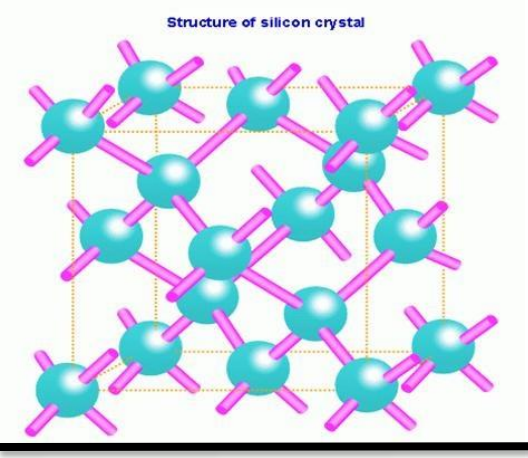
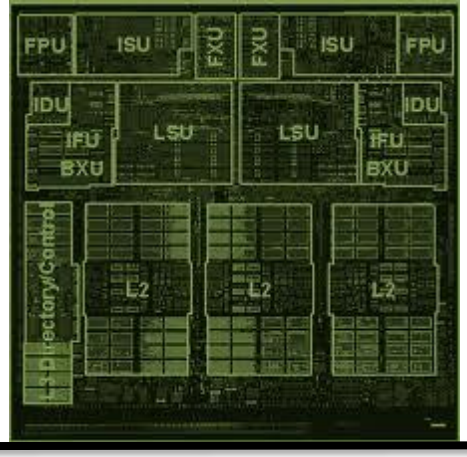
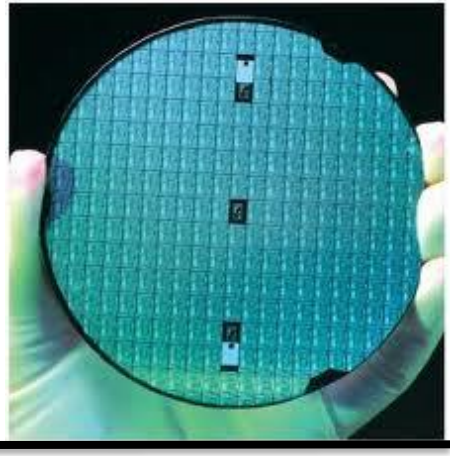






ECE 122A

VLSI Principles

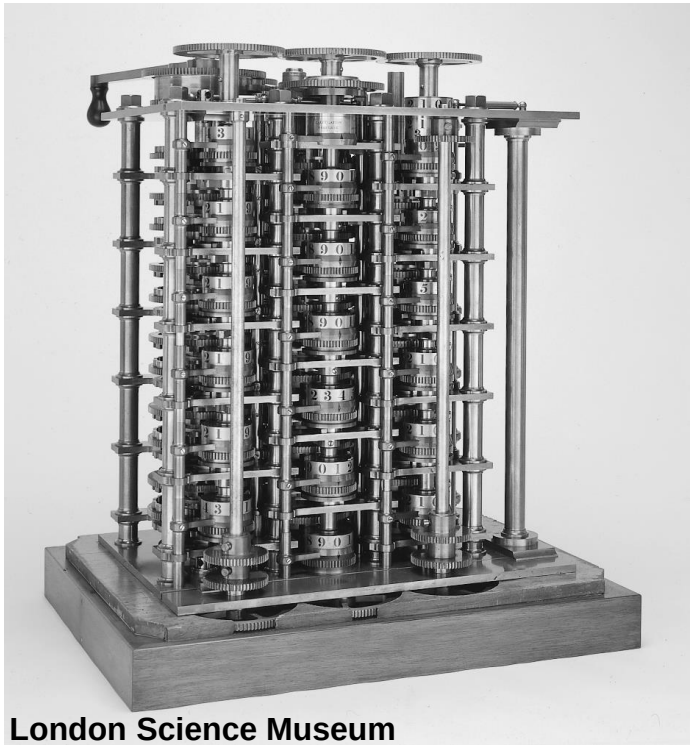
Lecture 2

Prof. Kaustav Banerjee
Electrical and Computer Engineering
University of California, Santa Barbara
E-mail: kaustav@ece.ucsb.edu

History of Computing…..The first computer

Charles Babbage

(1791-1871)



London Science Museum

The Babbage Difference Engine (1832)

25,000 parts

cost: £17,470

A mechanical digital calculator…

Mechanical computing devices

Used decimal number system

Could perform basic arithmetic operations

Even store and execute

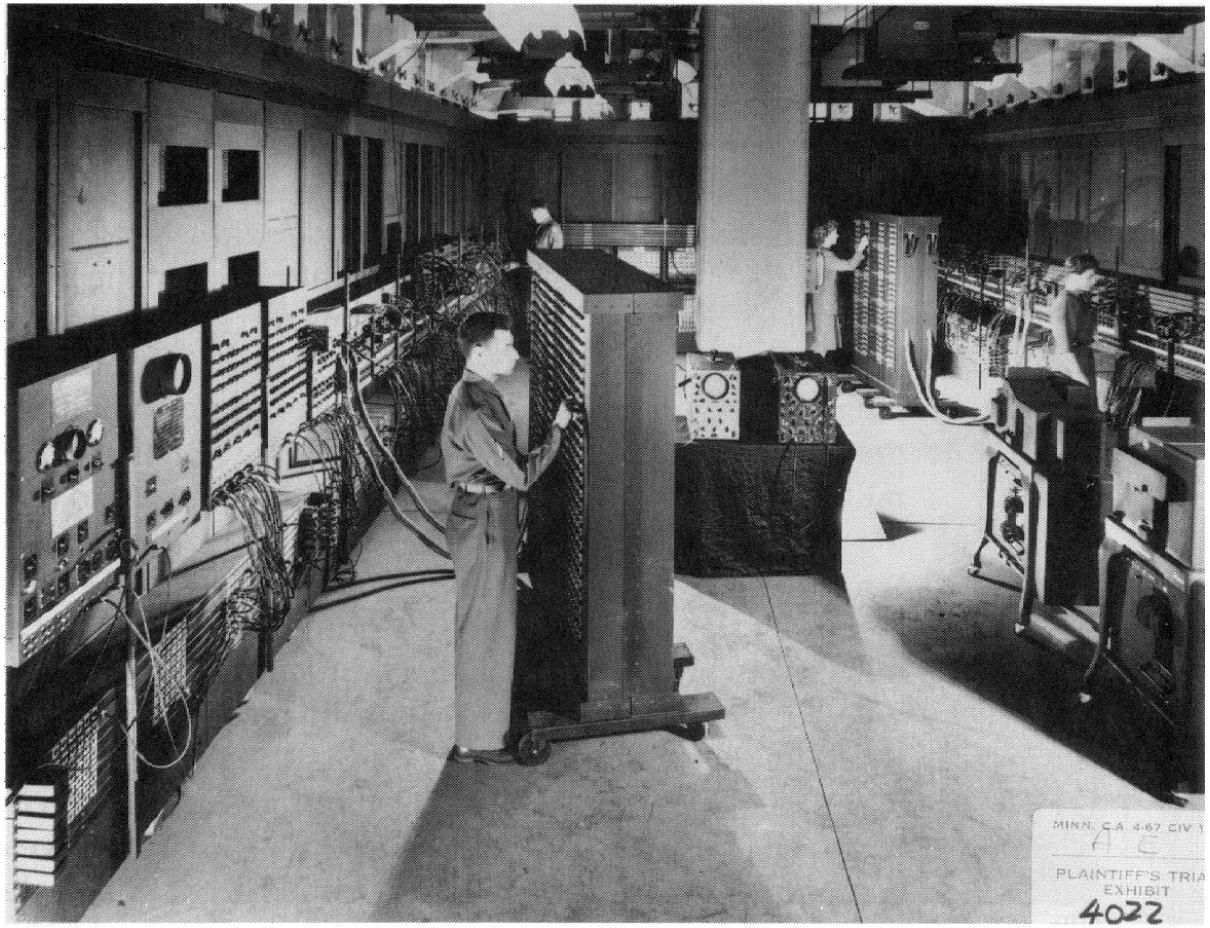
Problem: Too complex and expensive!

Further Reading….

- *http://en.wikipedia.org/wiki/Difference_engine*
- *How Computers Do Math*

  (ISBN: 0471732788) *Wiley*, Clive Maxfield and Alvin Brown.

ENIAC - The first electronic computer (1946)



Vacuum tube based computer…

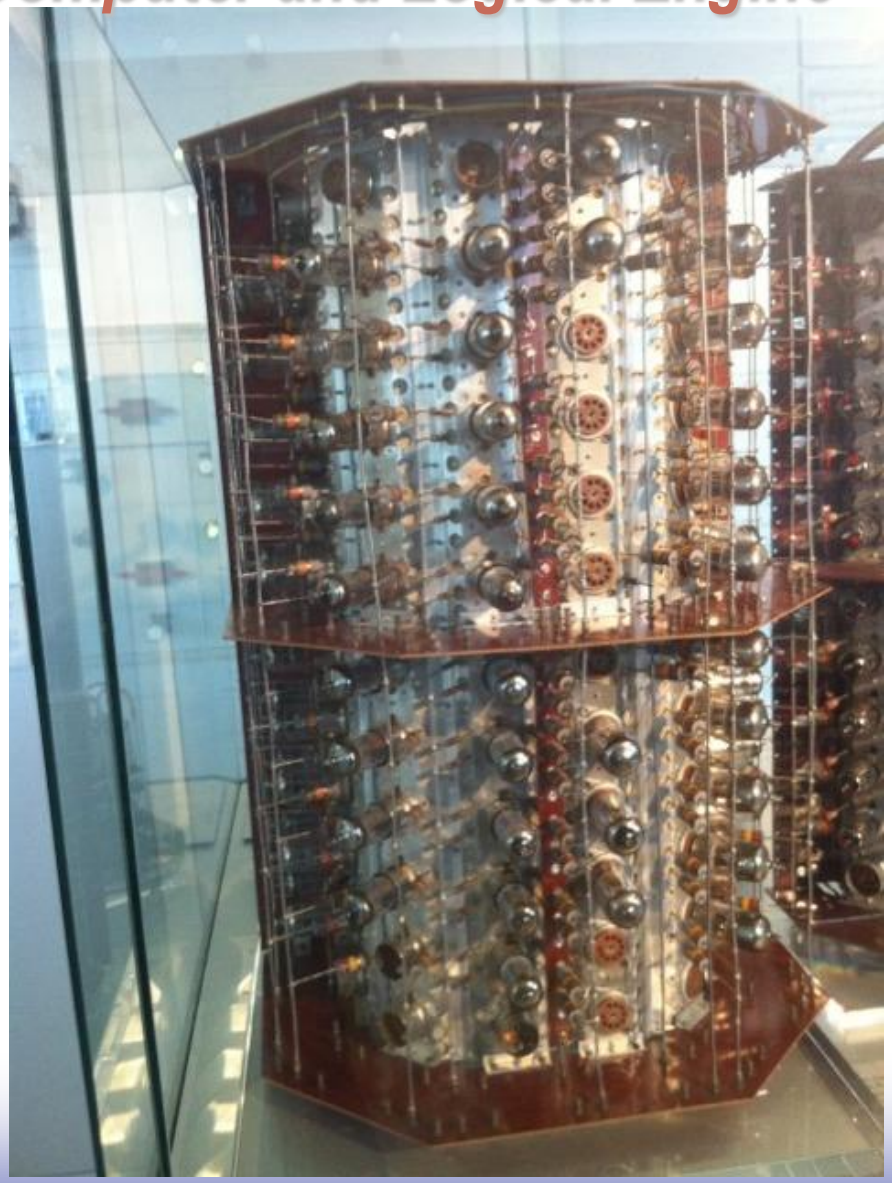
For Military applications…

80 ft long, 8.5 ft high, several ft wide…

With ~18,000 vacuum tubes!

Problem: Reliability issues and excessive power consumption!

Parts of the ORACLE Computer –Oak Ridge Automatic Computer and Logical Engine---Oak Ridge National Lab (1950-54)



Used vacuum tubes, transistors, and diodes.

Addition time: 70 microseconds
Multiplication time: 370-590 microseconds
Division time: 590 microseconds

These times include the storage access time, which was about 62 microseconds.

