香港中文大學
The Chinese University of Hong Kong

CENG5030

Part 1-1: Introduction

Bei Yu

(Latest update: January 7, 2019)

Spring 2019

Question

Why Energy Efficient Computing?

Question

In computing, where the energy consumption comes from?

Overview

Background: Digital Logic

Power Modeling

Power Reduction: First Glance

Overview

Background: Digital Logic

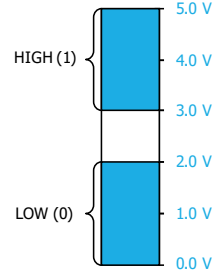
Power Modeling

Power Reduction: First Glance

Digital Logic

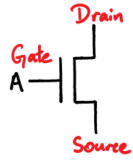
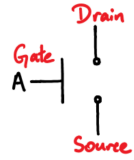
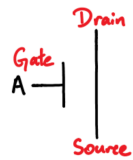
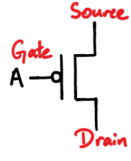
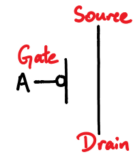
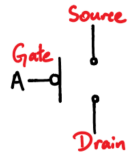
- Digital logic circuits operate on logical values, represented by voltage ranges.

Logic 0	Logic 1
False	True
Off	On
LOW	HIGH
No	Yes
Open switch	Closed switch



- Voltages between ground and a certain threshold represent the logical value 0.
- Voltages between a higher threshold and VDD represent the logical value 1.
- The threshold levels are design choices.
- If a voltage falls in the gap between the defined logical ranges, the result is undefined and there must be an error in the logic circuit that produced it.

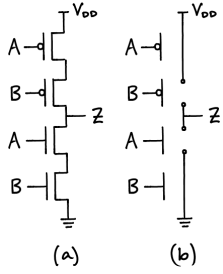
MOSFET Approximations

MOSFET Type	Logic Circuit Symbol	$A = 0$ Approximation	$A = 1$ Approximation
NMOS			
PMOS			

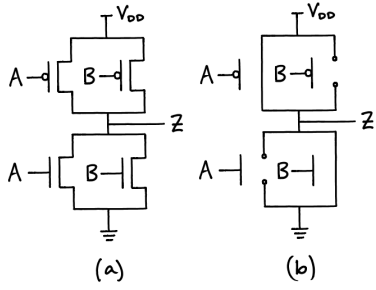
MOSFETs can be approximated as either open or short circuits between drain and source.

CMOS Logic Circuits

- CMOS logic circuits consist of complementary arrangements of NMOS and PMOS transistors.
- A CMOS circuit is reliable because its design guarantees that its output is always shorted to either ground or VDD but not both at the same time.
- As a consequence, the design also ensures that VDD is never shorted to ground through Z, which makes CMOS circuits power-efficient.

