Power (2)

Hanwool Jeong hwjeong@kw.ac.kr

Contents

- Introduction
- Dynamic Power
- Static Power

Dynamic Power

 \checkmark Power required for switching

Switching Energy in CMOS Inverter

- Half of the energy (=half of power) is dissipated in the pMOS transistor and the other half is delivered to the capacitor.
- Then the rest of half is dissipated in the nMOS



$$E_{\rm switching} = C_{\rm out} V_{\rm DD}^2$$

Operational Waveform



Short Circuit Current Energy in CMOS Inverter

- There is a time duration that nMOS and pMOS simultaneously. This component is not included in CV_{DD}² term.
- Provided that the slope of IN is not much small, the short circuit current portion is negligible.
- → It is important to note that signal slope matters!



 $E_{\text{shortcircuit}} = I_{\text{shortciurcuit}} \times t_{\text{on}}$

Energy to Power (1); **Considering Clock Period**

- We should first see "how often" digital circuit switches
- In general, digital circuit operates synchronous to "clock."
- That is, can we say

$$P_{\text{switching}} = E_{\text{switching}} / T_{\text{clk}} = C_{\text{out}} V_{\text{DD}}^2 f_{\text{clk}} ???$$

$$Digital Circuit$$

$$\int \\ \int \\ \int \\ CLK \\ T_{\text{clk}} = 1/f_{\text{clk}}$$

$$vitching = E_{switching} / T_{clk} = C_{out} V_{DD}^2 f_{clk} ????$$

Energy to Power (2); Activity Factor & Switching Period vs. Clock Period

- Note that a signal does not switch every clock.
- Define that a switching period $T_{sw} = 1/f_{sw}$ that says how often OUT experiences $0 \rightarrow 1$.
- f_{sw} is smaller than f_{clk} , so we can relate them utilizing $\alpha < 1$,



Switching Power

• In conclusion, we can say

$$P_{\rm switching} = \alpha C V_{\rm DD}^2 f_{\rm clk}$$

- C is effective capacitance of all nodes
- α is the probability that the circuit node transitions from 0 to 1.
- A clock has an activity factor of $\alpha = 1$ because it rises and falls every cycle.
- Most data has a maximum activity factor of 0.5 because it transitions only once each cycle.
- Static CMOS logic has been empirically determined to have activity factors closer to 0.1

Total Dynamic Power

• Dynamic energy:

$$E_{dynamic} = E_{switching} + E_{shortcircuit}$$

• Dynamic power:

$$P_{dynamic} = P_{switching} + P_{shortcircuit}$$
$$= \alpha C V_{DD}^{2} f_{clk} + P_{shortcircuit} \approx \alpha C V_{DD}^{2} f_{clk}$$

• We can reduce V_{DD} or C to reduce power consumption. What is the counter effect of reducing V_{DD} ?

Example 5.1

A digital system-on-chip in a 1 V 65 nm process (with 50 nm drawn channel lengths and $\lambda = 25$ nm) has 1 billion transistors, of which 50 million are in logic gates and the remainder in memory arrays. The average logic transistor width is 12 λ and the average memory transistor width is 4 λ . The memory arrays are divided into banks and only the necessary bank is activated so the memory activity factor is 0.02. The static CMOS logic gates have an average activity factor of 0.1. Assume each transistor contributes 1 fF/ μ m of gate capacitance and 0.8 fF/ μ m of diffusion capacitance. Neglect wire capacitance for now (though it could account for a large fraction of total power). Estimate the switching power when operating at 1 GHz.

How Can We Reduce Dynamic Power?

Considering

$$P_{dynamic} = \alpha C V_{DD}^2 f_{clk} + P_{shortcircuit}$$

- 1) Reducing α
- 2) Reducing C
- 3) Reducing V_{DD}
- 4) Reducing P_{shortcircuit}

1) Reducing α; Power Saving by Clock Gating

• Clock gating ANDs a clock signal with an enable to turn off the clock to idle blocks.



Activity Factor

- Define P_i to be the probability that node i is 1. $\overline{P_i} = 1 P_i$ is the probability that node i is 0.
- α_i , the activity factor of node i, is the probability that the node is 0 on one cycle and 1 on the next.

$$\alpha_i = \overline{P}_i P_i$$

- Completely random data has P = 0.5 and thus α = 0.25.
- Structured data may have different probabilities.
 - For example, the upper bits of a 64-bit unsigned integer representing a physical quantity such as the intensity of a sound or the amount of money in your bank account are 0 most of the time

Switching Probabilities

• Can you derive the activity factor of the output?