Credit: Deutsches Museum, Munich, Germany

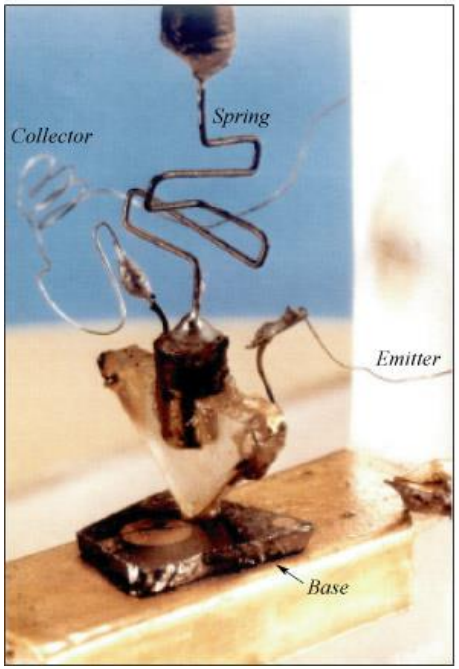
The Transistor Revolution…

http://www.nobelprize.org/educational/physics/integrated_circuit/history/

The first point contact transistor

William Shockley, John Bardeen, and Walter Brattain

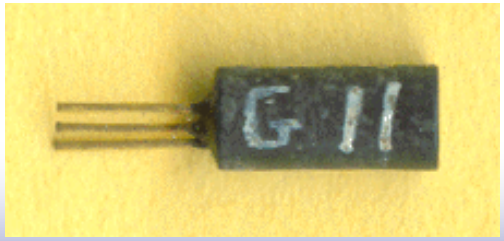
Bell Laboratories, Murray Hill, New Jersey (1947)

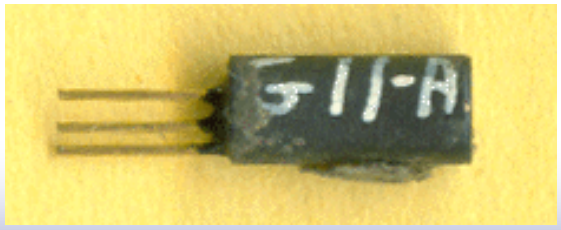


First transistor
Bell Labs, 1947
(Ge point contact-
Bipolar transistor)
Bardeen and
Brattian
-Nobel Laureates

BJT (1948)
Schockley
— Nobel Laureate

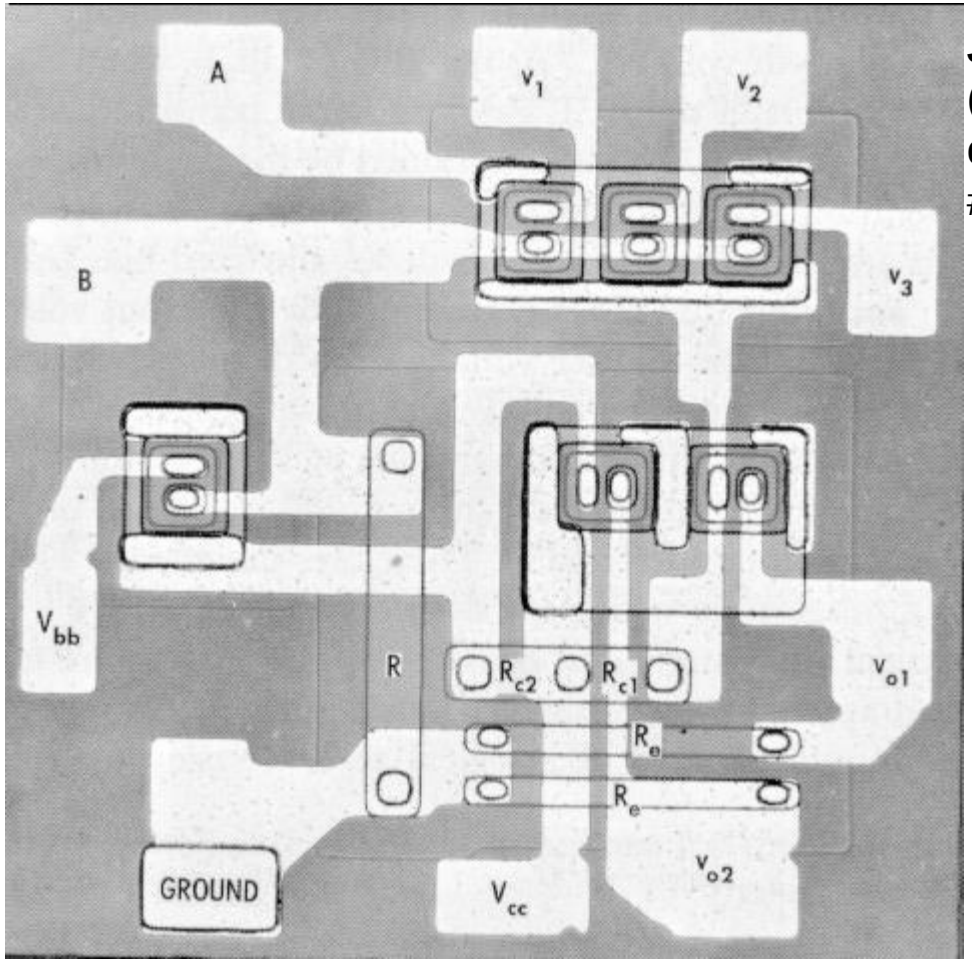
General Electric types G11 and G11A commercial point contact transistors





Transistor Size (1/8” OD X 3/8”)

The First Integrated Circuits



Jack Kilby, Texas Instruments (1958), the monolithic integrated circuit, or microchip (patent #3,138,743), **Nobel Prize in 2000**

Bipolar logic
Early 1960's (TTL, ECL)

TTLs offered higher integration density—composed largest fraction of semiconductor market until the 1980s

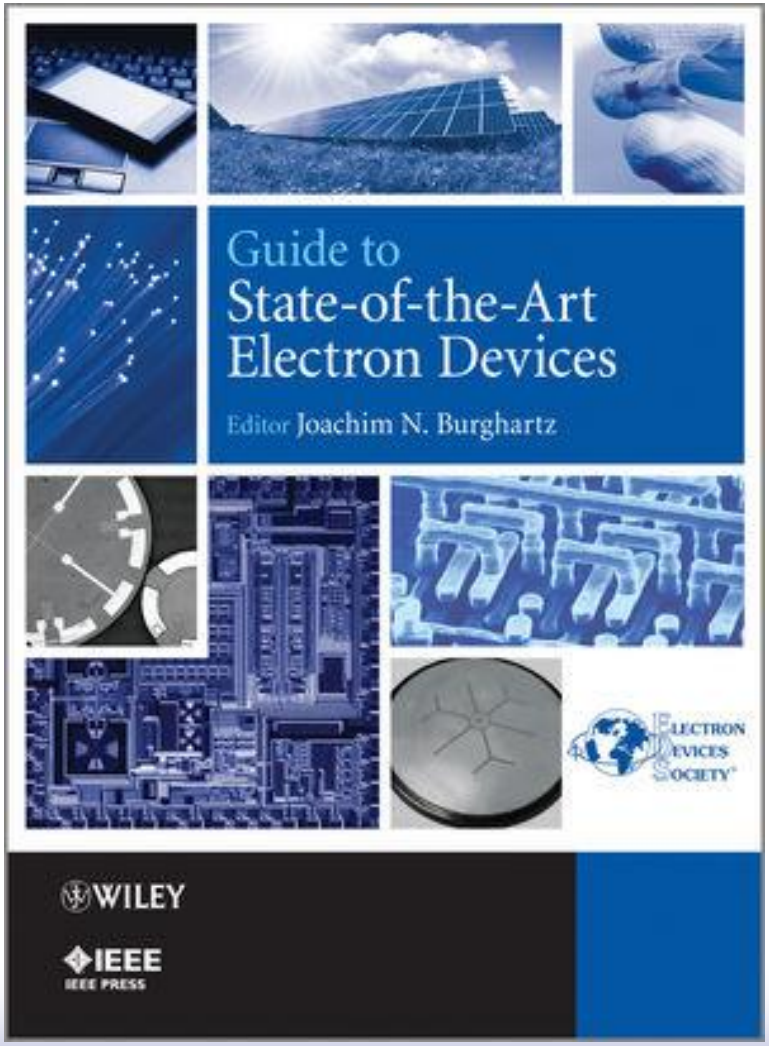
ECL 3-input Gate
Motorola 1966

Integrated Circuits (Early) History

- Invention of BJT (1948)
- First silicon transistor (1954)
- MOS transistor (1960)
- MOS integrated circuit (1962)
- DRAM cell (1968)
- Intel formed (1968) (Intel: short form of integrated electronics)
- AMD formed (1969)
- Microprocessor invented (1971)
- 32-bit microprocessors (1980)

For more historical perspectives….

http://www.ece.ucsb.edu/news/?i=4245



BJT vs FETs

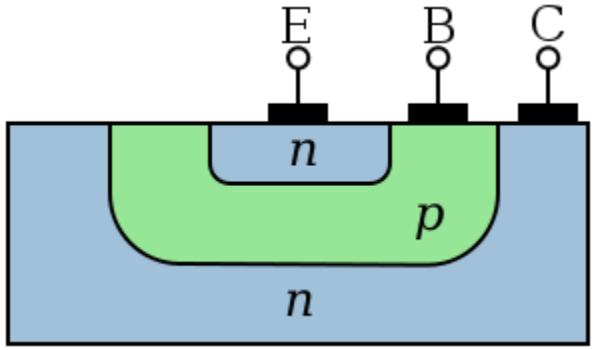
Both are based on basic properties of pn junctions

- A Bipolar Junction Transistor (BJT) is a 3 terminal device
- Uses the injection of minority carriers (under a forward bias)
- A BJT is a "bipolar" device (both electrons are holes are involved in its operation)
- It is an asymmetric device….why?

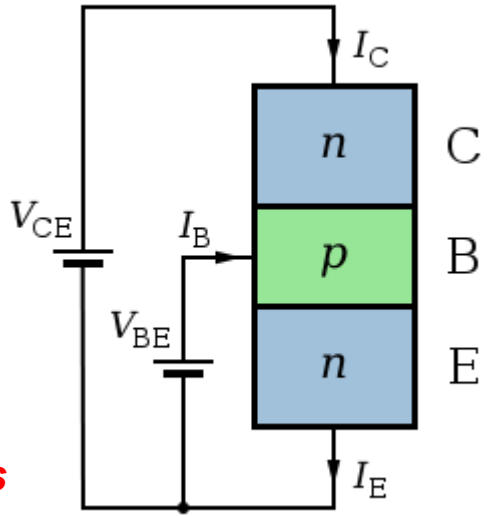
- A Field Effect Transistor (FET) is also a 3 terminal device (plus a substrate terminal)
- A FET is a "unipolar" device (majority carrier only)
- It is based on controlling the depletion width of a--- junction (JFET) or a Schottky Barrier (MESFET) through a control (gate) voltage

Power has been the main driver for various technologies…..vacuum tubes, BJT, PMOS, NMOS, CMOS…..???

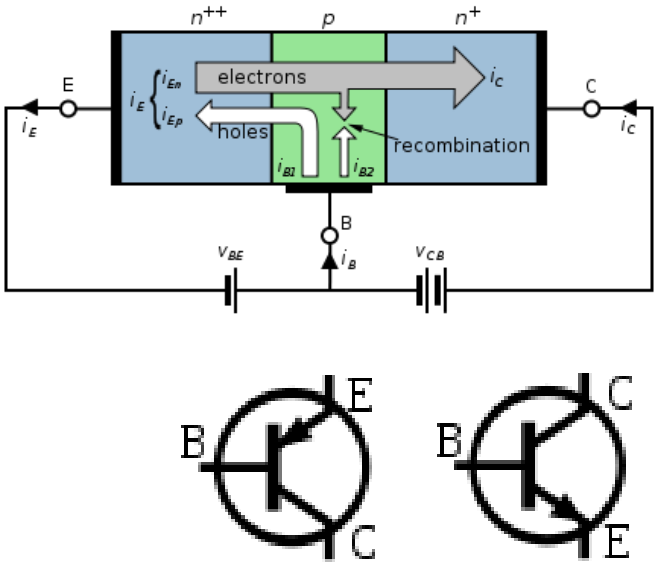
Recalling the BJT…..



Schematic cross-sectional view of a planar NPN bipolar junction transistor…emitter is more heavily doped than collector….depth of "p" region must be smaller than the diffusion length of electrons



Basic structure of an NPN BJT



Schematic symbols for PNP and NPN type BJTs

- **Both electrons and holes are involved in the operation…..hence the name "bipolar"**
- **For NPN BJT: electrons are injected from a high-concentration emitter (n++) into the p-type base…..where they are minority carriers that diffuse toward the collector (n+)…..hence BJTs are minority-carrier devices**
- **Quiescent power (drawn even in idle circuits) due to small base current limits large-scale integration**

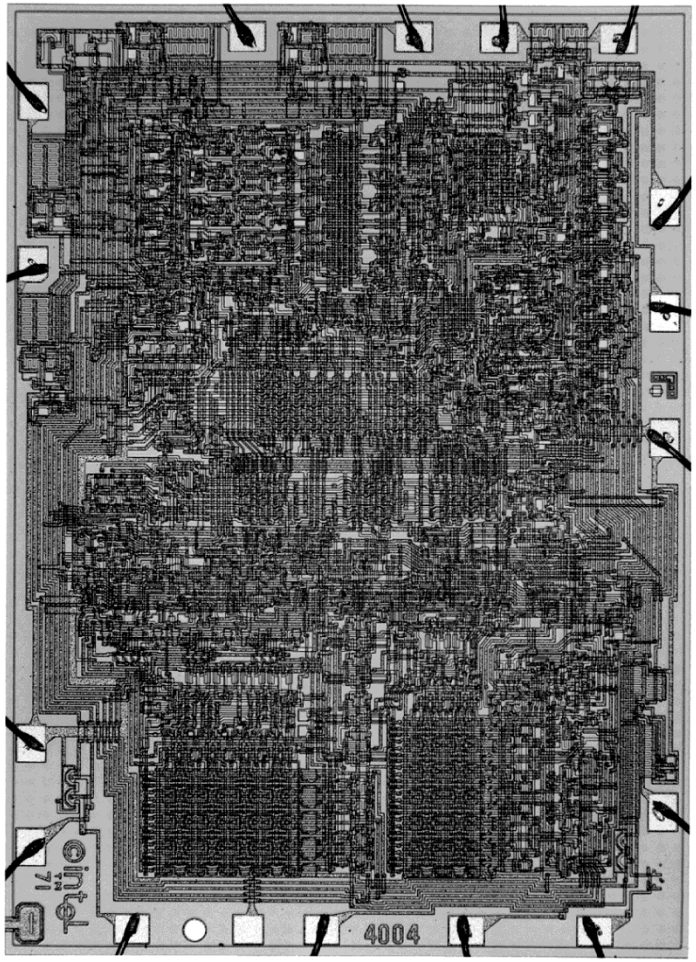
Other FETs…..