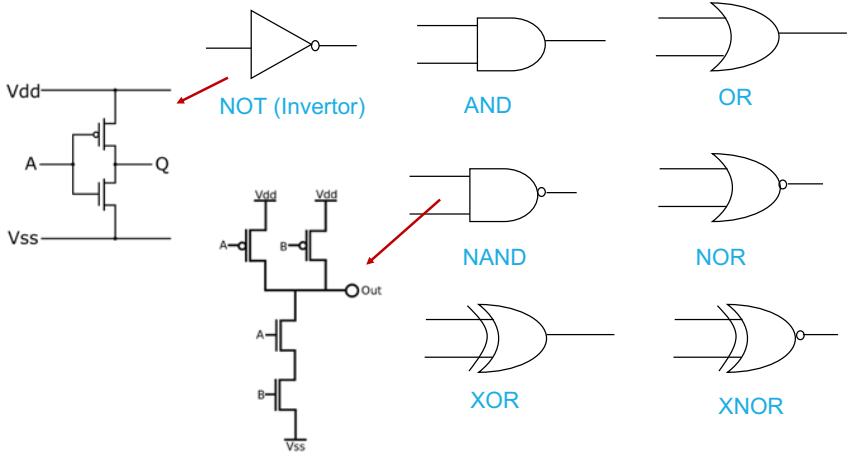


(a) Bad Design 1

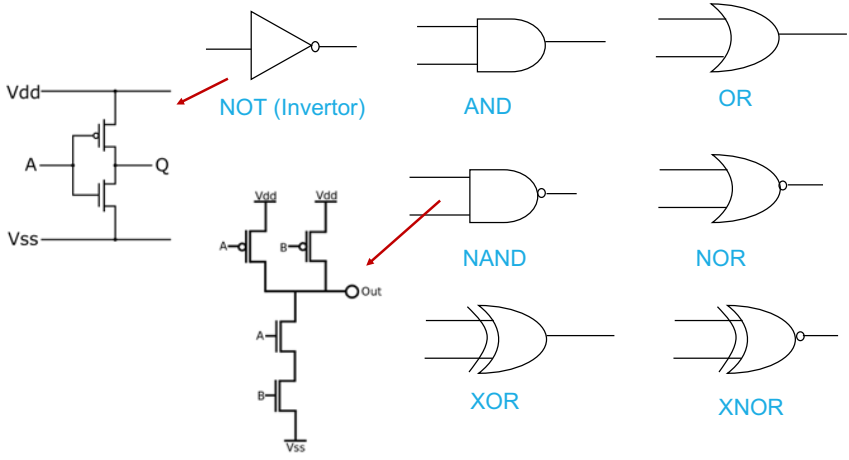


(b) Bad Design 2

Logic Gates


Vdd
A
Q
Vss
NOT (Invertor)
AND
OR
Vdd
Vdd
A
B
Out
A
B
Vss
NAND
NOR
XOR
XNOR

Logic Gates



Question:

What is the schematic view of an AND gate?

Question:

Please draw NOR gate schematic view.

Overview

Background: Digital Logic

Power Modeling

Power Reduction: First Glance

Dynamic Power

Dynamic Power Modeling

$$P \propto C \cdot V^2 \cdot A \cdot f$$

- C: total capacitance seen by the gate's outputs
- V: supply voltage
- A: activity of the gates in the system
- f: frequency of the system's operation

Dynamic Power

Dynamic Power Modeling

$$P \propto C \cdot V^2 \cdot A \cdot f$$

- C: total capacitance seen by the gate's outputs
- V: supply voltage
- A: activity of the gates in the system
- f: frequency of the system's operation

Question:

What's the most effective way to reduce dynamic power consumption?

Static (Leakage) Power

The power dissipated by a transisotr whose gate is intended to be off.

Static Power Modeling

$$P \propto V \cdot I_{leak}$$
$$I_{leak} \propto exp(-q \cdot V_{th})$$

- V_{th}: threshold voltage
- Minimum gate-to-source voltage that is needed to create a conducting path between the source and drain terminals
- Click here for an animation of threshold voltage.

Compensation of Voltage Scaling

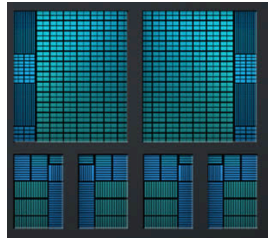
$$f_{max} \propto \frac{(V - V_{th})^2}{V}$$

- Maximum frequency is roughly linear in V
- Voltage should be larger than threshold voltage

Compensation of Voltage Scaling

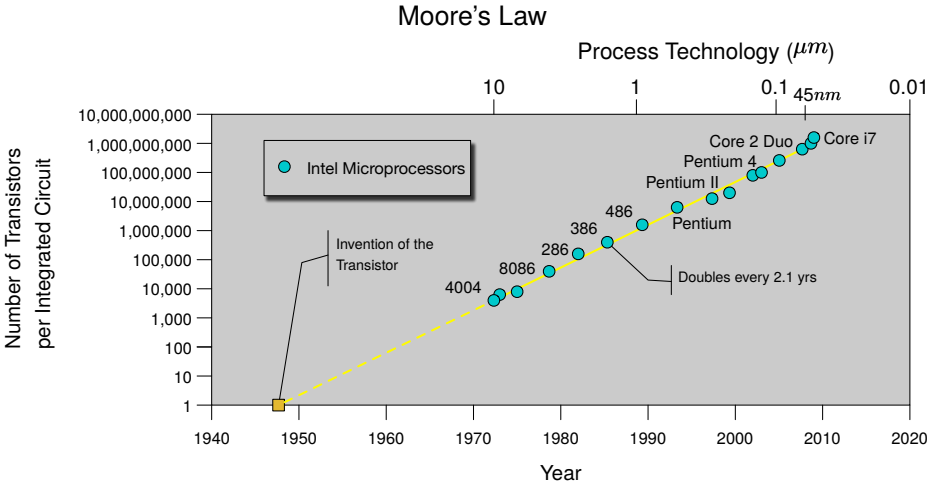
$$f_{max} \propto \frac{(V - V_{th})^2}{V}$$

