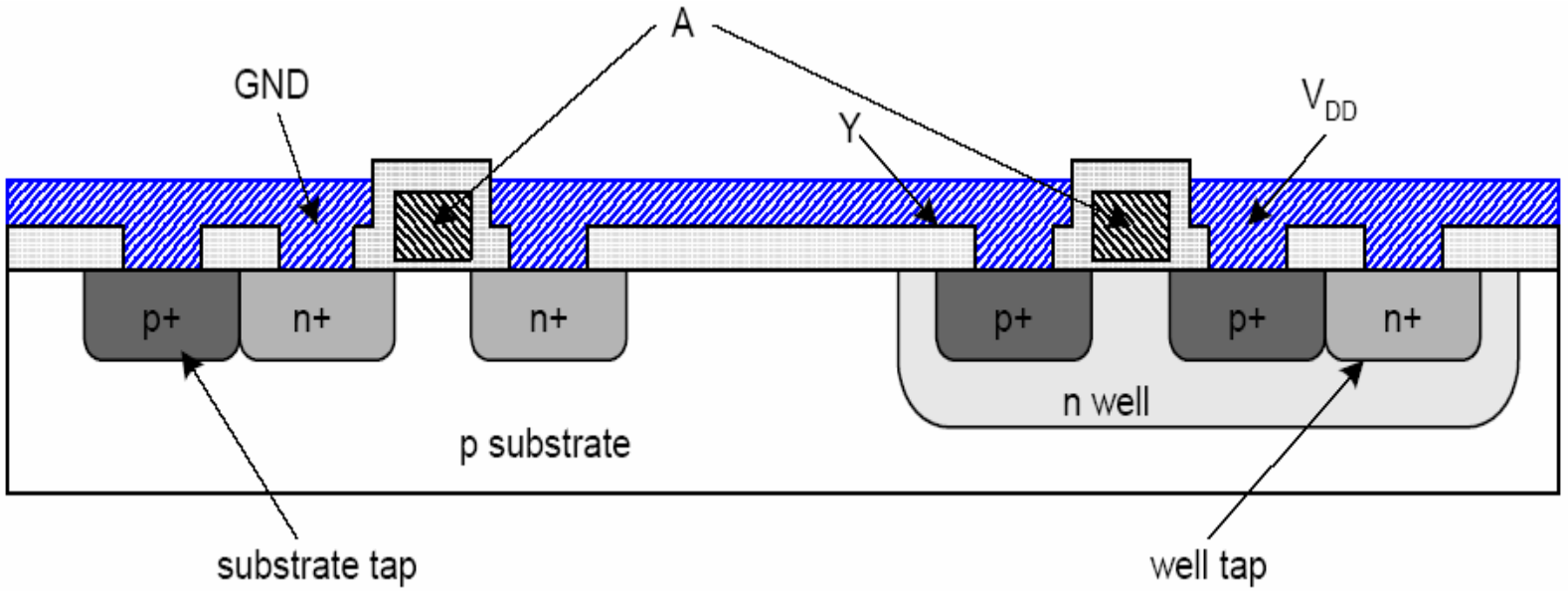
Inverter Cross-section

- Typically use p-type substrate for nMOS transistors
- Requires n-well for body of pMOS transistors



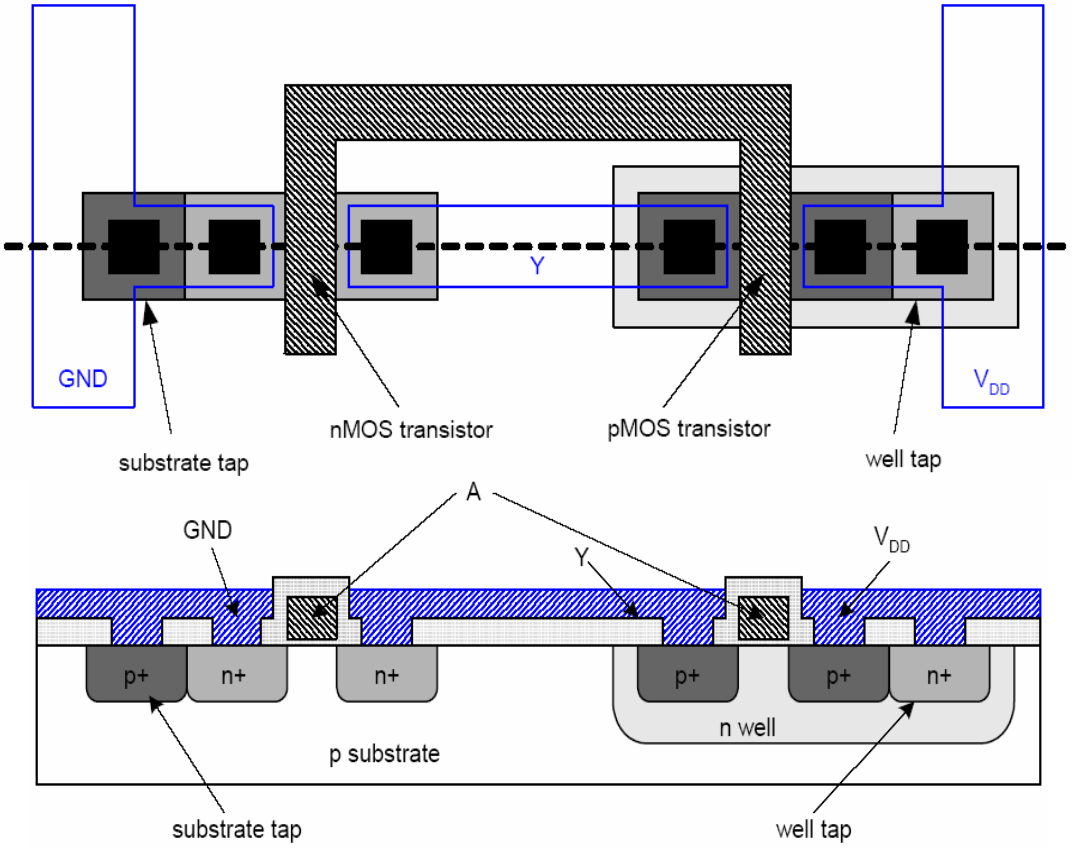
Well and Substrate Taps

- 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



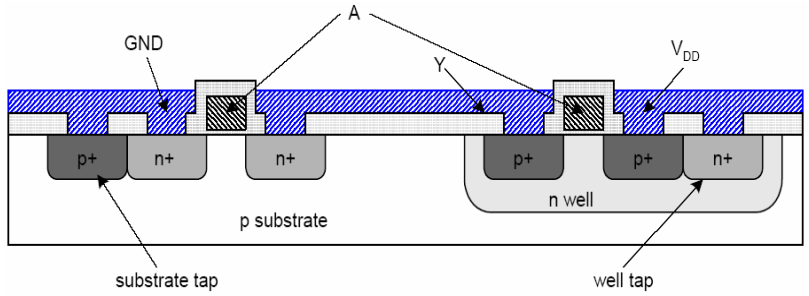
Inverter Mask Set

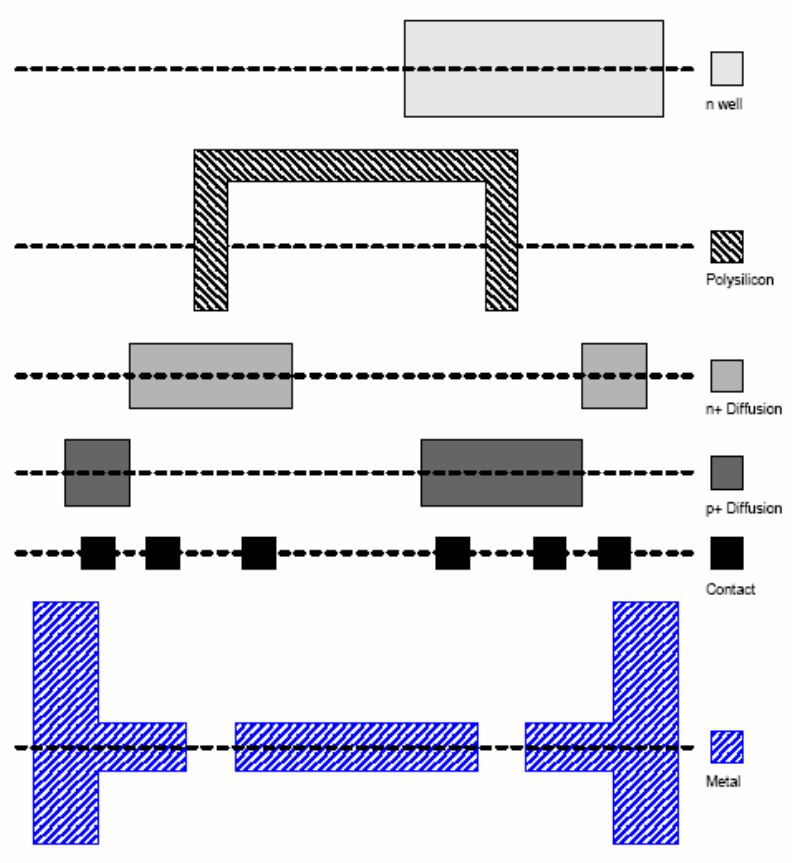
- Transistors and wires are defined by *masks*
- Cross-section taken along dashed line