- If the metal gate electrode is separated from the semiconductor by an insulator– metal-insulator-semiconductor FET (MISFET)
- Also called an IGFET (insulated gateFET)
- Most commonly called as MOSFET (metal-oxide-semiconductor FET)
- FETs have high input impedance—since the control voltage is applied to a reverse biased junction or Schottky Barrier or across an insulator-----they are better suited (than BJTs) for controlled switching between conducting (ON) and non-conducting (OFF) states---therefore better for digital circuit implementation….low-power
- FETs are also more integrable…processing perspective

Strange History of FETs…..

- Actually FET was invented (proposed) in 1925!! (by Julius Lilienfeld)
- But…..it never worked….Why?.....
  - Role of surface defects….dangling bonds etc. causing large number of surface states….Fermi level pinning
  - Silicon dioxide….a key material for Silicon's success…..
- Bardeen and Brattain accidentally discovered the first (bipolar) transistor: the Ge point contact transistor---- while trying to experimentally demonstrate the FET
- First MOSFET demonstrated in 1960 by Kahng and Atalla
- First logic gates using MOSFETs (both n-type and p-type…thus Complementary-MOS-FET)— Frank Wanlass at Fairchild in 1963

Intel 4004 Micro-Processor



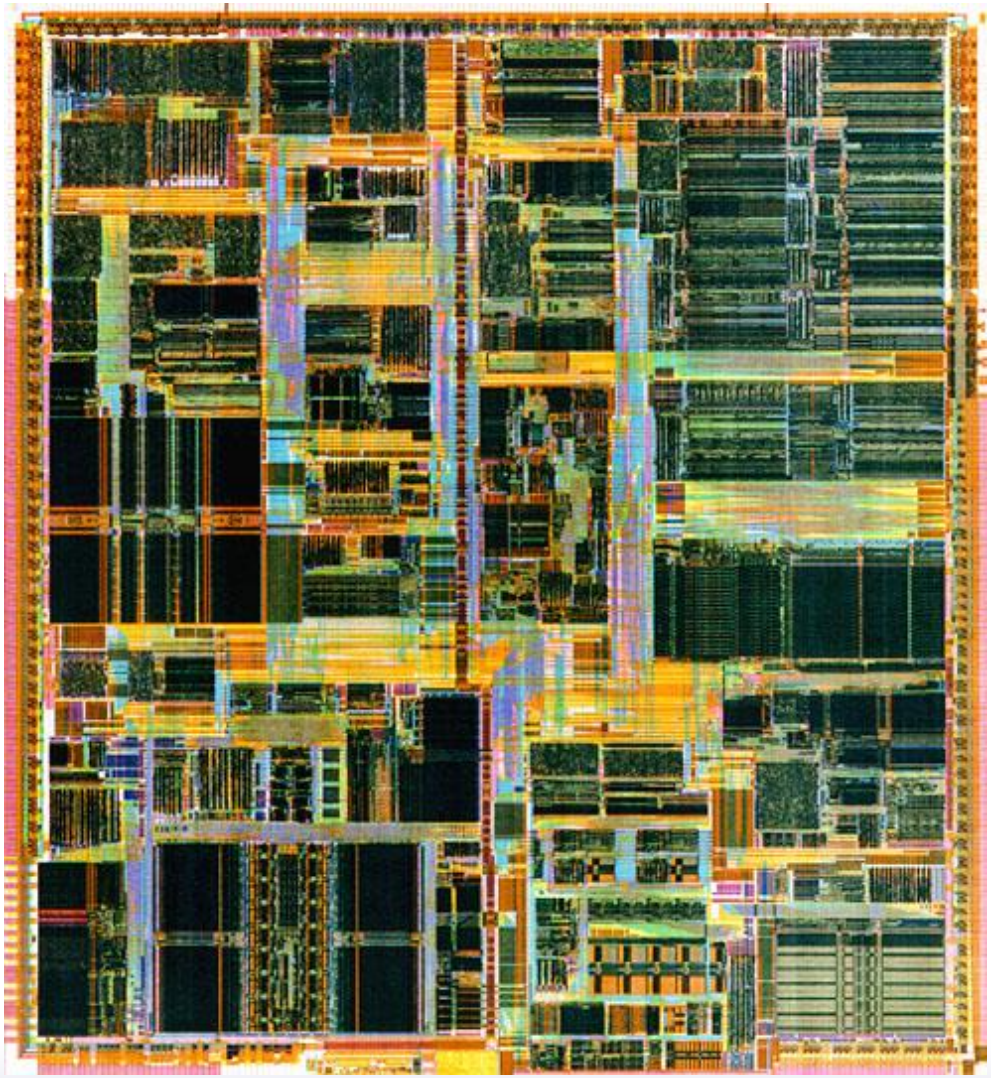
Nov 1971
2300 transistors
108 KHz operation
PMOS only (10 um process)

Note: *although NMOS was conceived in the early 1960s—application in wrist watches and portable electronics (of 1960s)---but device quality was poor, hence PMOS was used until 1974*

The 1971 4004 Microprocessor (10 um process) was PMOS only

---Intel 8080 (8-bit) Microprocessor (1974)
6 thousand transistors, 2 MHz
NMOS only (6 um process)

Intel Pentium (IV) Microprocessor



2003:
CMOS based
> 50 million transistors
> 3GHz operation

Economic Implications

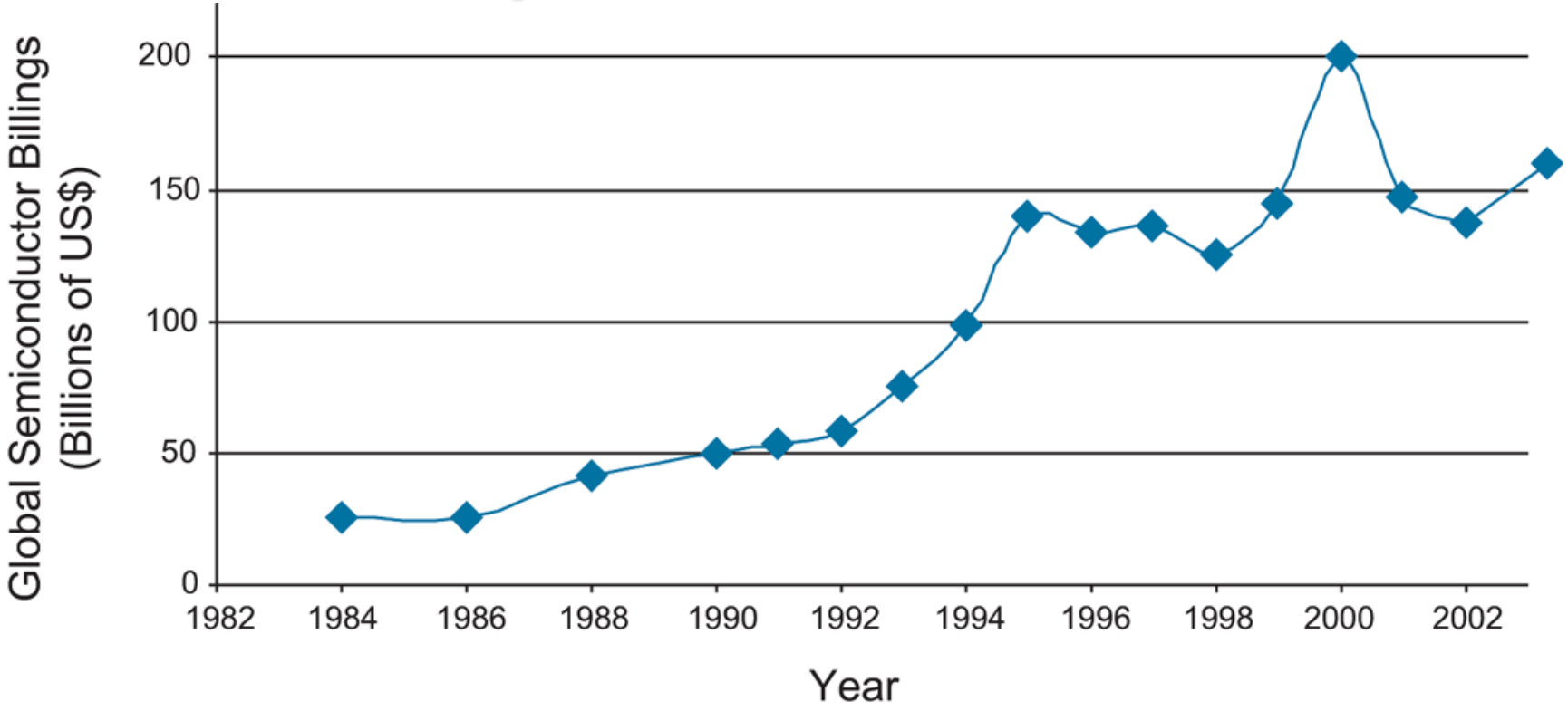


FIG 1.1 Size of worldwide semiconductor market

Source: Semiconductor Industry Association.

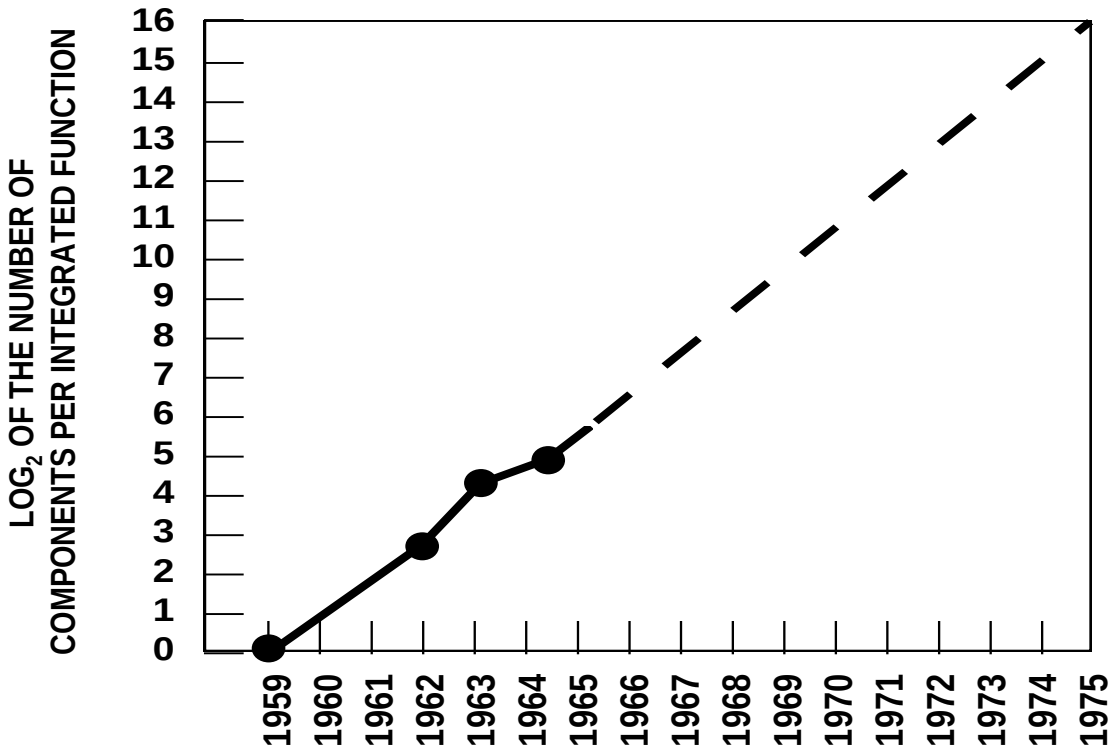
For more up to date info. go to the ***International Technology Roadmap for semiconductors (ITRS)*** *http://public.itrs.net/*

Moore's Law

- In 1965, Gordon Moore (co-founder of Intel) noted that the number of transistors on a chip doubled every 18 to 24 months

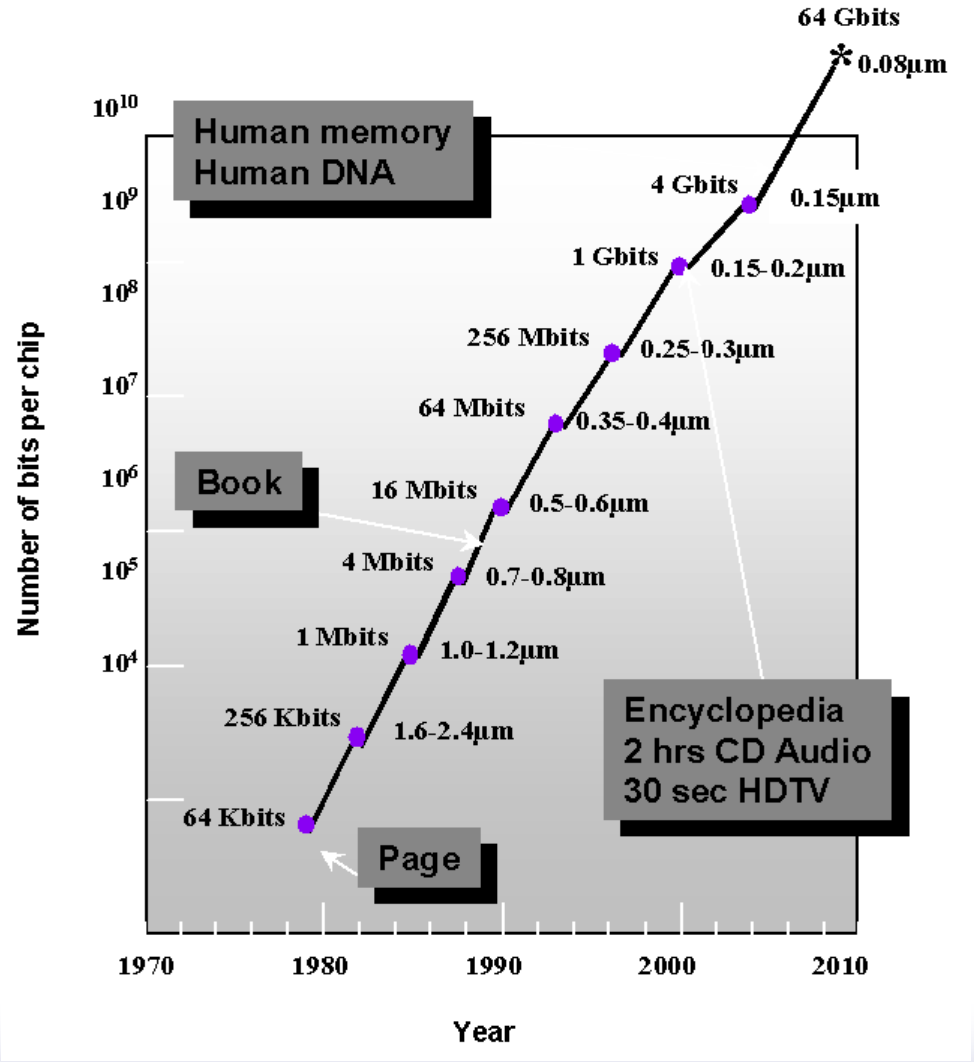
- He made a prediction that semiconductor technology will double its effectiveness every 18 (24) months

Moore's Law



Electronics, April 19, 1965.