- Maximum frequency is roughly linear in V
- Voltage should be larger than threshold voltage
- Motivation of parallel computing

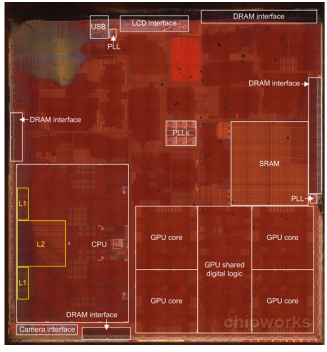


Apple A11 chip, in 2017.

Background: Moore's Law to Extreme Scaling

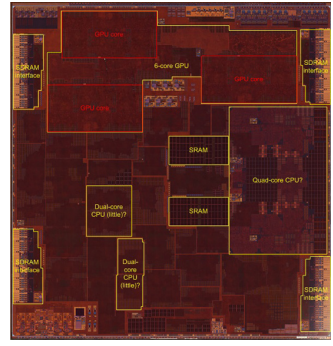


Background: Scaling of Apple SOC



Apple A7 (2013)

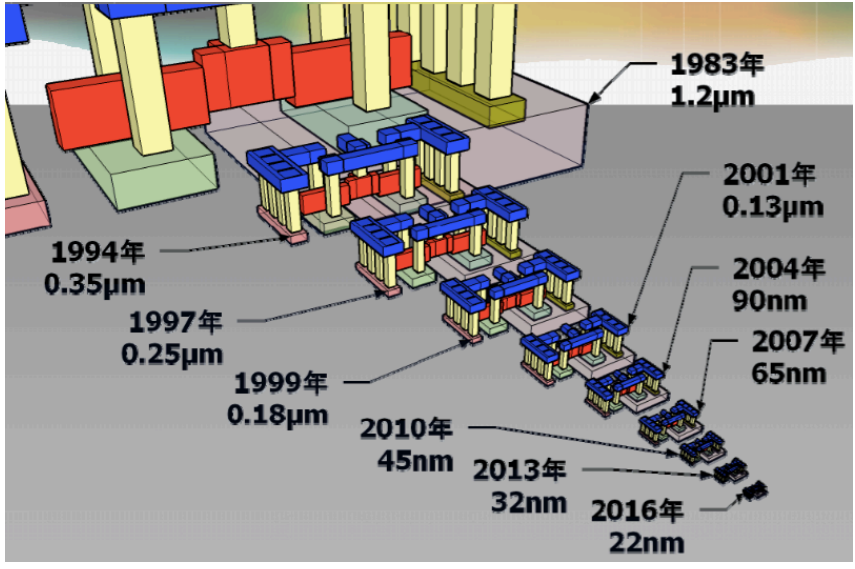
- 1,000,000,000 Transistors
- 102mm^2 die size
- 1.3GHz