Detailed Mask Views

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



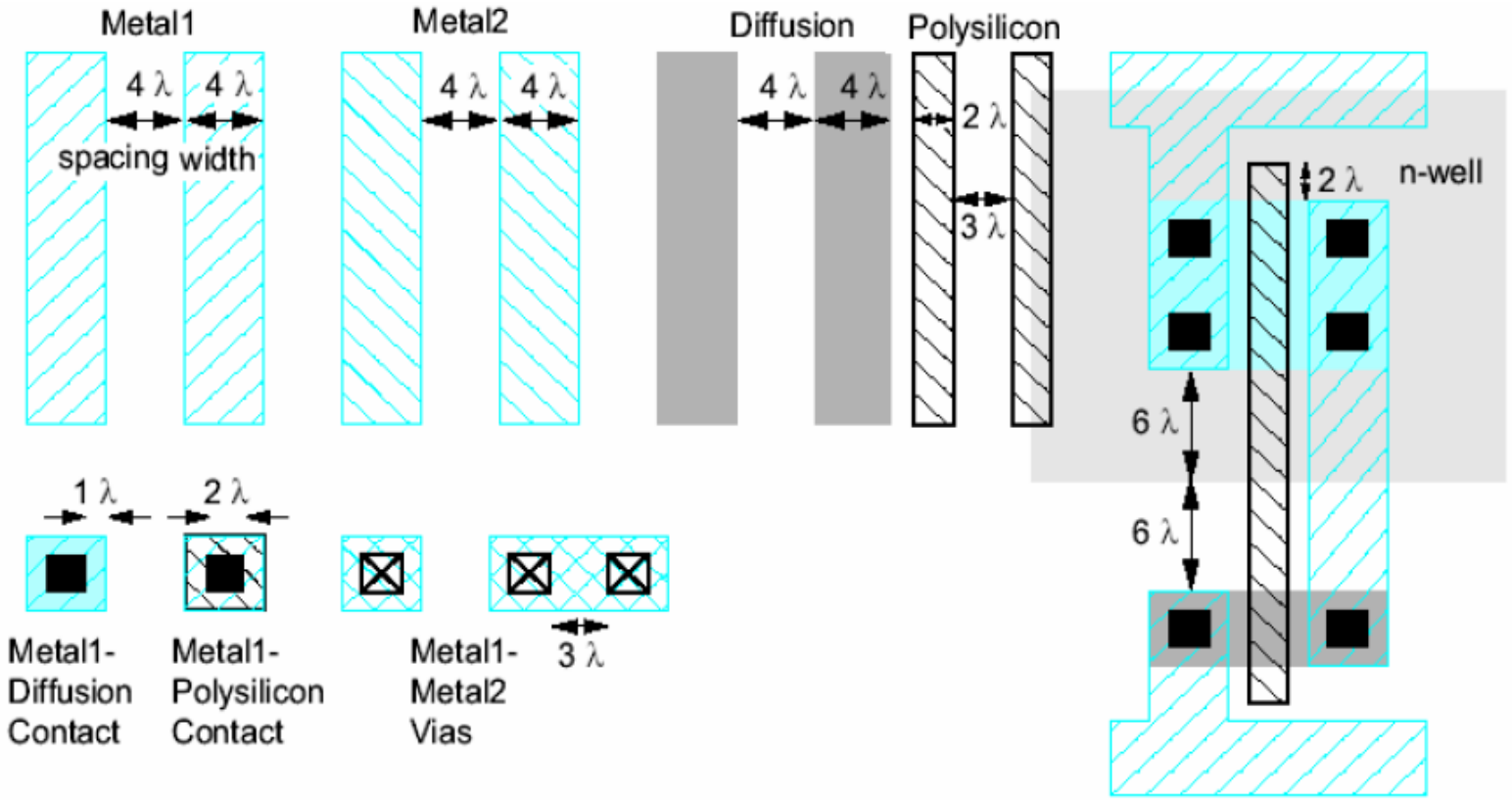


Layout

- 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. λ = 0.3 μm in 0.6 μm process

Simplified Design Rules

- Conservative rules to get you started



Outline

- An Overview CMOS Fabrication
- Integrated Circuit Layers
- MOSFETs
- CMOS Layers
- Designing FET Arrays

Integrated Circuit Layers

- *Physical design*: CMOS integrated circuits are electronic switching networks that are created on small area of a silicon wafer using a complex set of physical and chemical processes
  - Metal
  - Insulator
  - Substrate

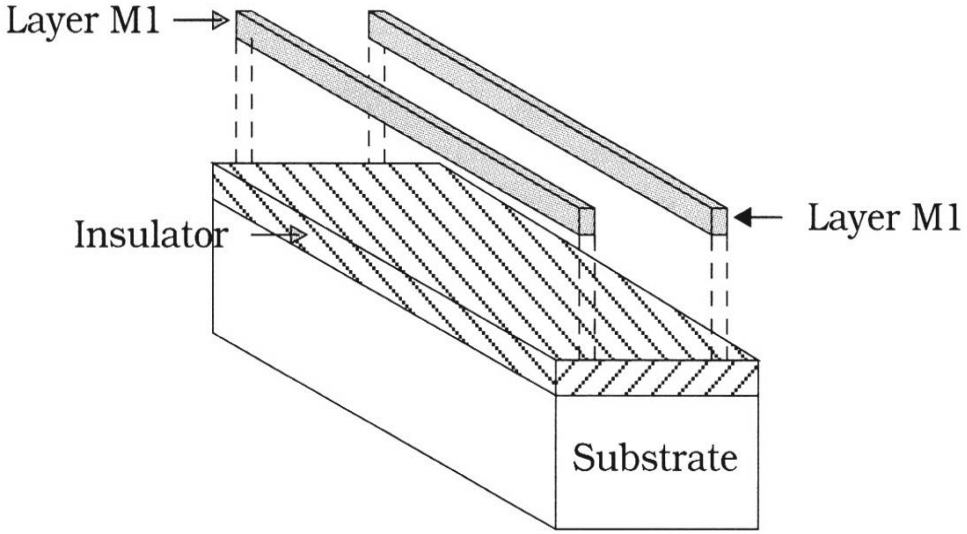
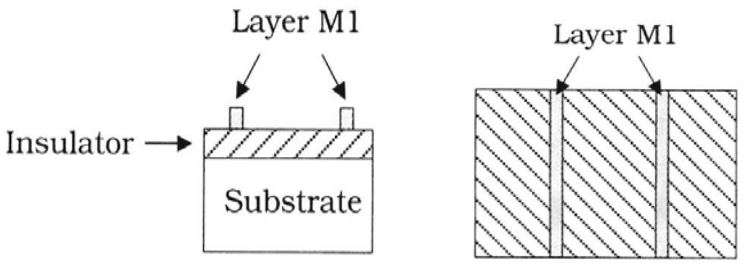


Figure 3.1 Two separate material layers

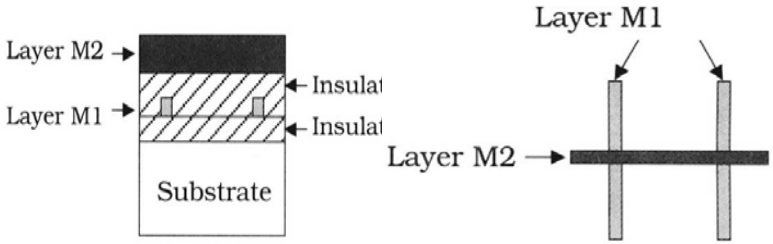
Three-dimensional Structure

- Combining the top and side views of an integrated circuit allows us to visualize the three-dimensional structure
  - » The side view illustrates the order of the stacking
  - » Insulating layers separate the two metal layers so that they are electrically distinct
  - » The patterning of each layer is shown by a top view perspective
- The stacking order is established in the manufacturing process, and can not be altered by the VLSI designer



(a) Side view (b) Top view

Figure 3.2 Layers after the stacking process is completed



(a) Side view (b) Top view

Figure 3.3 Addition of another insulator and a second metal layer

Interconnect Resistance and Capacitance

- Logic gates communicate with each other by signal flow paths from one point to another
  - » Using patterned metal lines
  - » Current flow is governed by the physical characteristics of the material and the dimensions of the line
  - » *Ohm's law*

$$V = IR \qquad (3.1)$$

  - » *Line resistance* R_{line}: a parasitic (unwanted) electrical element that cannot be avoided

$$A = wt \qquad (3.2)$$

$$R_{line} = \frac{l}{\sigma A} \qquad (3.3) \qquad (\sigma\text{:conductivity})$$

$$\text{Since} \quad \rho = \frac{1}{\sigma} \qquad (3.4) \qquad (\rho\text{ :resistivity})$$

$$\Rightarrow R_{line} = \rho \frac{l}{A} \qquad (3.5)$$

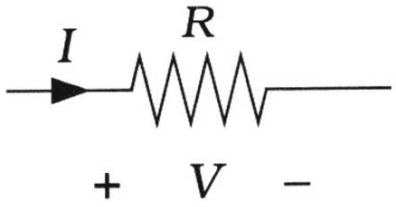


Figure 3.4 Symbol for a linear resistor

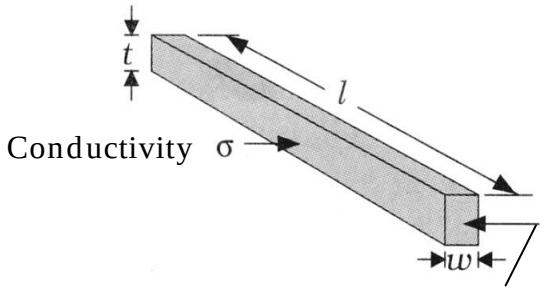


Figure 3.5 Geometry of a conducting line

Sheet Resistance Model

- Sheet resistance R_s, rewriting $R_{line} = \rho \frac{l}{A}$

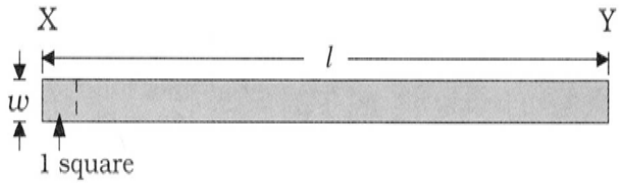
$$\Rightarrow R_{line} = \left(\frac{1}{\sigma t}\right)\left(\frac{l}{w}\right) \qquad (3.6)$$

$$R_s = \frac{1}{\sigma t} = \frac{\rho}{t} \qquad (3.7) \quad \text{(sheet resistance)}$$