Evolution in Complexity

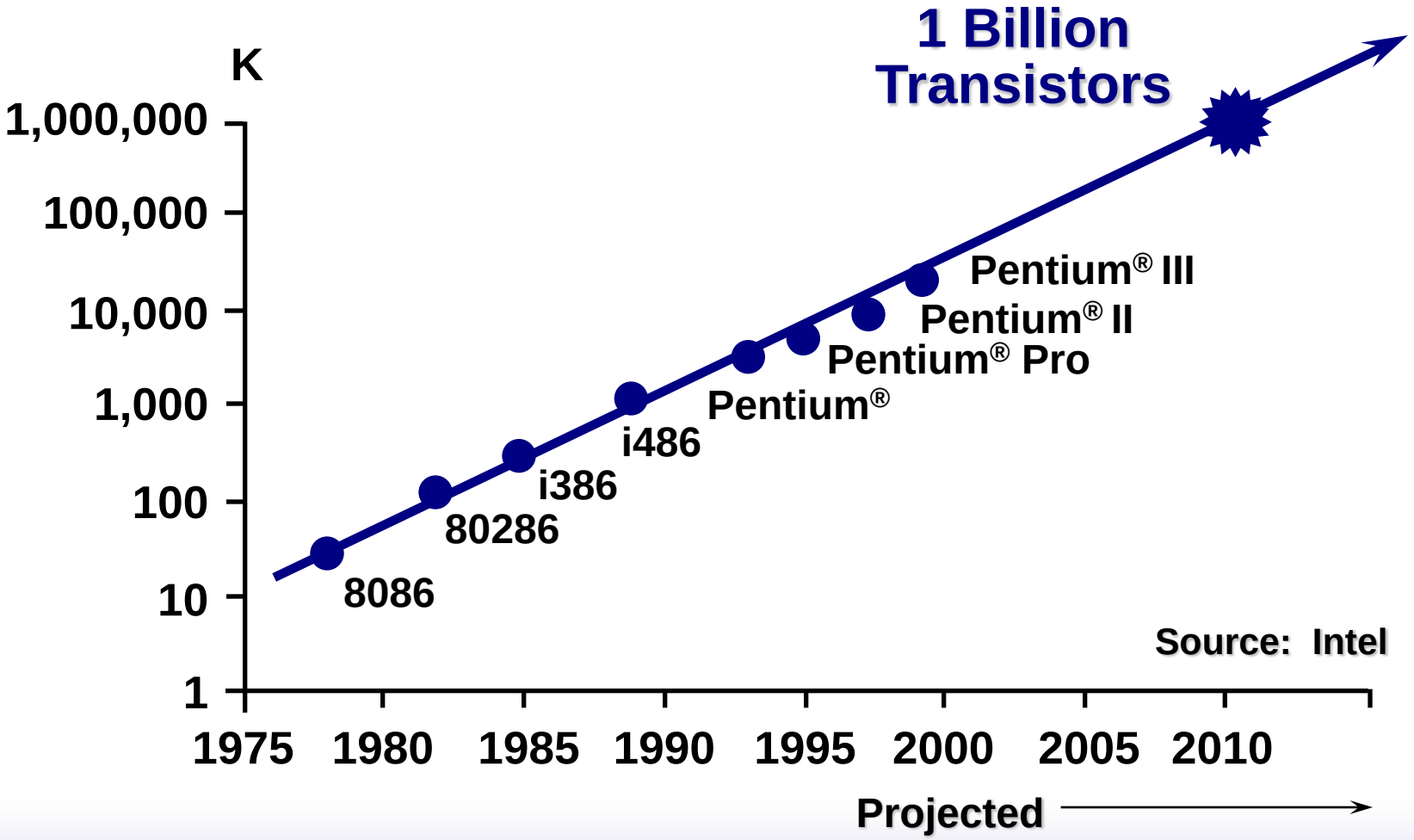


IC Classification

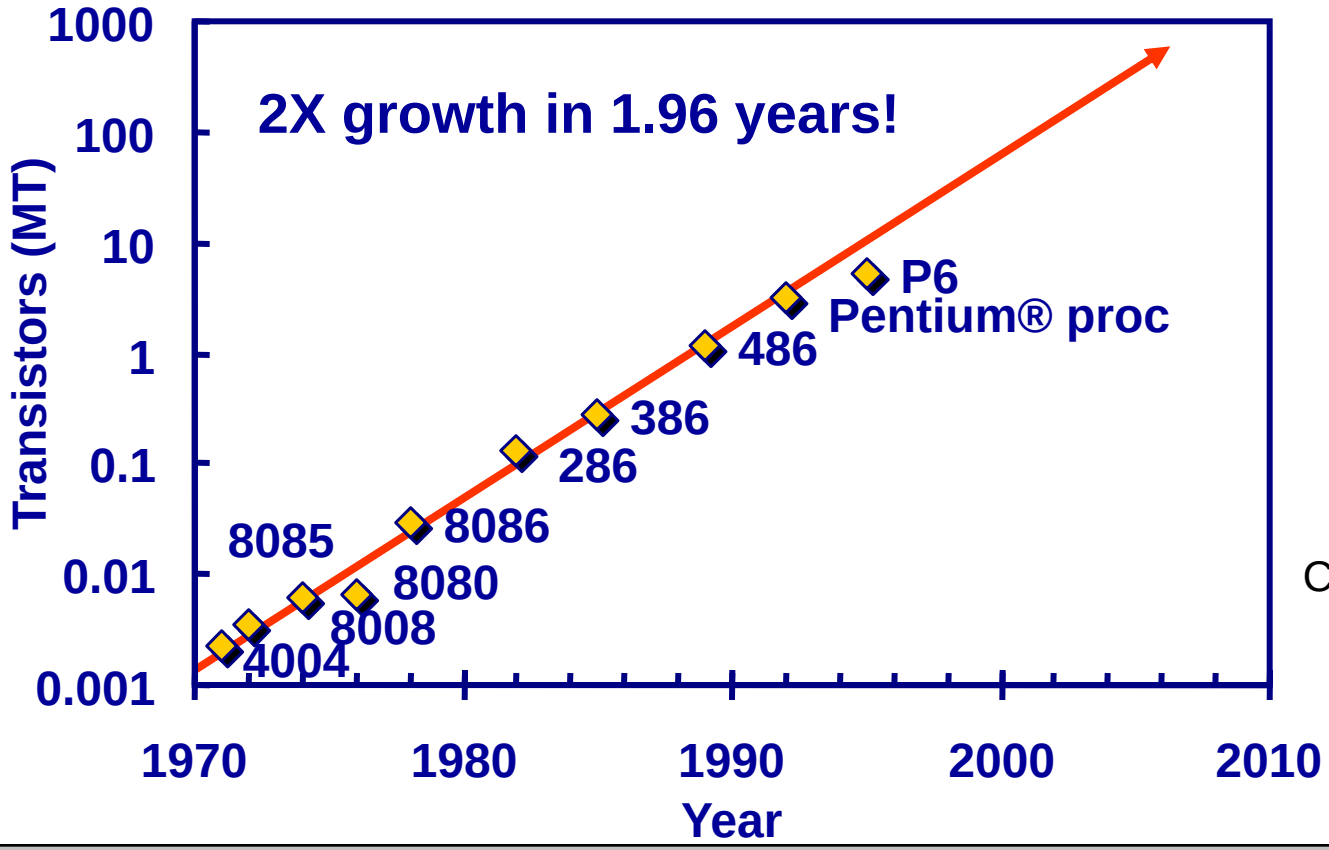
- Circuit size (transistor count)
- Circuit technology (BJT, BiCMOS, NMOS, CMOS)
- Design style
  - standard cell
  - gate array
  - custom
- Size classification (historical)

<100	SSI	1963
100-3000	MSI	1970
3000-30,000	LSI	1975
30,000-1,000000	VLSI	1980
>1,000000	ULSI	1990
>1 billion	GSI	2010