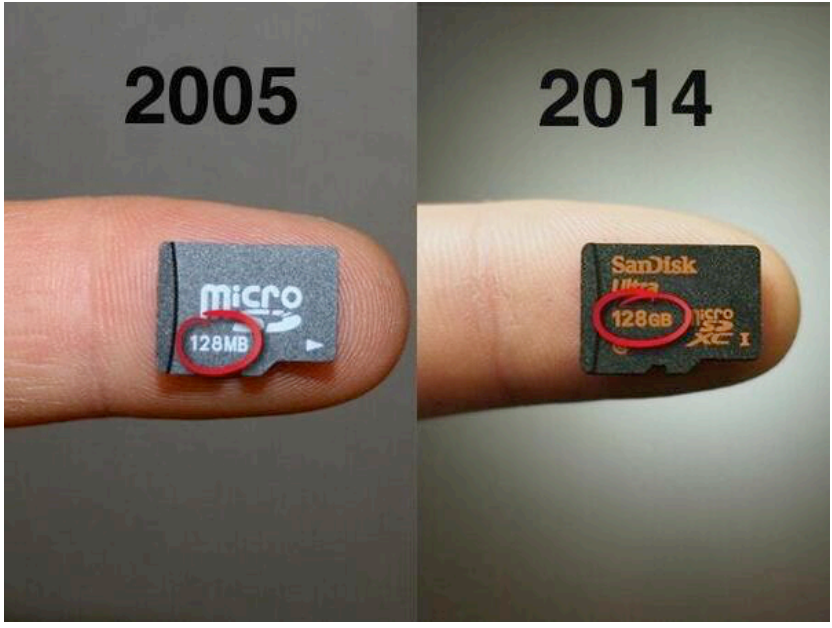


Apple A10 (2016)

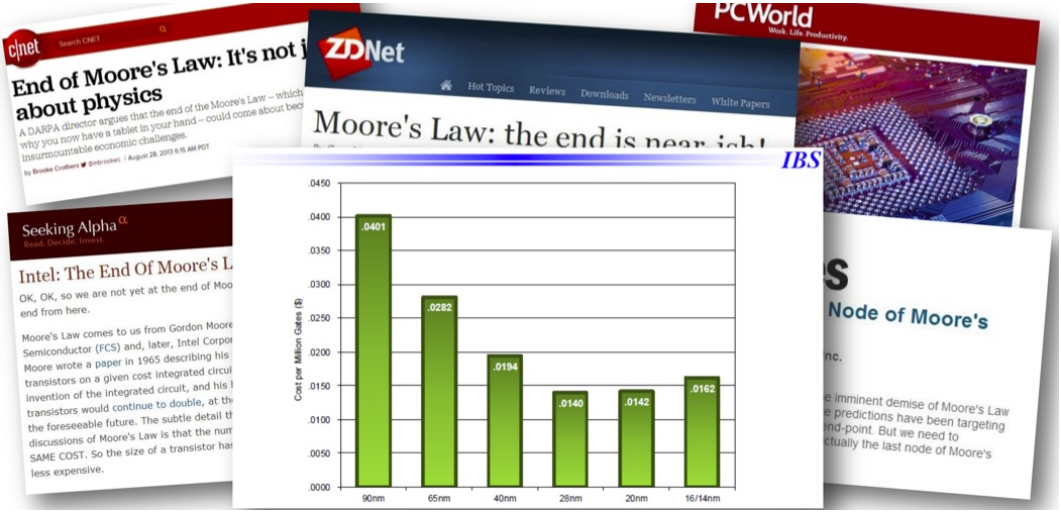
- 3,300,000,000 Transistors
- 125mm^2 die size
- 2.34GHz

Background: An Inverter Example




2005
2014
micro
128MB
SanDisk
128GB

Question if Moore's Law Can Continue



Question:

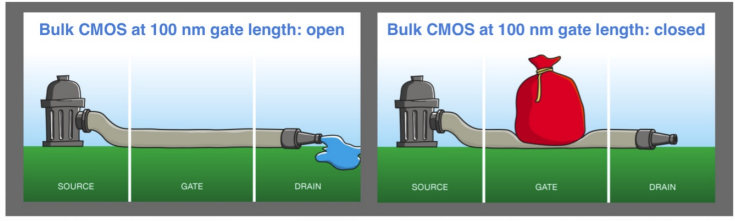
What is the bottleneck of further scaling?