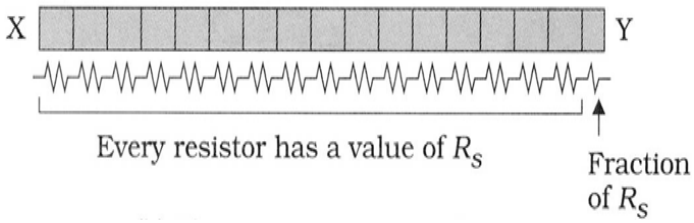
$$\Rightarrow R_{line} = R_s\left(\frac{w}{w}\right) = R_s \qquad (3.8)$$

$$\Rightarrow R_{line} = R_s n \qquad (3.9)$$

$$where \quad n = \frac{1}{w} \qquad (3.10)$$



(a) Top-view geometry



(b) Sheet resistance contributions

Figure 3.6 Top-view geometry of a patterned line

Capacitor

- Interconnect lines also exhibit the property of capacitance
  - In electronics, the element that stores charge is called *capacitor*

$$Q = CV \qquad (3.11)$$

C: F (farad), 1 F = 1 C/V

  - Since electric current is defined by the time derivative $I = (dQ/dt)$, differentiating gives the *I-V* equation

$$I = C\frac{dV}{dt} \qquad (3.12)$$

  - Capacitance exists between any two conducting bodies that are electrically separated
    - For the interconnect line, the conductor is isolated from the substrate by an insulating layer of silicon dioxide glass
    - So, the capacitance depends on the geometry of the line

$$C_{line} = \frac{\varepsilon_{ox} wl}{T_{ox}} \qquad (3.13) \qquad \text{(parallel-plate formula)}$$

Where ε_{ox} is the permittivity of the insulating oxide F / cm

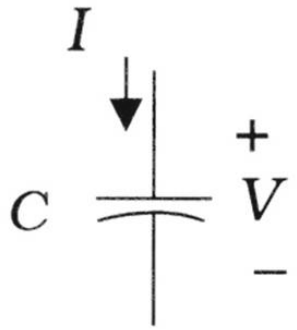


Figure 3.7 Circuit symbol for a capacitor

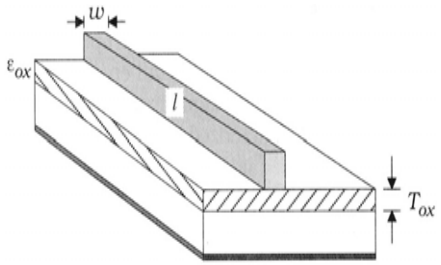


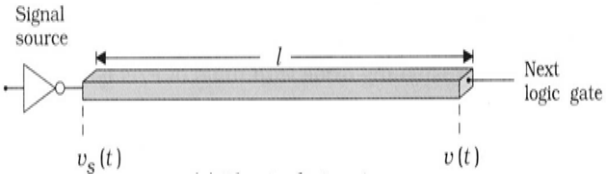
Figure 3.8 Geometry for calculating the line capacitance

Delay: RC Time Constant

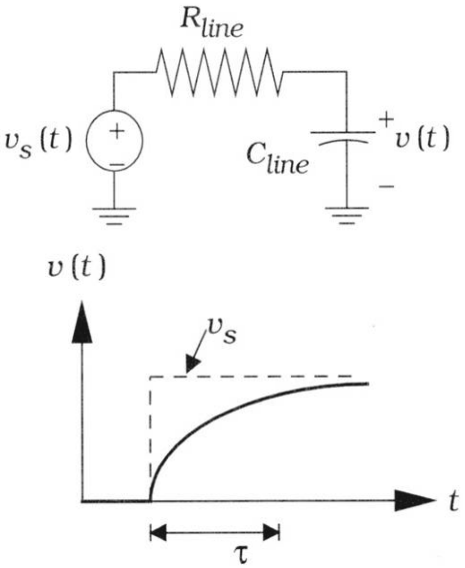
- The interconnect line exhibits both parasitic resistance R_{line} [Ω] and capacitance C_{line} [F]
  - » Forming the product of these two quantities gives

$$\tau = R_{line}C_{line}\ [s] \qquad (3.14)$$

- In high speed digital circuits, signals on an interconnect line are delayed by τ, which places a limiting factor on the speed of the network
  - » VLSI processing are directed toward minimizing both R_{line} and C_{line}
  - » Circuit designers are then faced with creating the fastest switching network within the limits of delay



(a) Physical structure



(b) Circuit model

Figure 3.9 Time delay due to the interconnect time constant

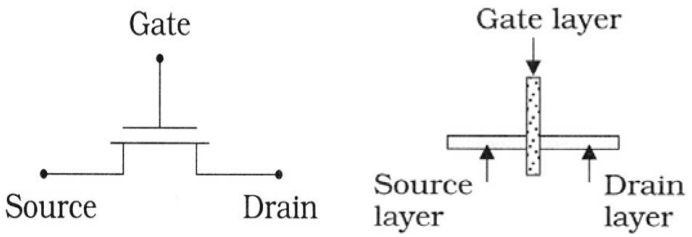
Outline

- An Overview CMOS Fabrication
- Integrated Circuit Layers
- MOSFETs
- CMOS Layers
- Designing FET Arrays

MOSFETs

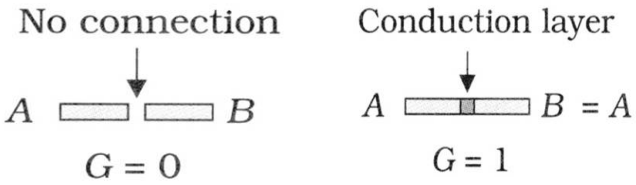
- MOSFET is a small area set of two basic patterned layers that together act like a controlled switch
  - » The voltage applied to the gate determines the electrical current flow between the source and drain terminals
- Assuming that the drain and source are formed on the same layer, then this behavior can be used to deduce that

The gate signal G is responsible for the absence or presence of the conducting region between the drain and source region



(a) nFET symbol (b) pFET symbol

Figure 3.10 nFET circuit symbol and layer equivalents



(a) Open switch (b) Closed Switch

Figure 3.11 Simplified operational view of an nFET

Physical Structure of MOSFET

- The drain and source regions are patterned into a silicon wafer
- *L* has units of centimeters [cm] and is called the **channel length** of the FET
- The width *W* of the drain and source regions is called the **channel width** and also has units of centimeters
- The aspect ratio of the FET is defined as (*W/L*) and is the most important parameter to the VLSI designer

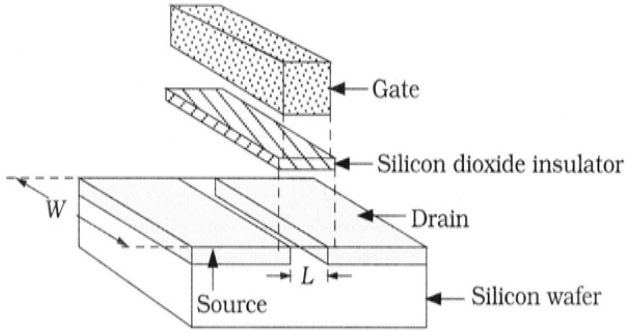
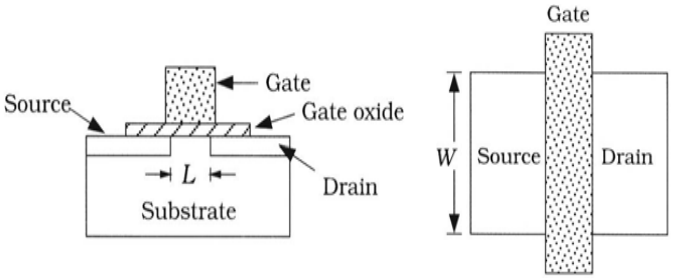


Figure 3.12 Layers used to create a MOSFET



(a) Side view (b) Top view

Figure 3.13 Views of a MOSFET

Electrical Conduction in Silicon (N-type)

- Silicon is a semiconductor because it can conduct small amounts of electrical current, making it a "partial" conductor
- *N-type*

$N_{Si} \approx 5 \times 10^{22}$ (The atomic density of Si cm^{-3})

$n_i \approx 1.45 \times 10^{10}\ cm^{-3}$ (3.16) (Intrinsic carrier density)

$n = p = n_i$ (3.17) (Pure silicon)

$np = n_i^2$ (3.18) (mass-action law)

$n_n = N_d\ cm^{-3}$ (3.19) (n-type donor atom, majority carrier)

$p_n = \frac{n_i^2}{N_d} cm^{-3}$ (3.20) (minority carrier)