Transistor Counts



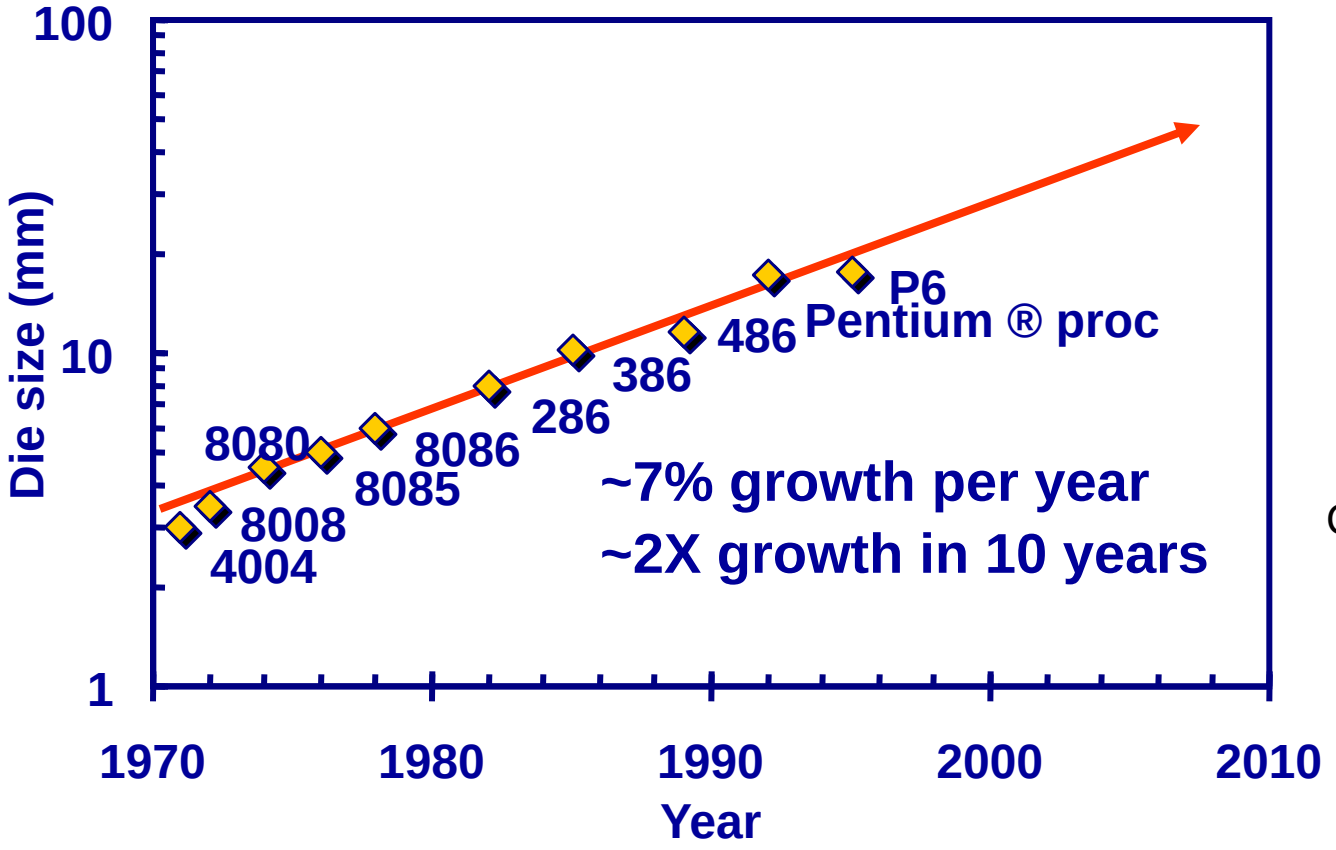
Moore's Law in Microprocessors



Courtesy, Intel

Transistors on Lead Microprocessors double every 2 years

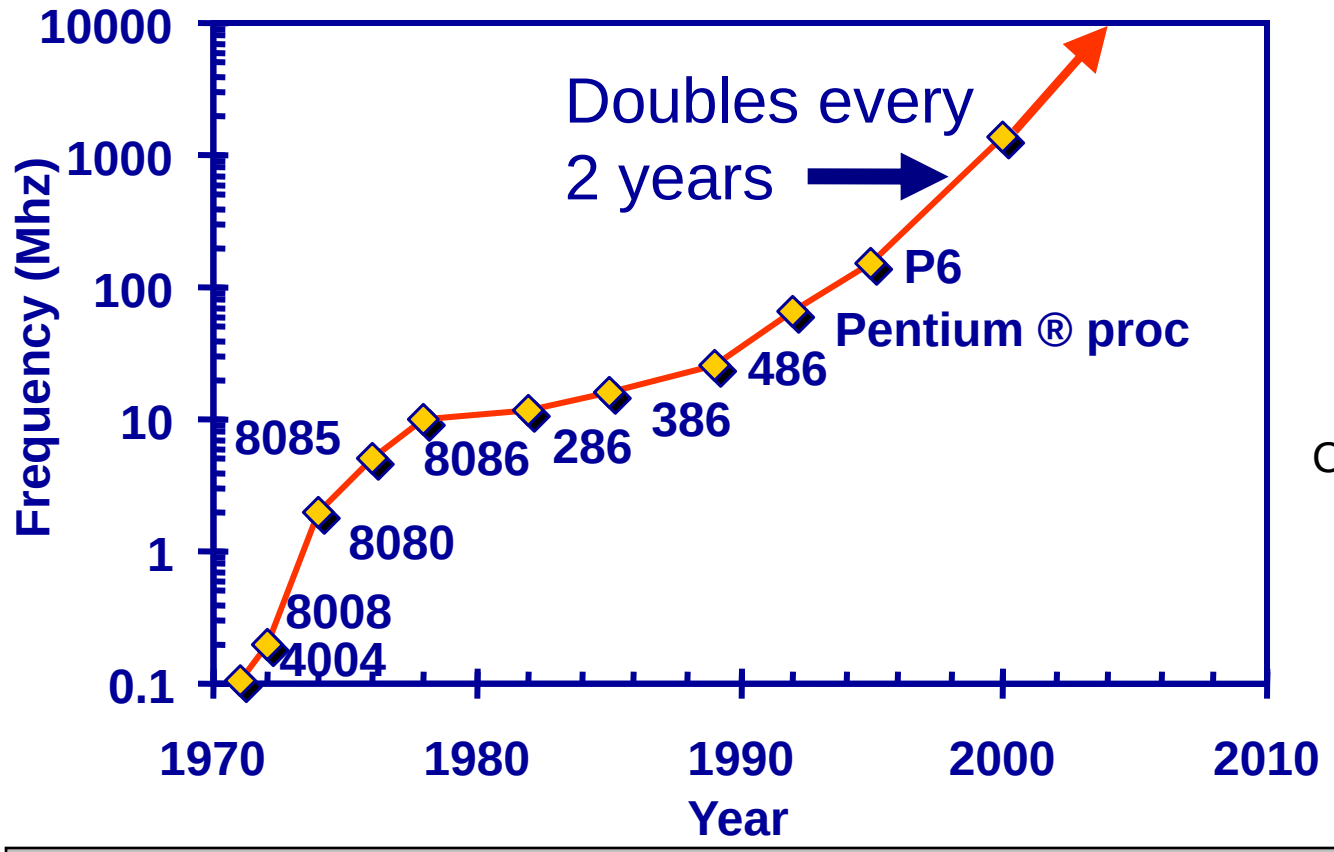
Die Size Growth



Courtesy, Intel

Die size grows by 14% to satisfy Moore's Law

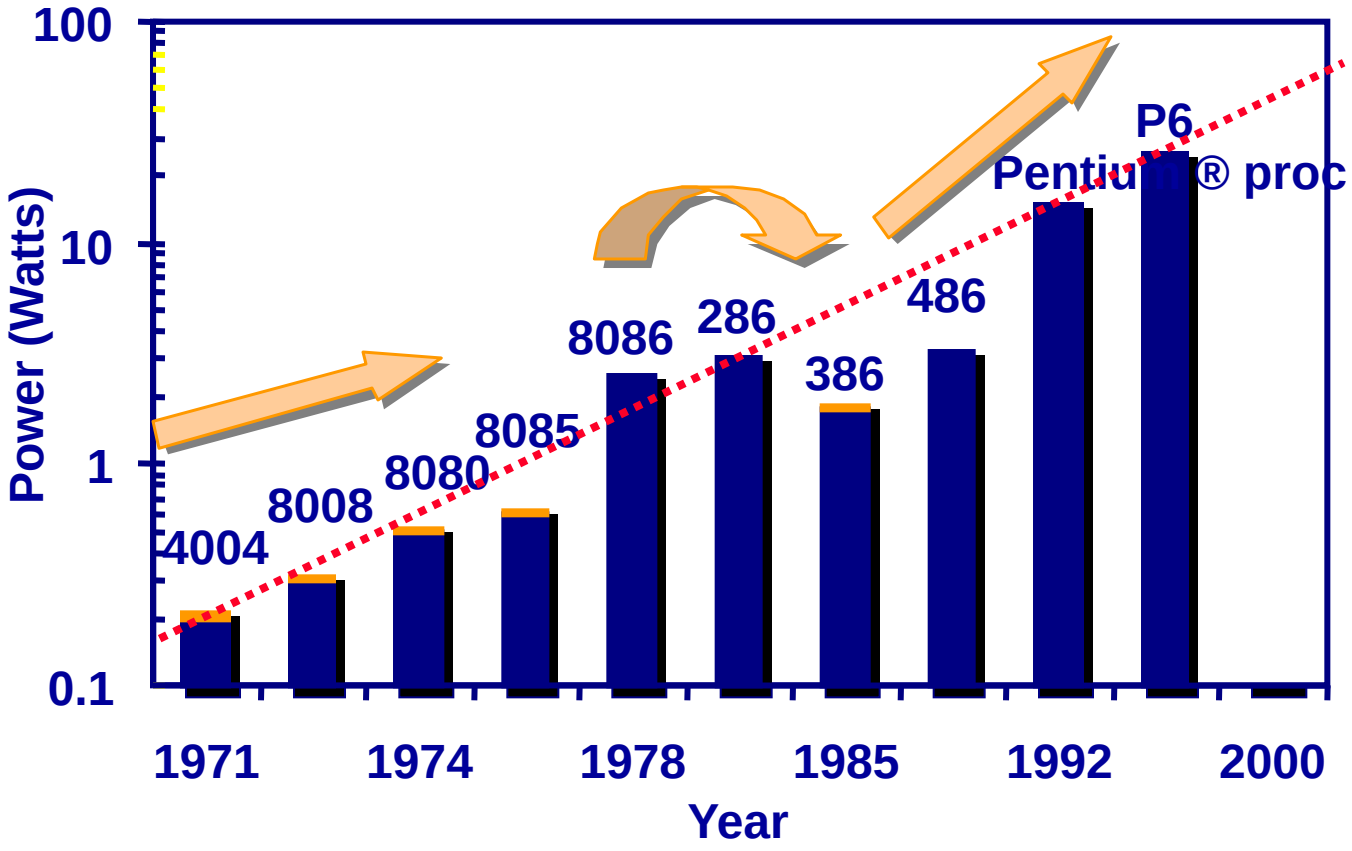
Frequency



Courtesy, Intel

Lead Microprocessors frequency doubles every 2 years

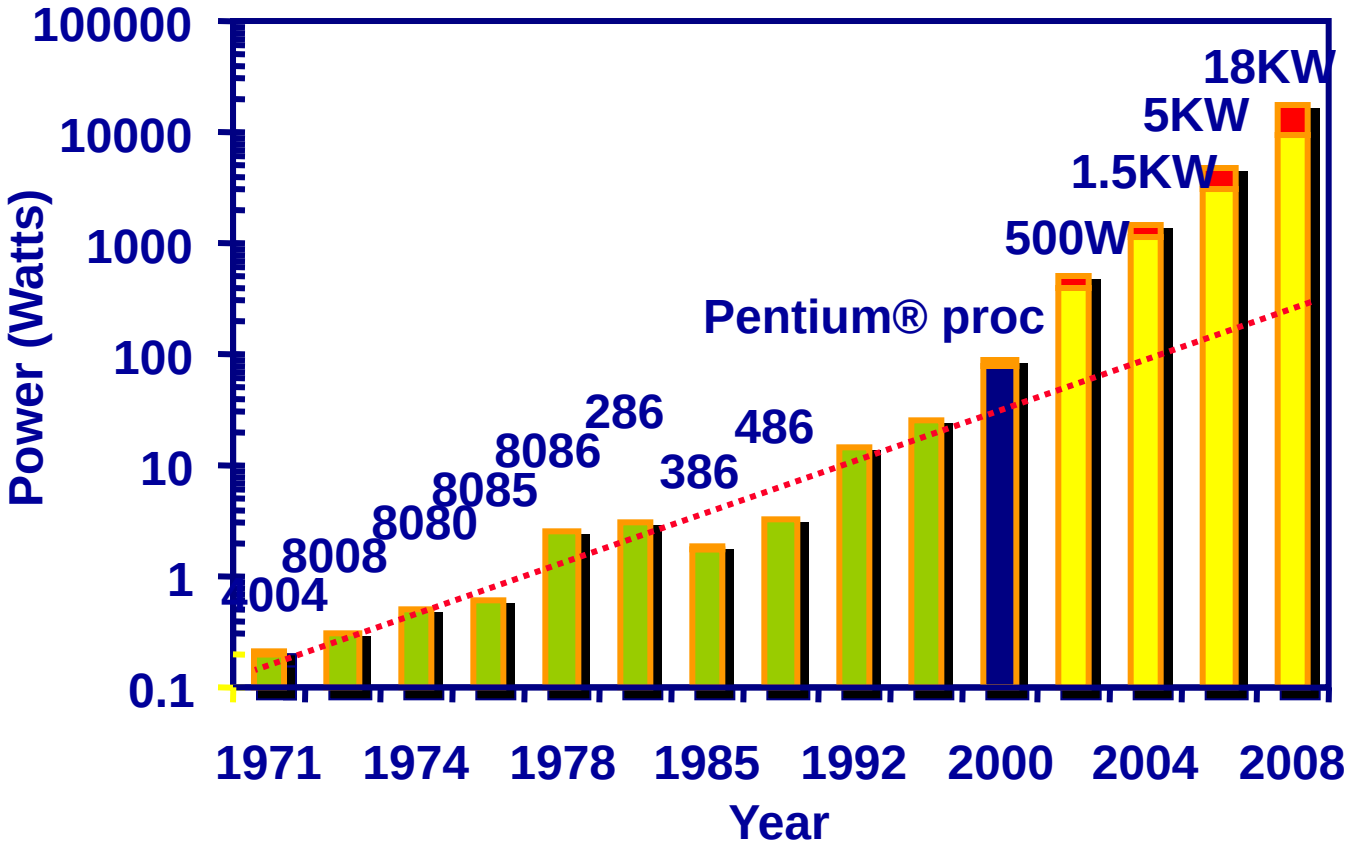
Power Dissipation



Courtesy, Intel

Lead Microprocessor power continues to increase

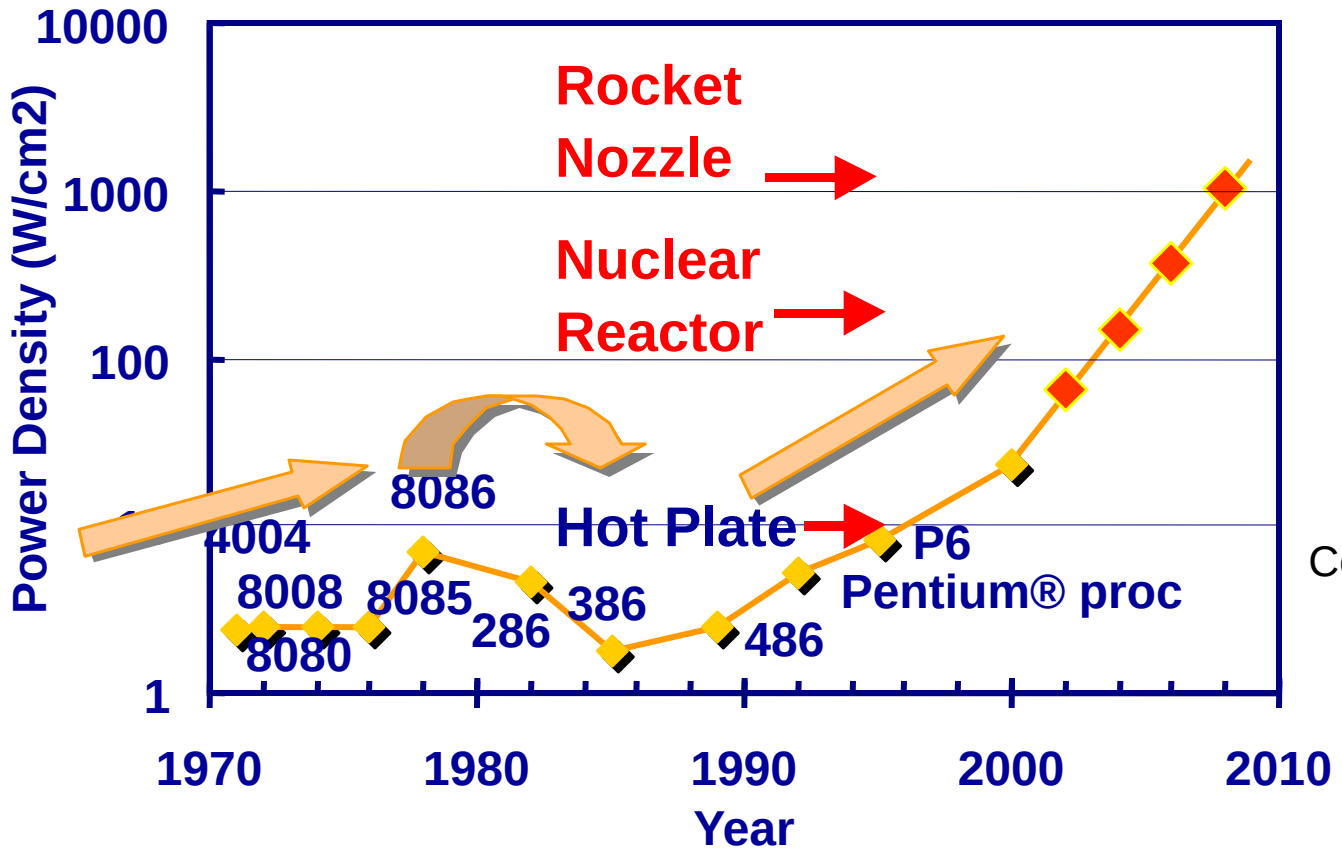
Power is a major problem….