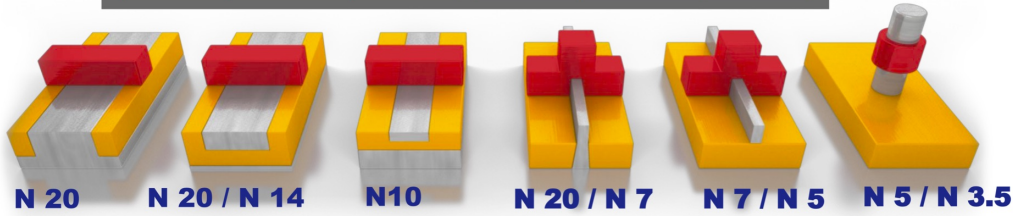
Static Power Challenges∗

Shrink scenarios for logic devices

ASML

Public
Slide 12
November 2014

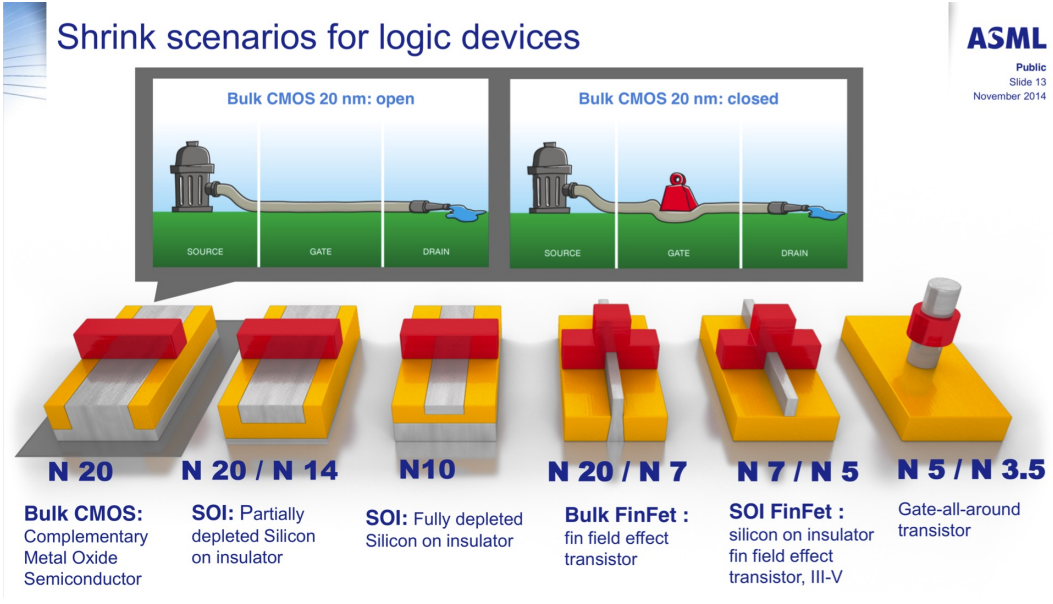




N 20	N 20 / N 14	N10	N 20 / N 7	N 7 / N 5	N 5 / N 3.5
Bulk CMOS: Complementary Metal Oxide Semiconductor	**SOI:** Partially depleted Silicon on insulator	**SOI:** Fully depleted Silicon on insulator	**Bulk FinFet :** fin field effect transistor	**SOI FinFet :** silicon on insulator fin field effect transistor, III-V	Gate-all-around transistor

∗M. Brink, “Many ways to shrink: The right moves to 10 nanometer and beyond”, 2014.

Static Power Challenges∗



∗M. Brink, “Many ways to shrink: The right moves to 10 nanometer and beyond”, 2014.

Overview

Background: Digital Logic

Power Modeling

Power Reduction: First Glance

Logic Level Techniques

Clock Gating

Turn off clock tree branches to latches or flip-flops whenever they are not used.

Half-frequency Clocks

Use both edges of the clock to synchronize events

Asynchronous Logic

Without global clock signals, the system can save considerable power. (Drawback?)

Others: gate sizing; wire-sizing; scaling voltages...

Architecture Level Techniques

Memory Systems

- Small cache in front of L1 cache – reduce total activities
- Memory banking: split memory into banks and only one bank
- Shut down part of memory – reduce static power

Buses

- Gray code – switches the least number of signals in neighbor data
- Transmitting the delta
- Data compression

Parallel and Pipeline