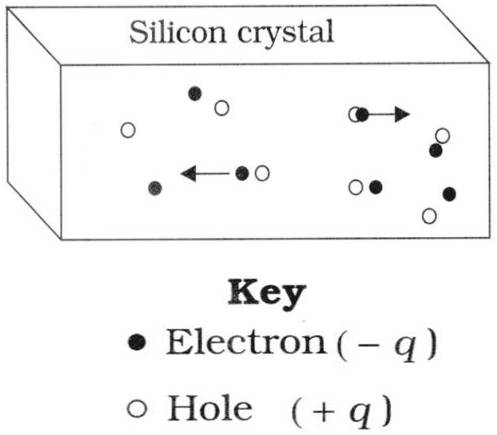


Figure 3.14 Creation of electron-hole pairs in silicon

Electrical Conduction in Silicon (P-type)

- *P-type*: has more positively charged holes than negatively charged electrons

$$p_p = N_a \qquad n_p = \frac{n_i^{\,2}}{N_a} \qquad (3.24) \qquad \text{(acceptor)}$$

$$\sigma = q(\mu_n n + \mu_p p) \qquad (3.25) \qquad \text{(conductivity)}$$

$$\mu_n = 1360 \qquad \mu_p = 480 \qquad (3.26) \qquad \text{(the electron and hole mobility } cm^2/\text{V-sec)}$$

$$\Rightarrow \sigma \approx 4.27\times10^{-6}\ [\Omega - cm]^{-1} \ or\ \rho \approx 2.34\times10^{5}\ [\Omega - cm]$$

$$\sigma = q\mu_n n_n \qquad (3.27) \qquad \text{(n-type, since } n_n >> p_n\text{)}$$

$$\sigma = q\mu_p p_p \qquad (3.28) \qquad \text{(p-type, since } p_p >> n_p\text{)}$$

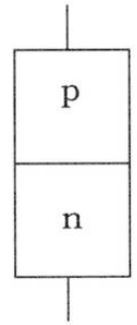
PN Junction

- In CMOS processing, most doped regions have both donors and acceptors
- To create an n-type region, we need $N_d > N_a$

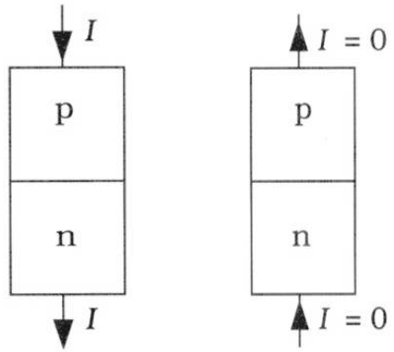
$$n_n = N_d - N_a \qquad p_n = \frac{n_i^2}{(N_d - N_a)} \tag{3.35}$$

- For a p-type region, we need $N_a > N_d$

$$p_p = N_a - N_d \qquad n_p = \frac{n_i^2}{(N_a - N_d)} \tag{3.36}$$



(a) A pn junction

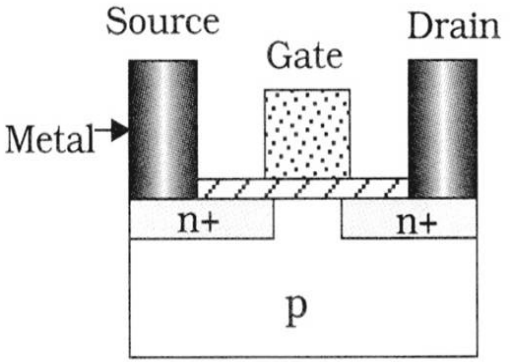


(b) Forward current (c) Reverse blocking

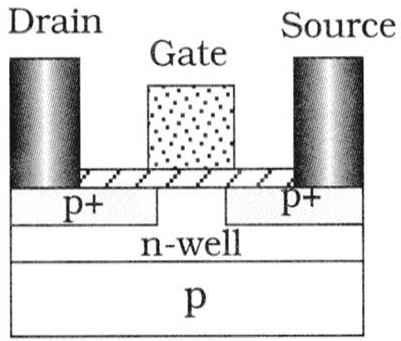
Figure 3.15 Formation and characteristics of a pn junction

nFET and pFET

- The polarity of a FET (n or p) is determined by the polarity of the drain and source regions
- *nFET*: the drain and source regions are labeled as "n+" to indicate that they are **heavily doped** as Figure 3.16 (a) showing
- *pFET*: the source and drain regions are p+ sections that are embedded in an n-type "well" layer as Figure 3.16 (b) showing
- All pn junction are used to prevent current flow between adjacent layers



(a) nFET cross-section



(b) pFET cross-section

Figure 3.16 nFET and pFET layers

Current Flow in a FET

- The creation of the conducting layer underneath the gate is due to the property of the capacitance that is built into the gate region of the MOSFET itself

$$Q = CV \qquad (3.38)$$

$$Since \quad C = \frac{\varepsilon A}{t_{ins}} \qquad (3.39) \qquad \text{(a parallel-plate structure)}$$

where ε is the permittivity of the insulator in units of F/cm

$$\Rightarrow C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \qquad (3.40) \qquad \text{(oxide capacitance)}$$

$$\Rightarrow C_G = C_{ox} A_G \qquad (3.41) \qquad \text{(gate capacitance)}$$

$$\varepsilon_{ox} = 3.9\varepsilon_0 \qquad (3.42) \qquad \text{(the permittivity of the glass insulating layer)}$$

$\varepsilon_0 \approx 8.854 \times 10^{-14} F/cm$ *is the perittivity of free space*

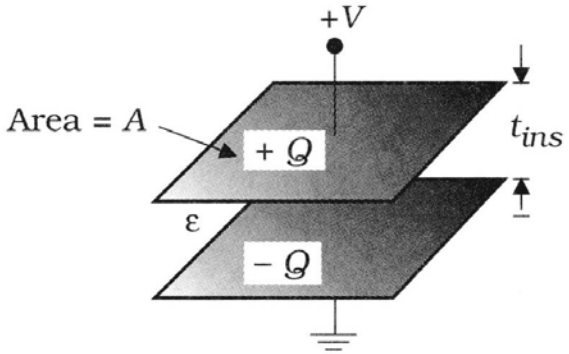


Figure 3.17 A parallel-plate capacitor

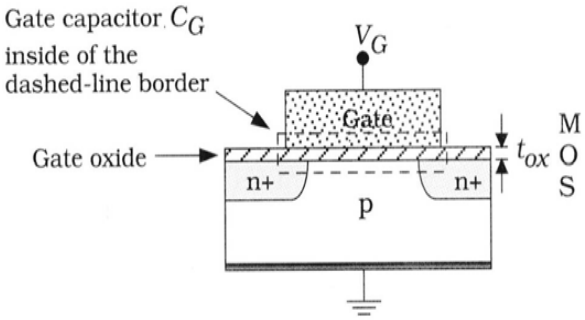


Figure 3.18 The gate capacitance in an n-channel MOSFET

Channel for NMOS Current Flow

- Applying a positive voltage to the gate as in Figure 3.19 (b)
  - » The capacitive MOS structure induces a layer of negatively charged electrons underneath the gate oxide → decided by the *threshold voltage*

$$Q_c = -C_G(V_G - V_{Tn}) \qquad (3.47) \qquad \text{(channel charge)}$$

$$\Rightarrow I = \frac{|Q_c|}{\tau_t}\ C/\text{sec} \qquad (3.48) \qquad \text{(channel current)}$$

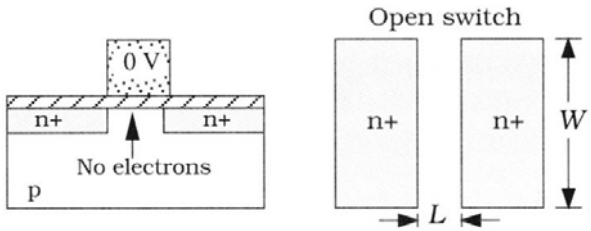
$$\text{Where } \tau_t = \frac{L}{v} \qquad (3.49) \qquad \text{(transit time)}$$

$$\Rightarrow I \approx \frac{C_G}{(L/v)}(V_G - V_{Tn}) = vC_{ox}W(V_G - V_{Tn}) \qquad (3.50)$$

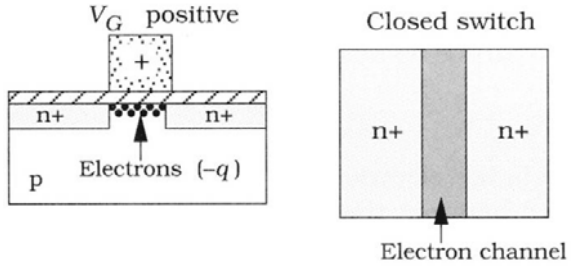
$$\text{Where } v = \mu_n E \qquad (3.51) \qquad \text{(the velocity of a charged particle moving in a FET)}$$

$$\Rightarrow E = \frac{V}{L} \qquad (3.52) \qquad \text{(electric field between n+)}$$

$$\Rightarrow I = \mu_n C_{ox}\left(\frac{W}{L}\right)(V_G - V_{Tn})V \qquad (3.53)$$