Courtesy, Intel

Power delivery and dissipation will be prohibitive

Power density



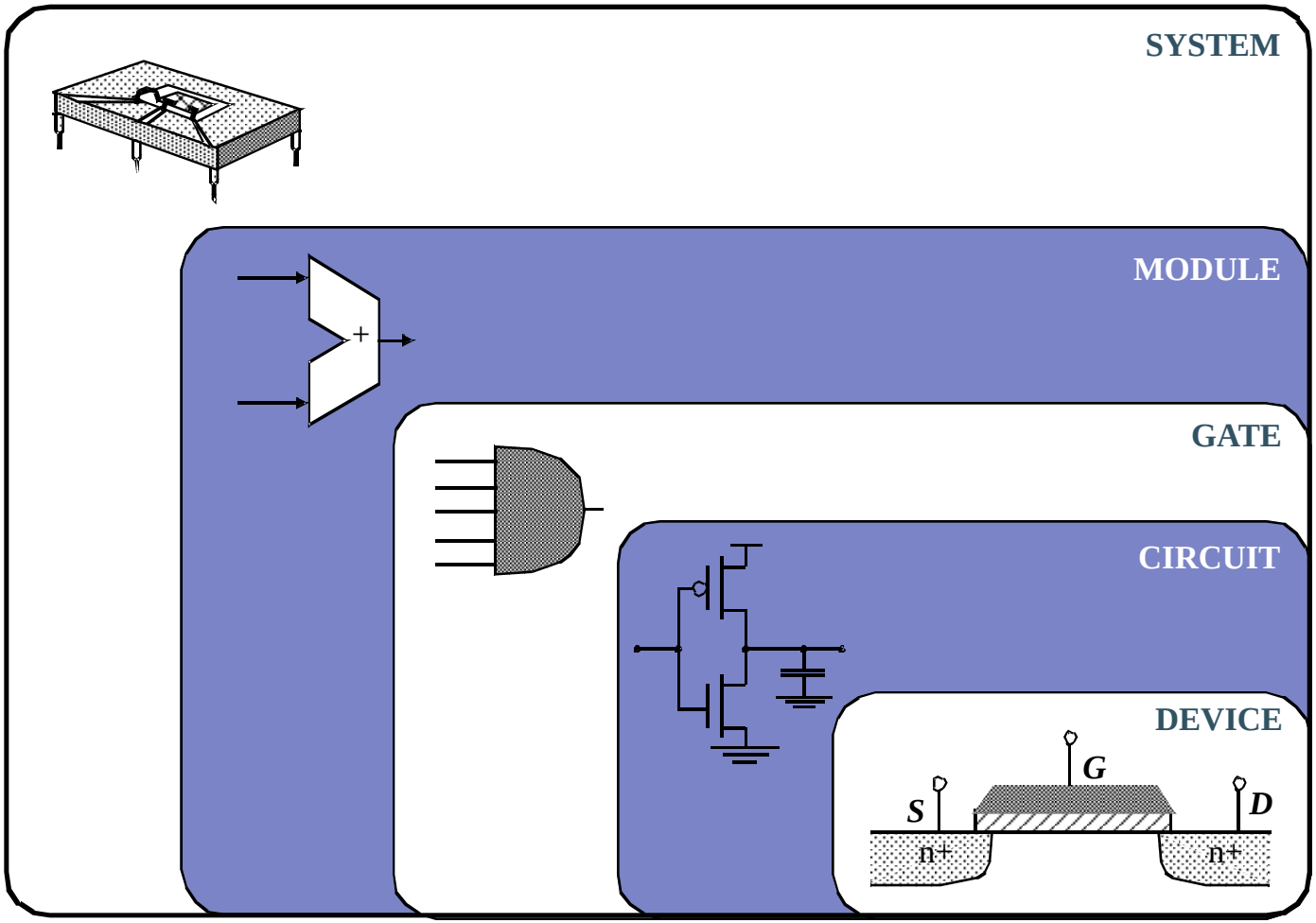
Courtesy, Intel

Power density too high to keep junctions at low temp

Why Scaling?

- Technology shrinks by 0.7/generation
- With every generation can integrate 2x more functions per chip; chip cost does not increase significantly (Dennard's (IBM) scaling law—transistors become faster, consume less power and become cheaper)
- Cost of a function decreases by 2x
- But …
  - How to design chips with more and more functions?
  - Design engineering population does not double every two years…
- Hence, a need for more efficient design methods
  - Exploit different levels of abstraction

Design Abstraction Levels



Silicon and Dopant Atoms

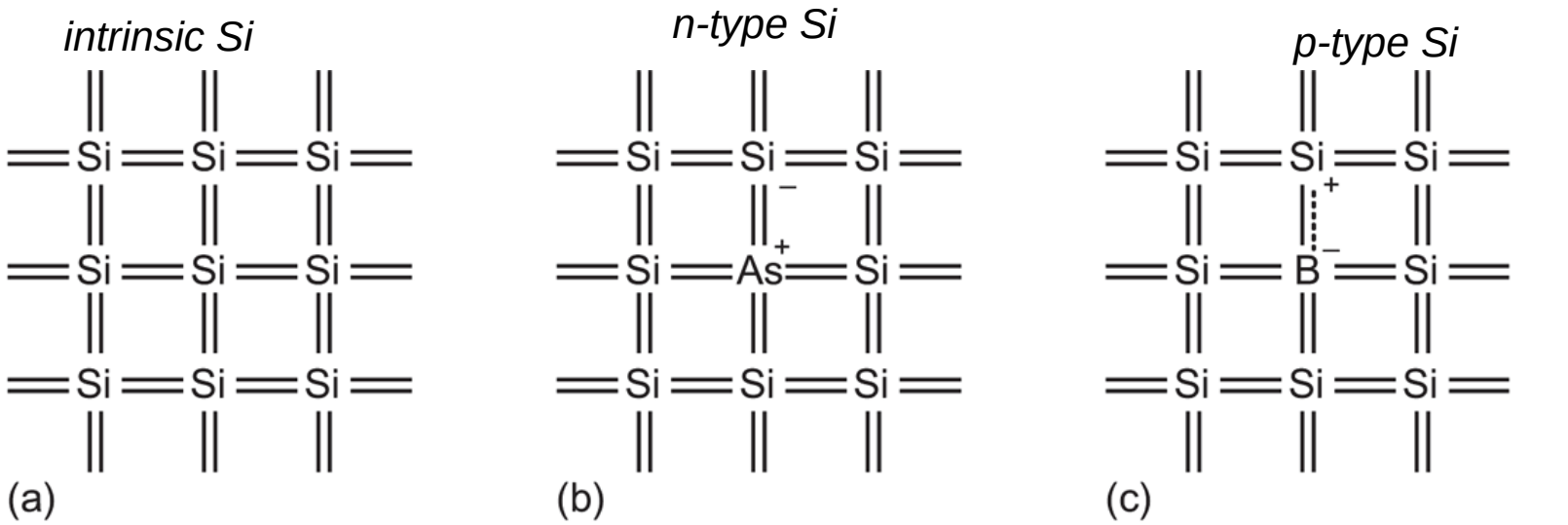
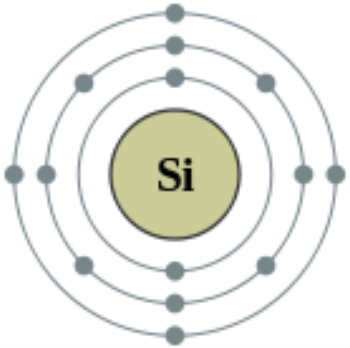


FIG 1.6 Silicon lattice and dopant atoms

Atomic number of Si: ***14***

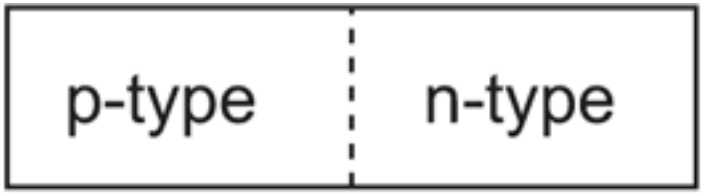
Eighth *most abundant element on earth!!*

Electronic configuration: $[1s]^2[2s]^2[2p]^6[3s]^2[3p]^2$

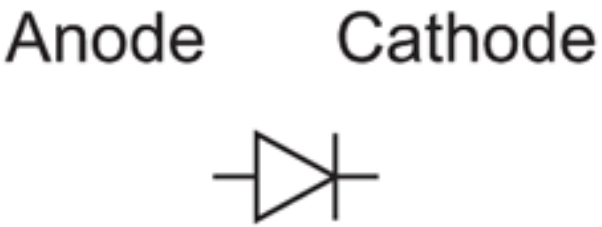


4 valence electrons

Pn-junction Diode



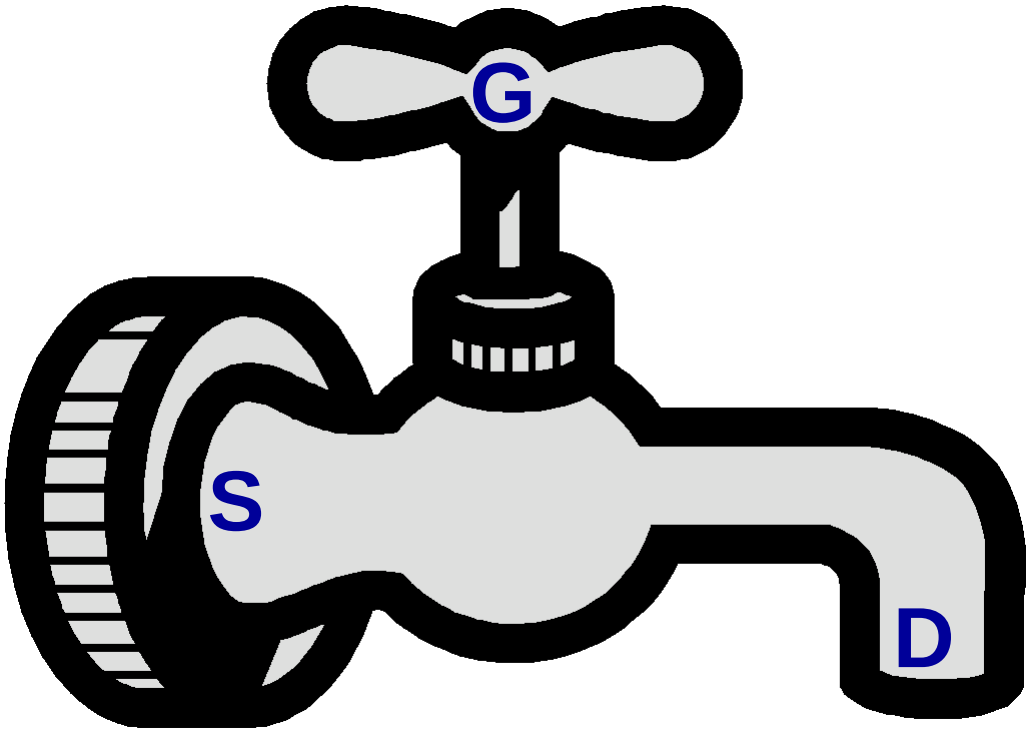
Forward Bias:
connect Vdd to p-type and GND to n-type (current flows)



Reverse Bias:
connect Vdd to n-type and GND to p-type (no current)

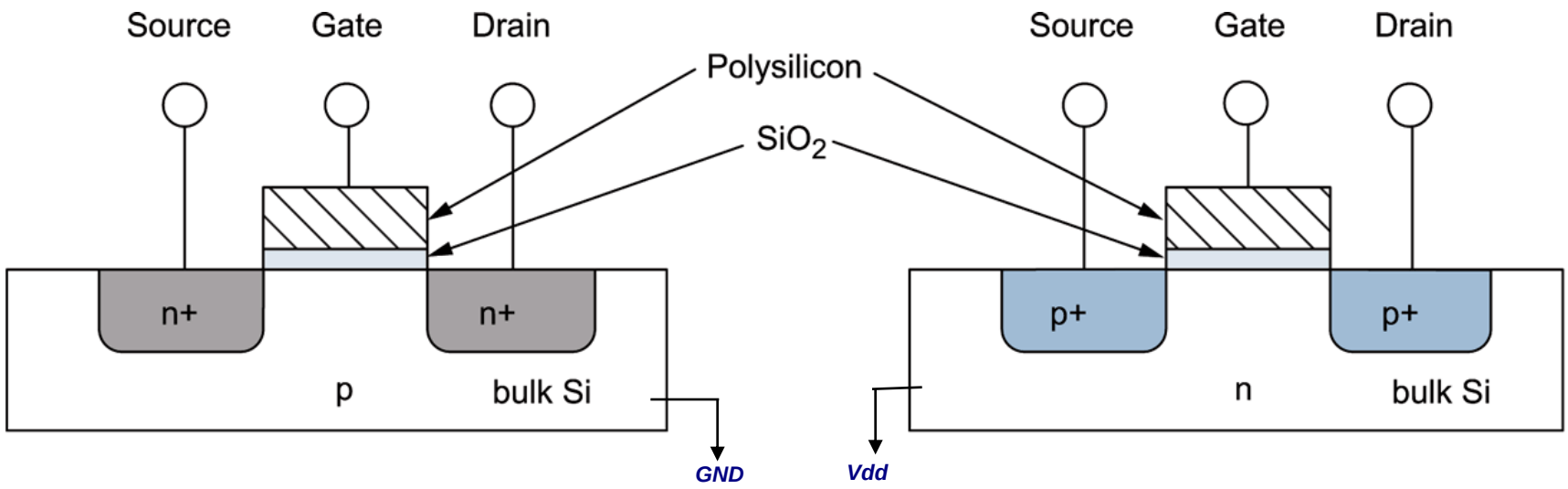
FIG 1.7 p-n junction diode structure and symbol

MOS Transistor is like a tap….