Gate	Ργ
AND2	$P_{\mathcal{A}}P_B$
AND3	$P_A P_B P_C$
OR2	$1 - \overline{P}_{\mathcal{A}}\overline{P}_{B}$
NAND2	$1 - P_A P_B$
NOR2	$\overline{P}_{\!\mathcal{A}}\overline{P}_B$
XOR2	$P_{\mathcal{A}}\overline{P}_{B} + \overline{P}_{\mathcal{A}}P_{B}$

Example 5.2

Figure 5.8 shows a 4-input AND gate built using a tree (a) and a chain (b) of gates. Determine the activity factors at each node in the circuit assuming the input probabilities $P_A = P_B = P_C = P_D = 0.5$.



Glitches Related to α

• It is only the case when we assume zero propagation delay



 However, in reality, gates sometimes make spurious transitions called glitches when inputs do not arrive simultaneously.



Activity factor can be above 1, increasing the power consumption

2) Reducing C; Gate Sizing

- With logical effort, we can optimize the delay of digital circuit.
- Then, how about the power? How does the gate sizing affect on the P_{dynamic}? You can consider the C effect in P_{dynamic}.
- With the assumption that unit inverter has gate cap 3C, we can represent the gate cap of a general logic gate with its g, p,and x.
- Then we can represent its dynamic power analytically to examine the effect of gate sizing on the power
- Then we can see how we can adjust sizing for reducing energy.

Revisit Drive

Generally, drive x is defined as (when C_{in} = 1 for the unit inverter)

$$x = C_{in}/g$$



Dynamic Energy vs. Gate Sizing

- With the assumption that unit inverter has gate cap 3C, then a gate with logical effort g, parasitic delay p, and drive x has
 - gx times as much gate capacitance
 - px times as much diffusion capacitance.
- Then can you derive the dynamic energy for circuit below?



Energy_i =
$$CV_{DD}^2$$

= $V_{DD}^2 (p_i x_i 3C + \sum_j g_j x_j 3C + C_{wire})$
= $3CV_{DD}^2 x_i (p_i + g_i h_i + c_i) = 3CV_{DD}^2 x_i d_i$

Dynamic Energy vs. Delay

We can sum up the energy of all nodes i,
 → Must consider the activity difference among nodes.

$$E = \sum_{i} \alpha_{i} 3CV_{DD}^{2}x_{i}d_{i}$$

Then, we can define the normalized dynamic energy for C and V for given process, as follows by diving 3CV_{DD}²
 ➔ To only see the effect of sizing

$$E = \sum_{i} \alpha_{i} x_{i} d_{i}$$

- You can derive delay vs. Energy curve using above equation.
- Or, you can seek to minimize E such that the worst-case arrival time is less than some delay D

Example 5.3

Generate an energy-delay trade-off curve for the circuit from Figure 4.37 as delay varies from the minimum possible ($D_{\min} = 23.44 \tau$) to 50 τ . Assume that the input probabilities are 0.5.



Dynamic Energy vs. Delay Curve Example



3) Reducing V_{DD}; Clustered Voltage Scaling

- Can we reduce V_{DD} ?
- Reducing V_{DD} reduces on current
 Delay is increased.
- But you remember critical path? How about lowering V_{DD} selectively to non-critical path?



Issue at V_{DDL} to V_{DDH} Interface

 Even the output of the first stage inverter is high, the pMOS in the second stage inverter can be turned on if V_{DDH} – V_{DDL} > V_{th} and burn current.



➔ To handle this problem, level shifter or level converter is required.

Level Shifter (Level Converter)

• Below shown is the level converter. How can it shift V_{DDL} signal into V_{DDH} signal?



Clustered Voltage Scaling



Summary on Dynamic Power

- $P_{dynamic} = \alpha CV_{DD}^2 f_{clk} + P_{shortcircuit} \approx \alpha CV_{DD}^2 f_{clk}$
- Determination of activity factor and clock gating
- Energy-delay trade-off in gate sizing
- V_{DD} can be lowered by clustering V_{DD} domain but LC is required