(a) Zero gate voltage



(b) Positive gate voltage

Figure 3.19 Controlling current flow in an nFET

Linear Resistance R_n of the n-device

- The linear resistance R_n of the device is

Note: $\Rightarrow I = \mu_n C_{ox}\left(\frac{W}{L}\right)(V_G - V_{Tn})V$

$$R_n = \frac{V}{I} = \frac{1}{\beta_n (V_G - V_{Tn})} \qquad (3.54)$$

$$Where\ \beta_n = \mu_n C_{ox}\left(\frac{W}{L}\right) \qquad (3.55)$$

(device transconductance A/V^2)

nFET open: R → ∞
nFET closed: R → R_n

- In fact, MOSFETs are intrinsically non-linear devices

$$R_n = R_{c,n}\left(\frac{L}{W}\right) \qquad (3.56)$$

$$Where\ R_{c,n} = \frac{1}{\mu_n C_{ox}(V_G - V_{Tn})} \qquad (3.57)$$

Outline

- An Overview CMOS Fabrication
- Integrated Circuit Layers
- MOSFETs
- CMOS Layers
- Designing FET Arrays

N-well Process

- CMOS fabrication process: In simplest terms, this refers to the sequence of steps that we use to take a bare “wafer” of silicon to the finished form of an electronic integrated circuit

  - p-substrate
  - n-well
  - n+ (nFET drain/source)
  - p+ (pFET drain/source)
  - gate oxide
  - gate (polysilicon)

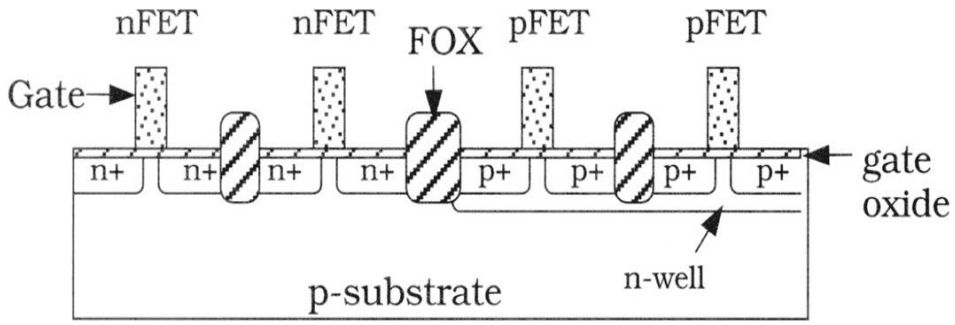


Figure 3.23 MOSFET layers in an n-well process

Fabrication Process

- Modern processes tend to allow for five or more metal interconnect layers to ease the problem of massive wiring in complex circuits

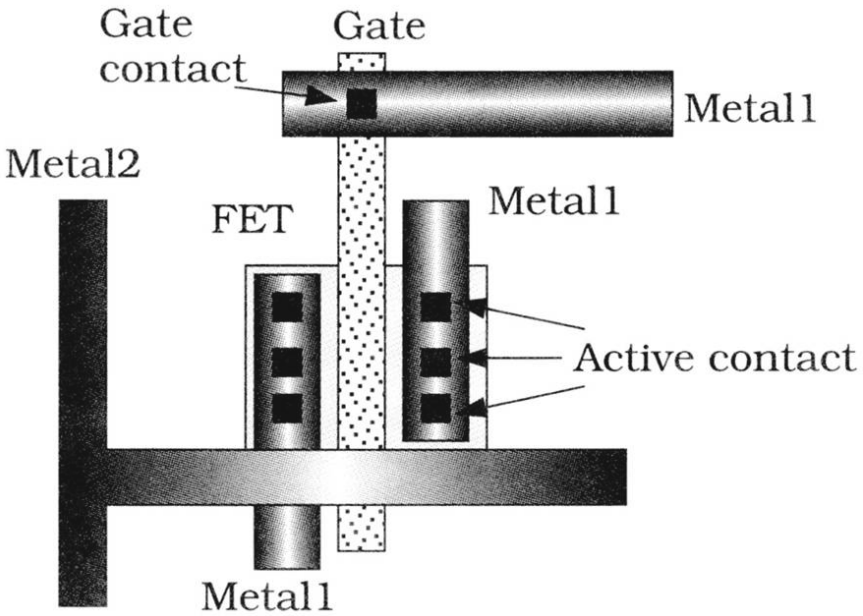


Figure 3.26 Interconnect layout example

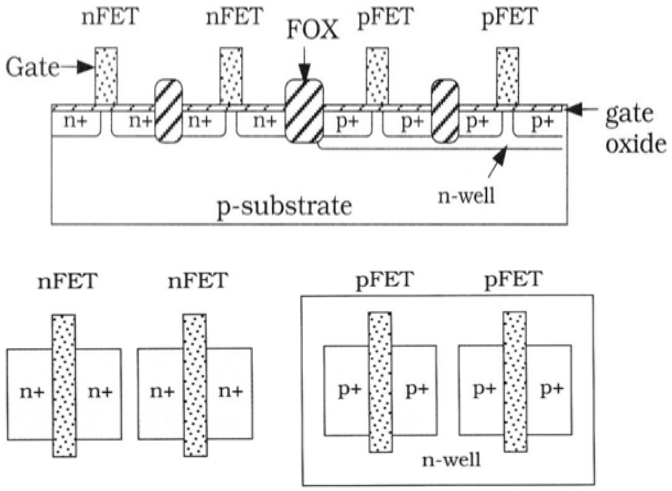


Figure 3.24 Top view FET patterning

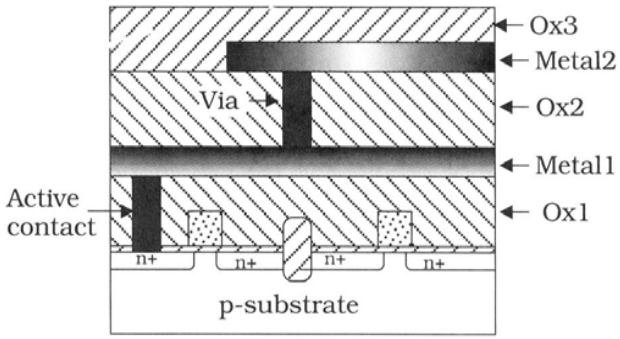


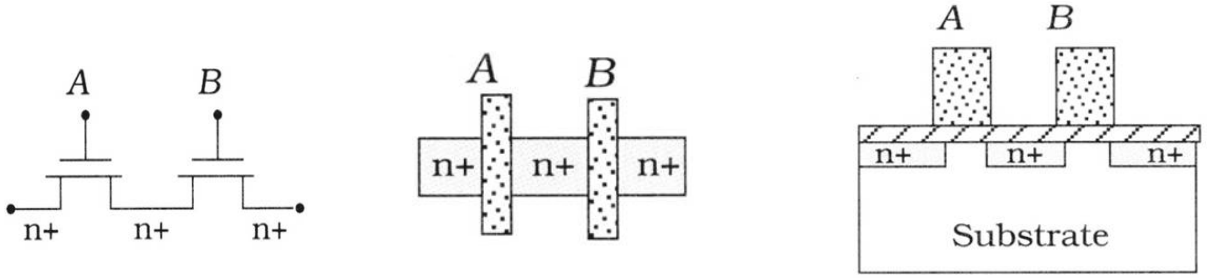
Figure 3.25 Metal interconnect layers

Outline

- An Overview CMOS Fabrication
- Integrated Circuit Layers
- MOSFETs
- CMOS Layers
- Designing FET Arrays

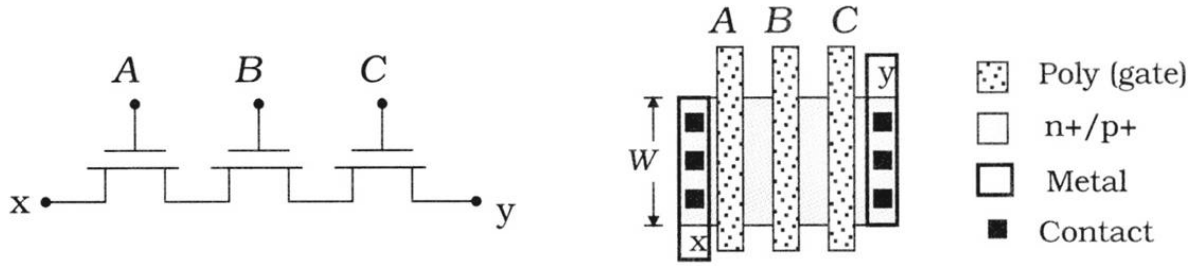
CMOS Switching Networks (1/2)

- CMOS logic gates are switching networks that are controlled by the input variables



(a) Schematic (b) Surface pattern (c) Side view

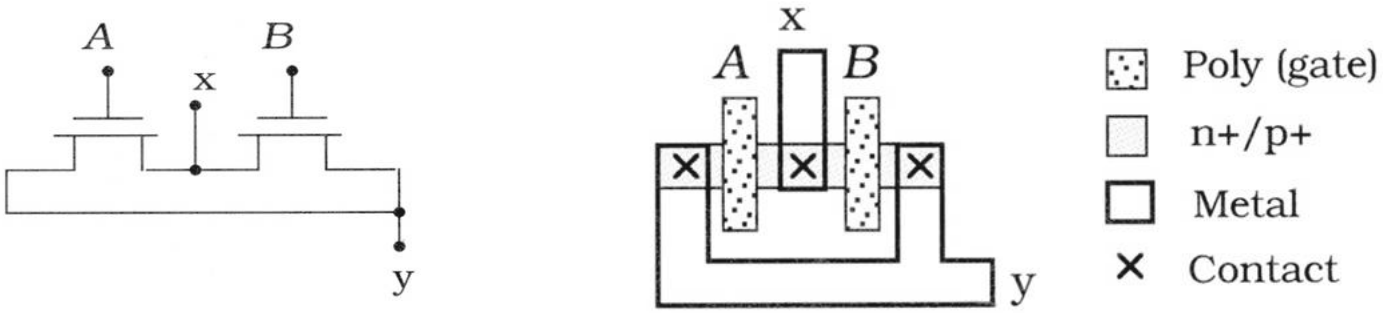
Figure 3.27 Silicon patterning for two series-connected nFETs