Think about how you want your tap to function….

MOS (Metal-Oxide-Semiconductor) Transistor



NMOS

(a tap that turns ON with clockwise movement of gate)

(a)

Vgs > Vt

(Vt is +ve)

PMOS

(a tap that turns ON with anti-clockwise movement of gate)

(b)

Vgs < Vt

(Vt is -ve)

FIG 1.8 nMOS transistor (a) and pMOS transistor (b)

Gate controls the flow of charge from source to drain…..

Vt is the minimum voltage (threshold voltage) required to turn ON transistors

MOS Transistor as a Switch

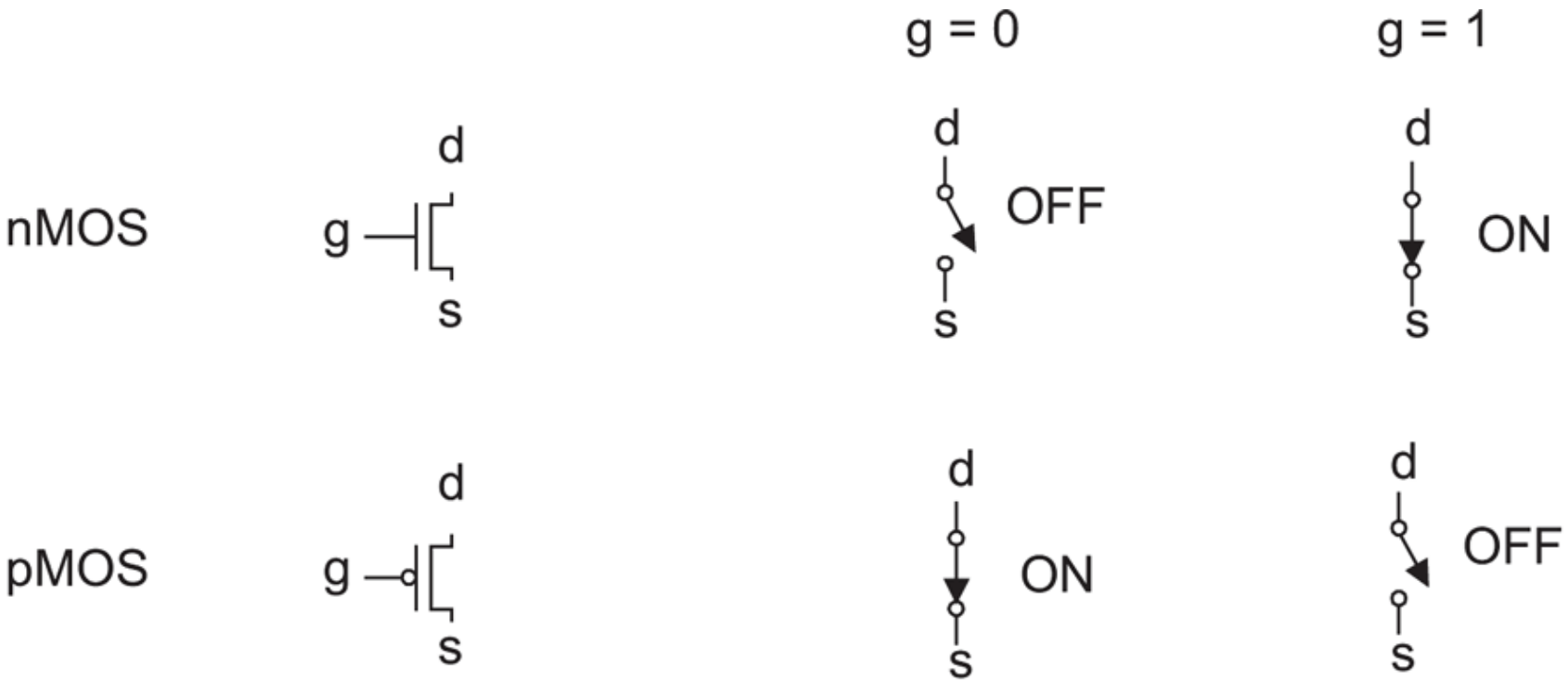


FIG 1.9 Transistor symbols and switch-level models

Inverter (NOT Gate)

Table 1.1 Inverter truth table	
A	Y
0	1
1	0

Transistor Level Implementation

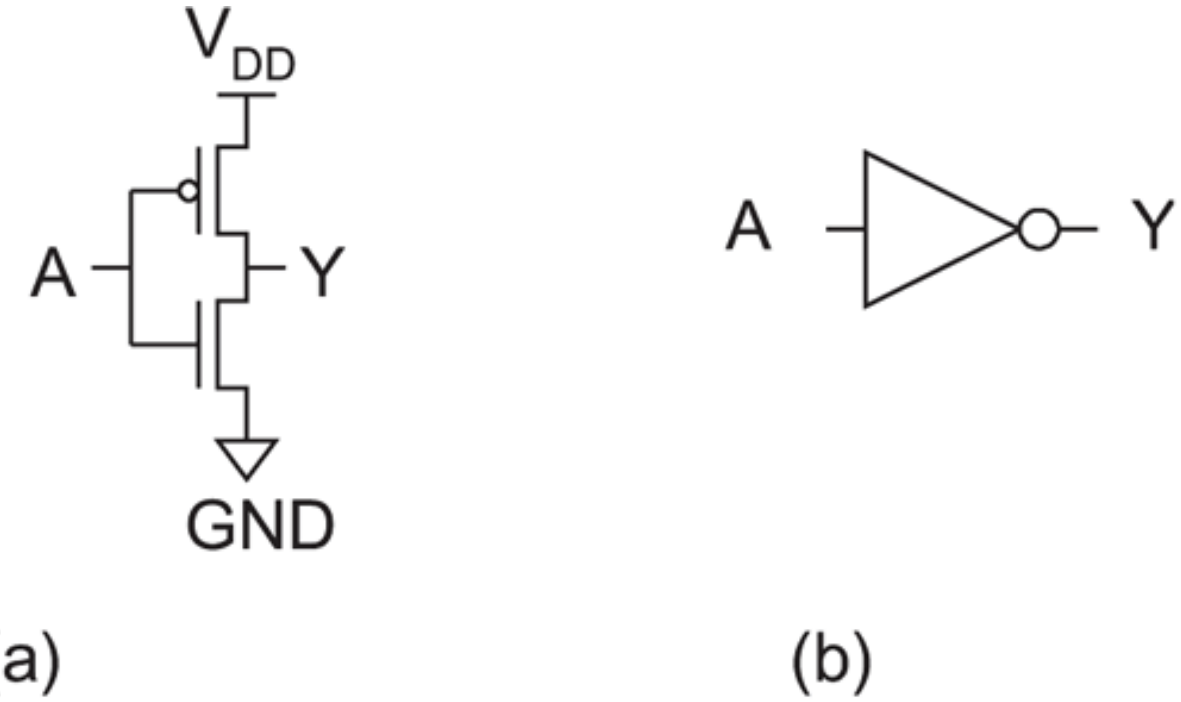


FIG 1.10 Inverter schematic (a) and symbol (b) $Y = \overline{A}$

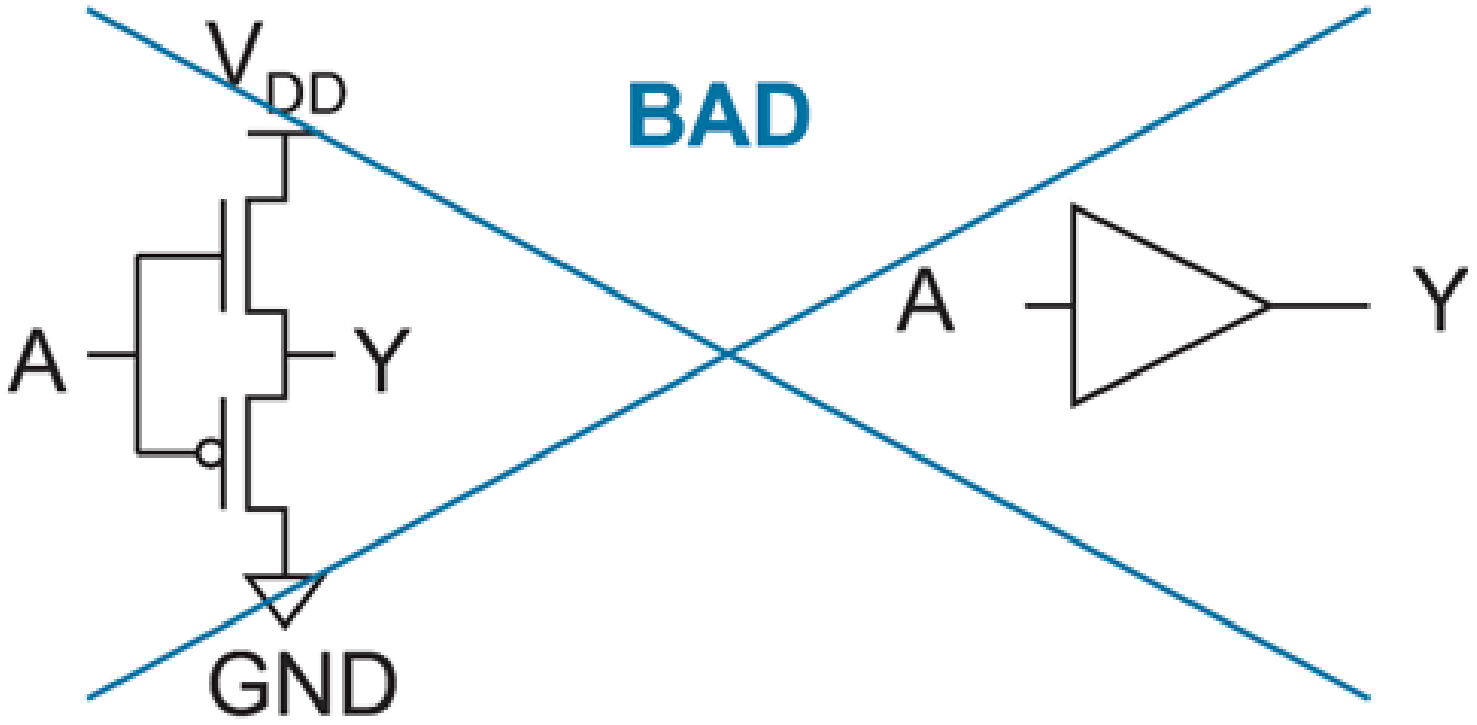
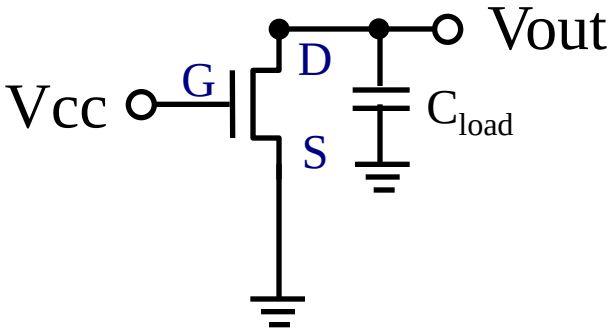


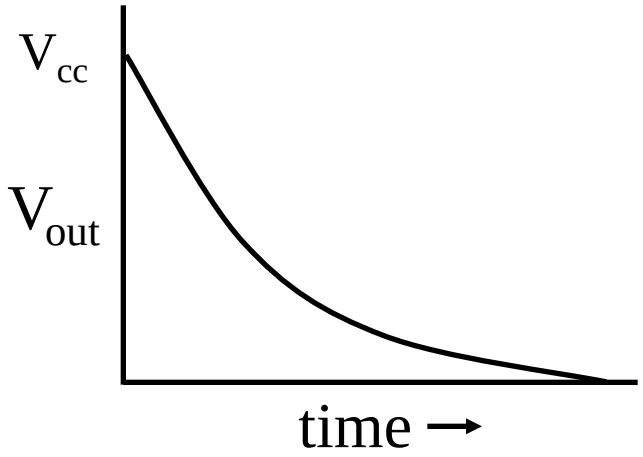
FIG 1.21 Bad noninverting buffer

MOS voltage levels

Case 1: NMOS discharges capacitor

- Initially: Vout = Vcc (capacitor fully charged)
- V_{GS} of NMOS = Vcc
- What is final Vout?



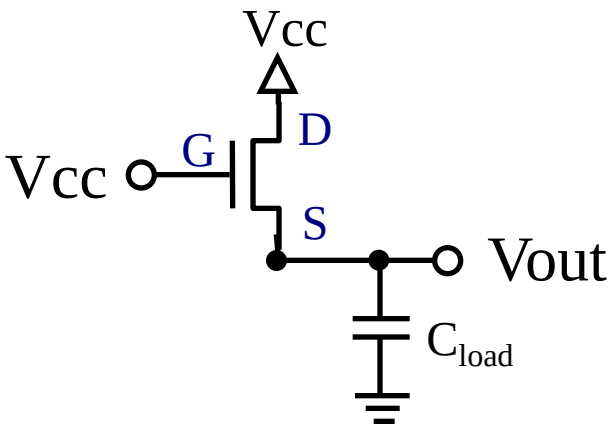


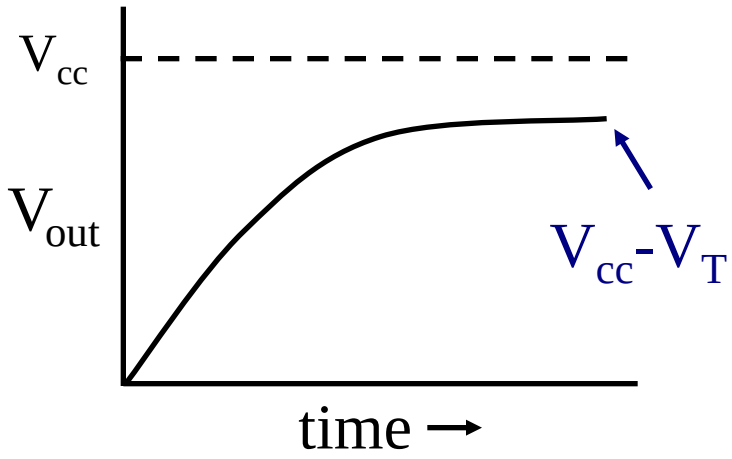
- NMOS remains on since $V_{GS} > V_T$
- Final output voltage $V_{out} = 0$ V
  - **Value at source (=0) is transferred to the drain (output)….completely**

MOS voltage levels

Case 2: NMOS charges capacitor

- Initially: Vout = 0
- Initial V_{GS} of NMOS = Vcc
- What is final Vout?