(a) Schematic (b) Surface pattern

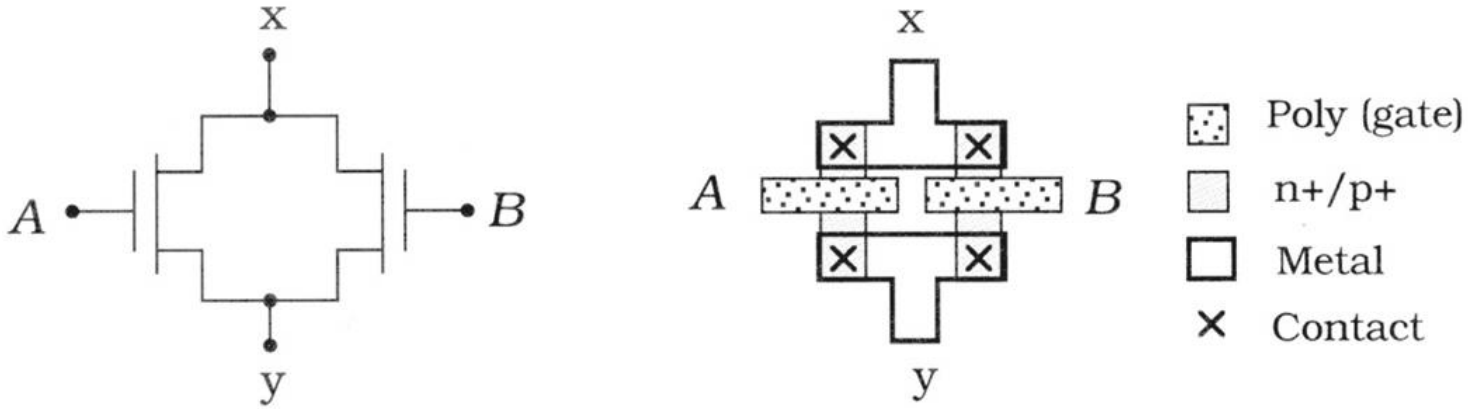
Figure 3.28 Three series-connected nFETs

CMOS Switching Networks (2/2)



(a) Schematic (b) Surface pattern

Figure3.29 Parallel-connected FET patterning

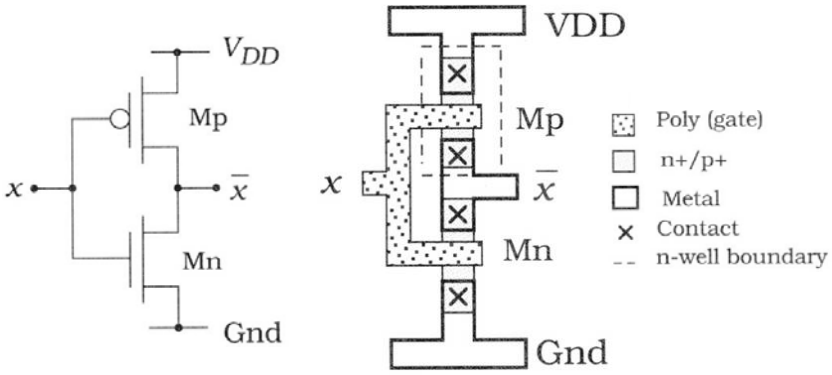


(a) Schematic (b) Surface pattern

Figure 3.30 Alternate layout strategy for parallel FETs

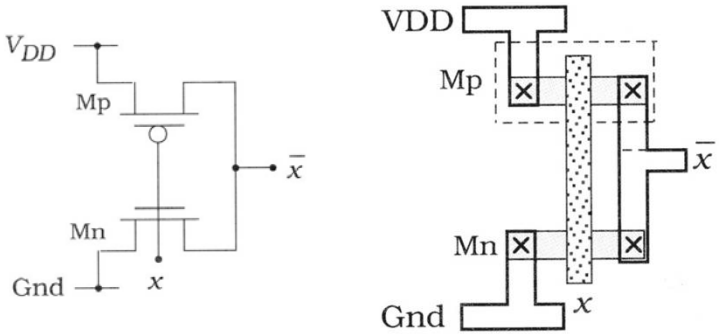
Basic Gate Designs

- Design note
  - Both the power supply (VDD) and ground (Gnd) are routed using the Metal layer
  - n+ and p+ regions are denoted using the same fill pattern. The difference is that pFETs are embedded within an n-well boundary
  - Contacts are needed from Metal to n+ or p+ since they are at different levels in the structure



(a) Circuit (b) Layer Patterning

Figure 3.31 Translating a NOT gate circuit to silicon



(a) Circuit (b) Layer Patterning

Figure 3.32 Alternate layout for a NOT gate

Basic Gate (1 / 3)

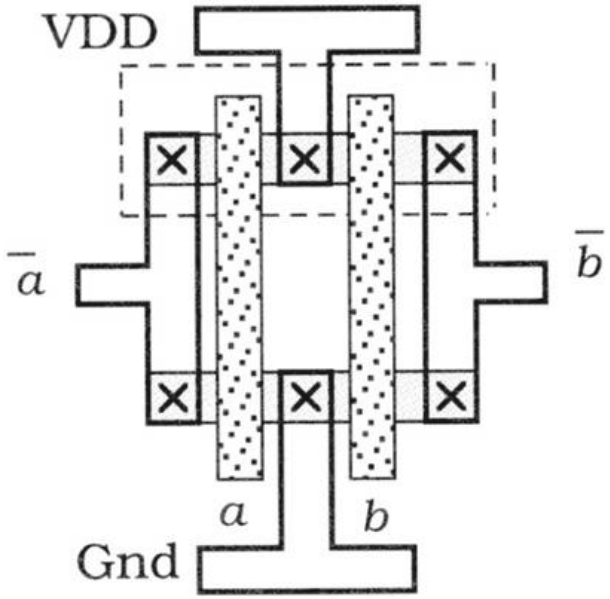
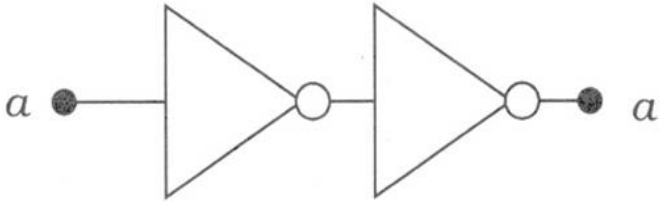
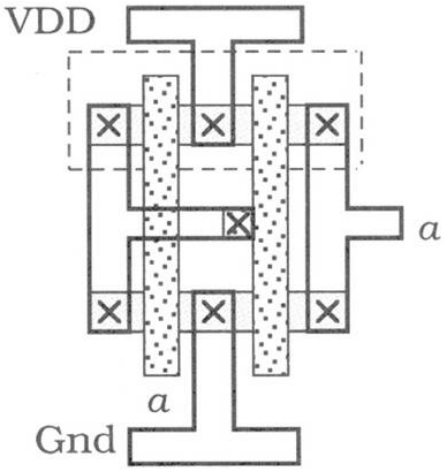


Figure 3.33 Two NOT gates that share power supply and ground



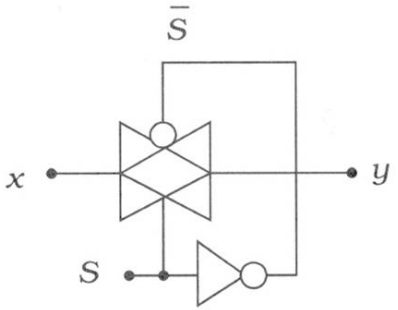
(a) Logic diagram



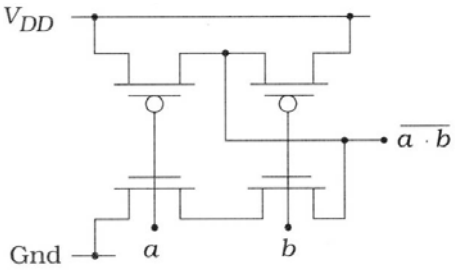
(b) Layout

Figure 3.34 Non-inverting buffer

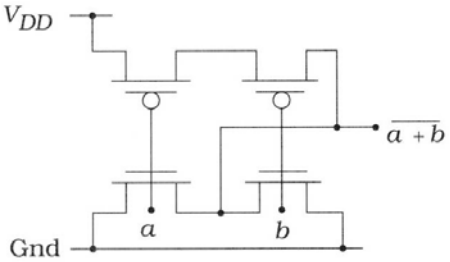
Basic Gate (2 / 3)



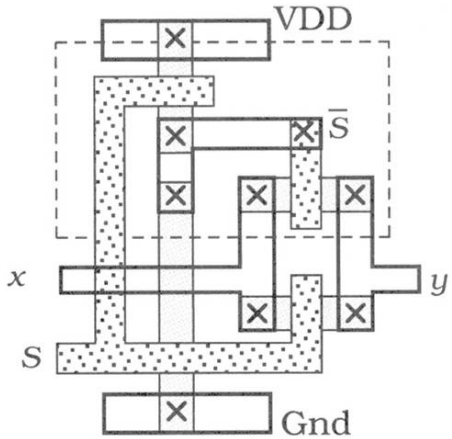
(a) Logic diagram



(a) Circuit

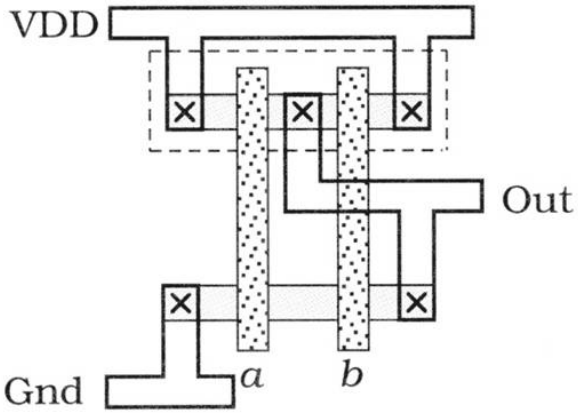


(a) Circuit



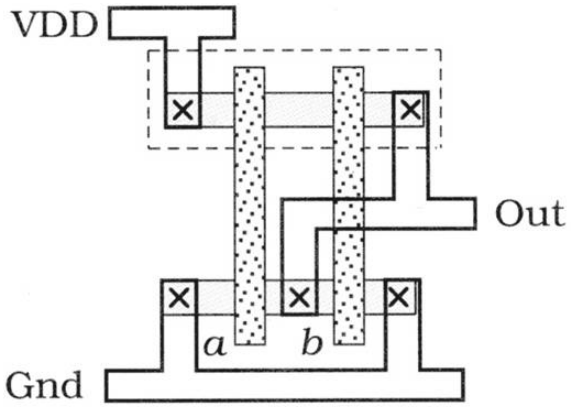
(b) Layout

Figure 3.35 Layout of a transmission gate with a driver



(b) Layout

Figure 3.36 NAND2 gate design



(b) Layout

Figure 3.37 NOR2 gate design

Basic Gate (3 / 3)

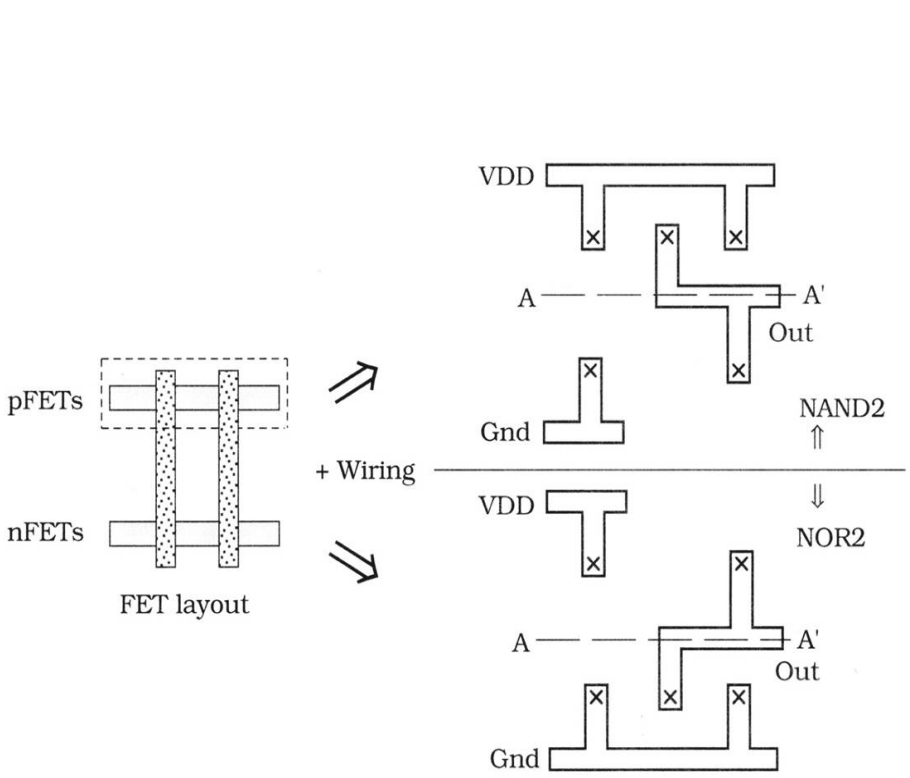
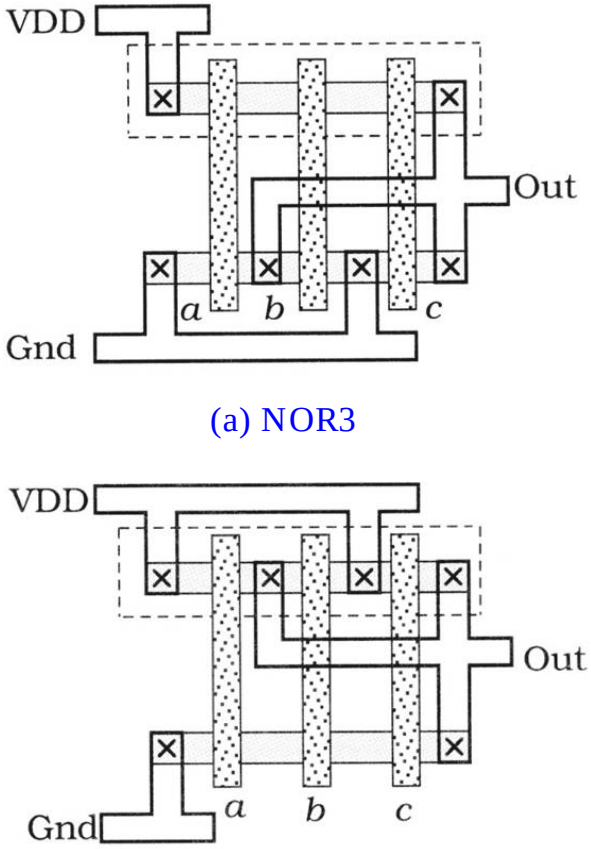


Figure 3.38 NAND2-NOR2 layout comparison



(a) NOR3

(b) NAND3

Figure 3.39 Layout for 3-input gates

Complex Logic Gate (1 / 2)

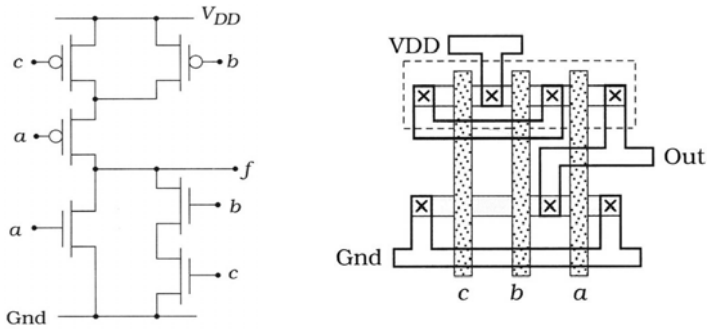
- Figure 3.40 shows the function
  - » The signal placement order is critical to obtaining the logic output

$$f = \overline{a + b \cdot c} \qquad (3.68)$$

- Dual logic: In physical circuit, suppose that we flip the metal wiring pattern around an imaginary horizontal line. We will get

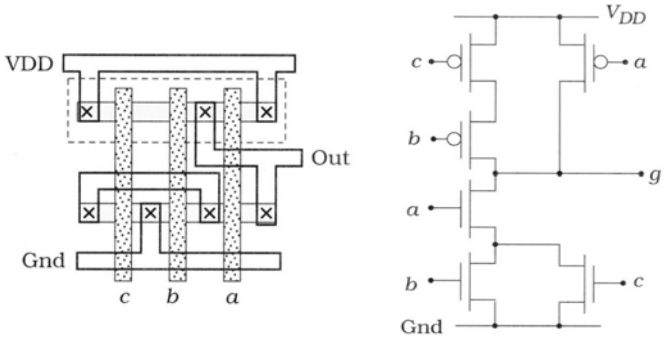
$$g = \overline{a \cdot (b + c)} \qquad (3.69)$$

- This is the same relationship that we found for the NOR-NAND gates



(a) Circuit (b) Patterning

Figure 3.40 Extension of layout technique to a complex logic gate

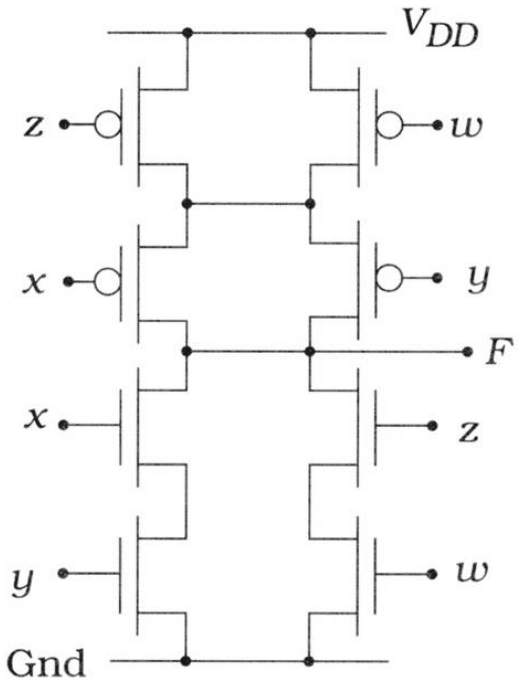


(a) Patterning (b) Circuit

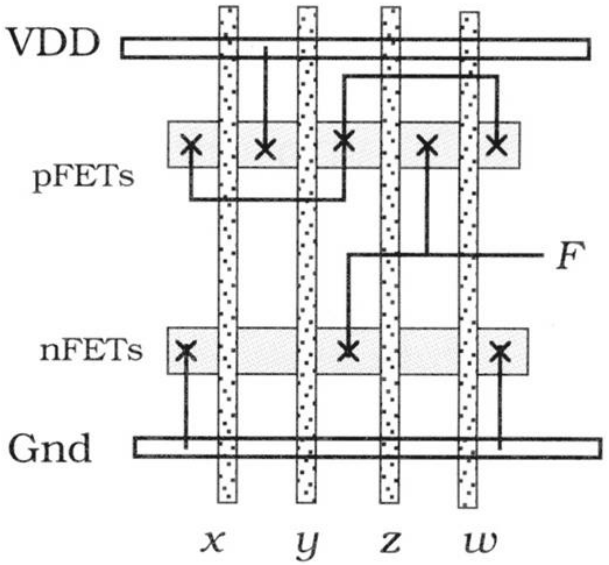
Figure 3.41 Creation of the dual network