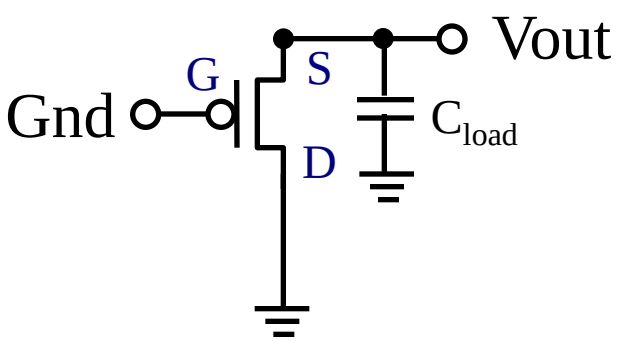


- NMOS remains on until $V_{GS} = V_T$
- Final output voltage $V_{out} = V_{cc} - V_T$
  - **Value at drain (=1) not transferred completely to the source (output)….**

MOS voltage levels

Repeat for PMOS:

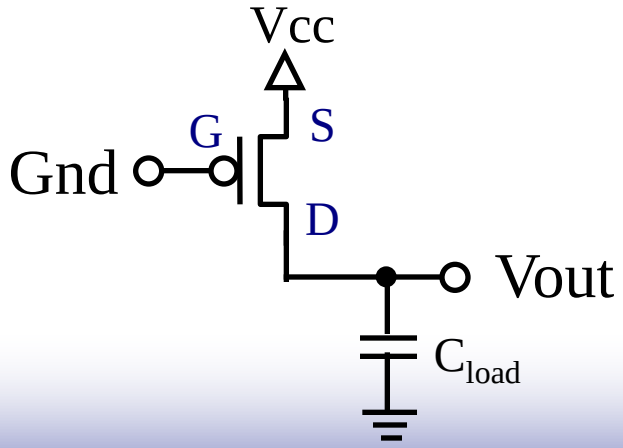
- **Case 1: PMOS discharging capacitor**



- PMOS on until $V_{GS} = -V_T$
- $V_{out} = |V_T|$

Value at drain (=0) is not transferred completely to the source (output)….

- **Case 2: PMOS charging capacitor**



- PMOS always on ($V_{GS} = -V_{CC}$)
- $V_{out} = V_{CC}$

Value at source (=1) is transferred to the drain (output)….completely

MOS voltage levels

- **NMOS summary**
  - Transfers logic ‘0’ completely (good for discharging a node)
  - Does not transfer logic ‘1’ completely (bad for charging a node)
- **PMOS summary**
  - Transfers logic ‘1’ completely
  - Does not transfer logic ‘0’ completely
- **Result:**
  - **NMOS used for pull-down, PMOS for pull-up**

Switch Behavior of NMOS and PMOS

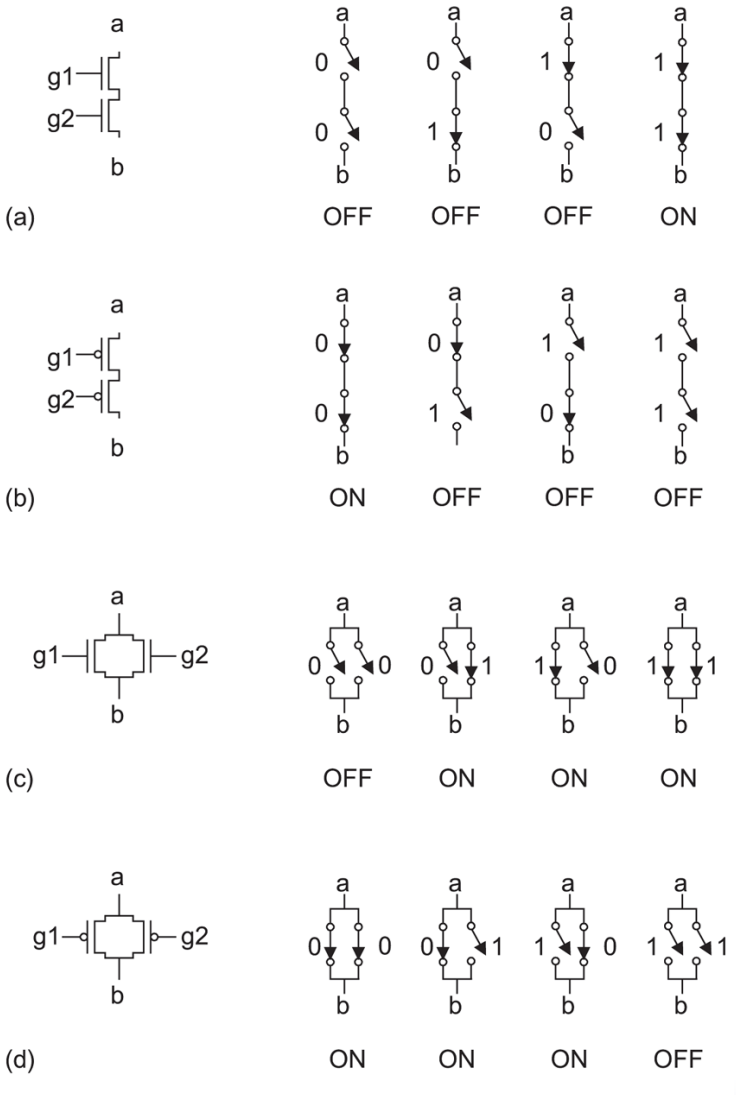


FIG 1.14 Connection and behavior of series and parallel transistors

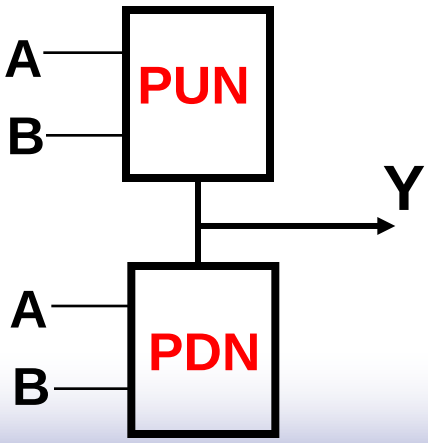
To establish a path between "a" and "b", both g1 AND g2 must be ON

To establish a path between "a" and "b", at least g1 OR g2 must be ON

2-input NAND Gate

Table 1.2 NAND gate truth table

A	B	pull-down network	pull-up network	Y
0	0	OFF	ON	1
0	1	OFF	ON	1
1	0	OFF	ON	1
1	1	ON	OFF	0

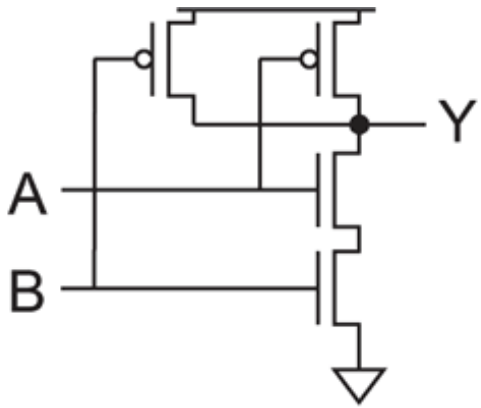


2 NMOS transistors must be in series….

2 PMOS transistors must be in parallel….

CMOS NAND Implementation

De Morgan's Law…..



(a)

(b)

$$Y = \overline{A \cdot B} = \overline{A} + \overline{B}$$

FIG 1.11 2-input NAND gate schematic (a) and symbol (b) $Y = \overline{A \bullet B}$

CMOS 3-input NAND Implementation

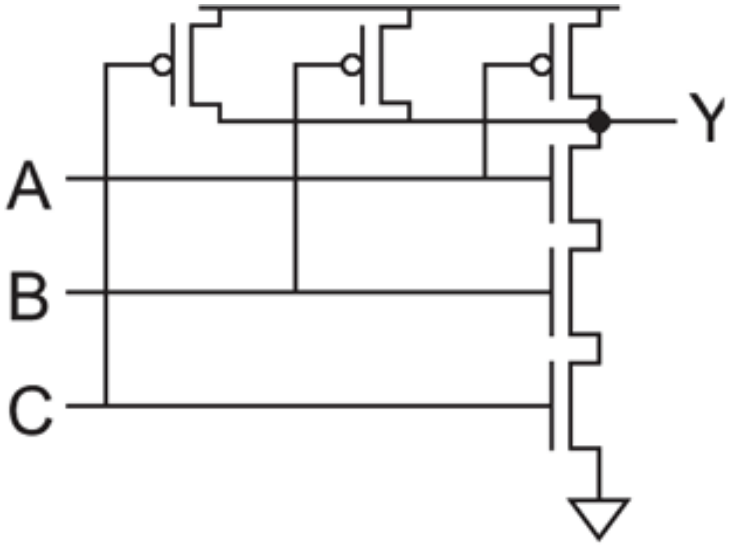


FIG 1.12 3-input NAND gate schematic

$$Y = \overline{A \bullet B \bullet C}$$

Y=0, when A=B=C=1

Hence, A, B, C are in series for the NMOS (pull-down network)

Y=1, when A or B or C=0

Hence, A, B, C are in parallel for the PMOS (pull-up network)

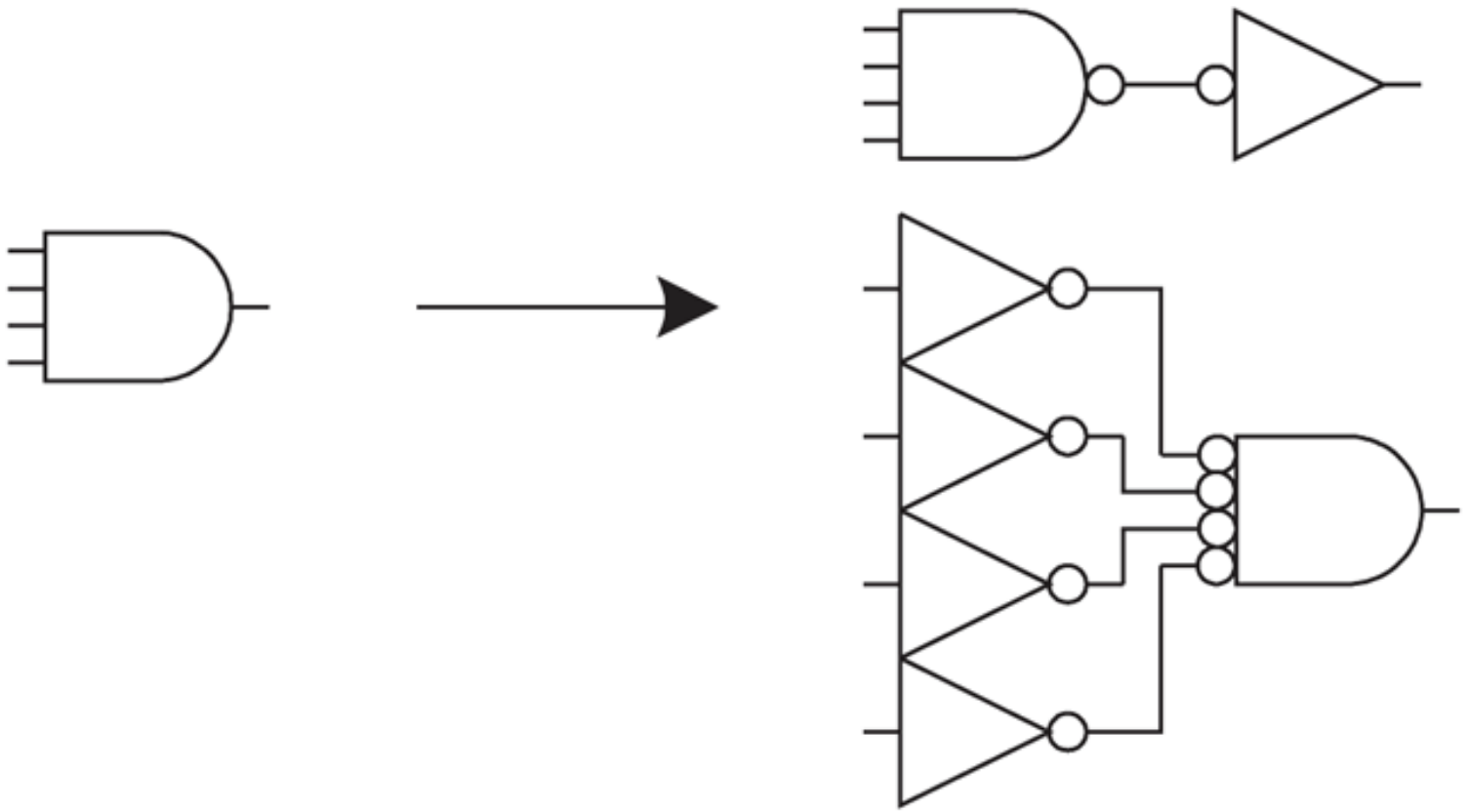


FIG 1.22 Various implementations of a CMOS 4-input AND gate