Complex Logic Gate (2 / 2)

- A logical function $F = \overline{x \cdot y + z \cdot w}$ (3.70)



(a) Circuit



(b) Layout wiring

Figure 3.42 A general 4-input AOI gate

General Discussion

- A basic techniques that it was possible to share n+ or p+ regions among several transistors
  - » Randomly placed polygons should be avoided
  - » It can reduce the area and wiring complexity
  - » A power supply (VDD), a ground (VSS) connection, and pFETs will be embedded in n-wells around VDD
  - » nFETs are closer to the ground rail
- One approach to layout is based on the concept of simple *stick diagram*
  - » It often used to perform quick layouts or
  - » To study large complex routing problems
- Moreover, any CMOS circuit can be translated into an *equivalent graph* consisting of edges and vertices

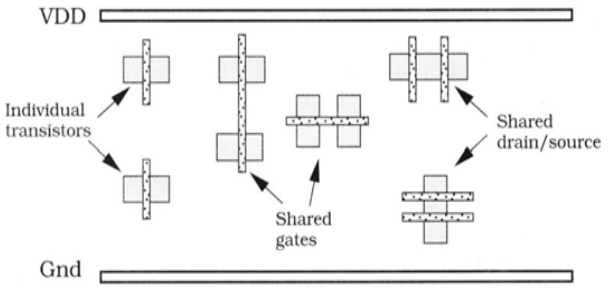


Figure 3.43 General gate layout geometry

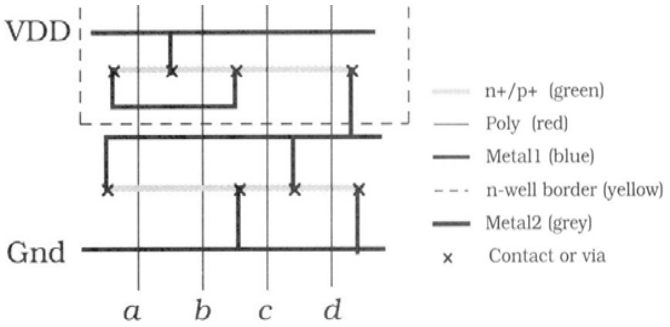


Figure 3.44 Basic stick layout diagram

Euler Graph (1 / 2)

- *Euler graph*
  - The drain and source nodes x and y of the transistor translate to connection nodes called **vertices**
  - An **edge** that corresponds to the signal flow path
  - If it is possible to trace the entire graph without passing over an edge more than once, then it is possible to use common n+/p+ regions for nFETs/pFETs
- *Euler path*
  - It is used to construct the Euler graph
  - If an Euler path cannot be found, then it means that it is not possible to use FET chains to build the circuit

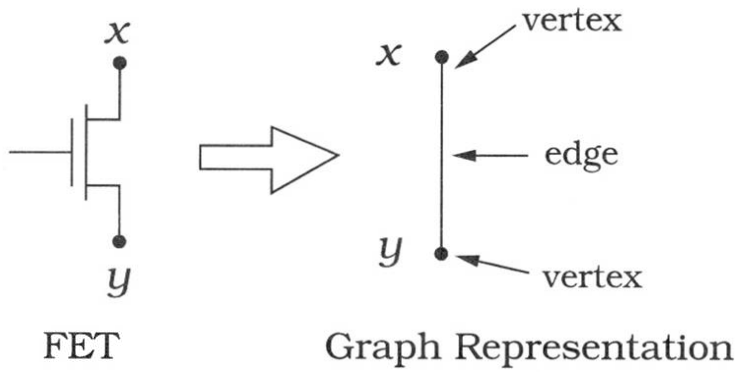
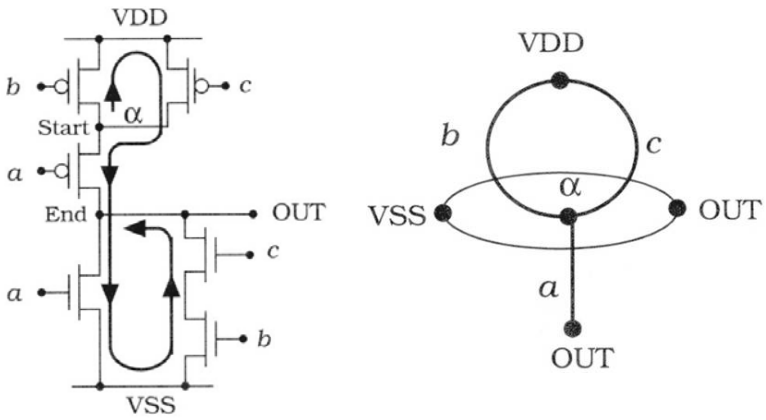


Figure 3.45 Representation of a FET in graph theory



(a) Circuit with path (b) Euler graph

Figure 3.46 Construction of an Euler graph

Euler Graph (2 / 2)

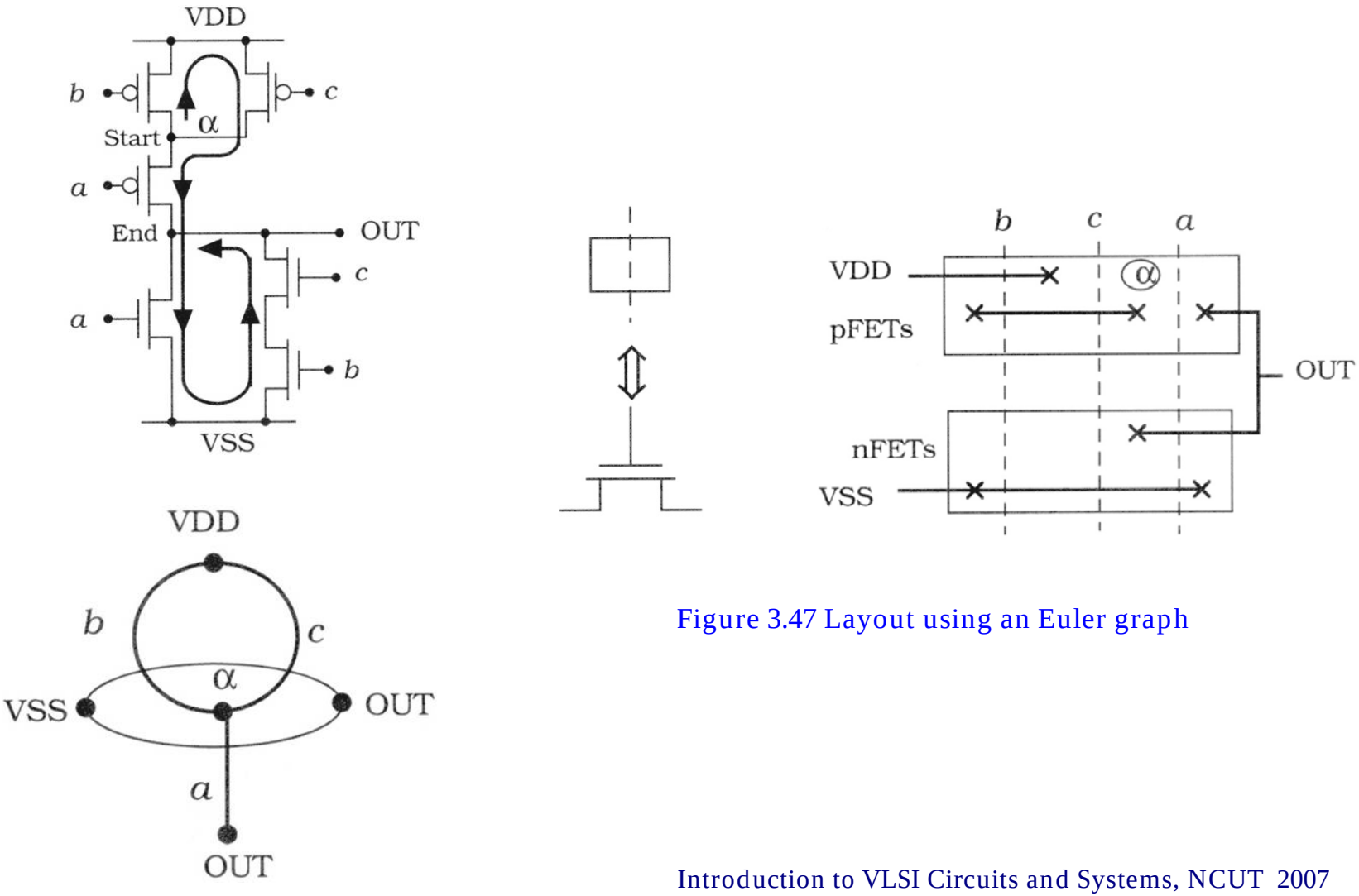


Figure 3.47 Layout using an Euler graph