Basics of Energy & Power Dissipation
Outline

- Basic Concepts
- Dynamic power
- Static power
- Time, Energy, Power Tradeoffs
- Activity model for power estimation
  - Combinational and sequential logic
Reading

- http://www.xbitlabs.com/articles/cpu/display/core-i5-2500t-2390t-i3-2100t-pentium-g620t.html

- Goal: Understand
  - The sources of power dissipation in combinational and sequential circuits
  - Power vs. energy
  - Options for controlling power/energy dissipation
Where Does the Power Go in CMOS?

- **Dynamic** Power Consumption
  - Caused by switching transitions → cost of switching state

- **Static** Power Consumption
  - Caused by leakage currents in the absence of any switching activity

- Power consumption per transistor changes with each technology generation
  - No longer reducing at the same rate
  - What happens to power density?
n-channel MOSFET

- $V_{gs} < V_t$ transistor off - $V_t$ is the threshold voltage
- $V_{gs} > V_t$ transistor on

- Impact of threshold voltage
  - Higher $V_t$, slower switching speed, lower leakage
  - Lower $V_t$, faster switching speed, higher leakage

- Actual physics is more complex but this will do for now!
Abstracting Energy Behavior

- How can we abstract energy consumption for a digital device?

- Consider the energy cost of charge transfer
Switch from one state to another

To perform computation, we need to switch from one state to another

Logic 1: Cap is charged
Logic 0: Cap is discharged

Connect the cap to GND thorough an ON NMOS
Connect the cap to VCC thorough an ON PMOS

The logic dictates whether a node capacitor will be charged or discharged.
Dynamic Power

- Dynamic power is used in charging and discharging the capacitances in the CMOS circuit.

\[
C_L = \text{load capacitance}
\]
Switching Delay

\[ V_s = V_{dd}(1 - e^{-\frac{t}{RC}}) \]

\[ \tau = RC \]

Charging to logic 1
Switching Delay

\[ V_c = V e^{-\frac{t}{RC}} \]

Discharging to logic 0
Switching Energy

Energy drawn from supply:

\[ E_{vDD} = \int_{0}^{\infty} i_{vDD}(t)V_{DD} \, dt = V_{DD} \int_{0}^{\infty} \left( \frac{d(C_Lv_{out})}{dt} \right) \, dt = C_LV_{DD} \int_{0}^{\infty} dv_{out} = C_LV_{DD}^2 \]

Energy stored in capacitor:

\[ E_C = \int_{0}^{\infty} i_c(t)v_{out} \, dt = \int_{0}^{\infty} \frac{d(C_Lv_{out})}{dt}v_{out} \, dt = C_L \int_{0}^{\infty} v_{out} \, dv_{out} = \frac{1}{2} C_LV_{DD}^2 \]

Energy dissipated in resistor:

\[ E_R = E_{vDD} - E_C = \frac{1}{2} C_LV_{DD}^2 \]

Also,

\[ E_R = \int_{0}^{\infty} i_r^2(t)R \, dt = \int_{0}^{\infty} i_r(t) \frac{V_{DD} - v_{out}}{R} \, dt = \int_{0}^{\infty} \frac{d(C_Lv_{out})}{dt} \left( V_{DD} - v_{out} \right) \, dt = \frac{1}{2} C_LV_{DD}^2 \]

\[ => \text{Independent of } R \]

\[ \text{Courtesy: Prof. A. Roychowdhury} \]
Power Vs. Energy

- Energy is a rate of expenditure of energy
  - One joule/sec = one watt
- Both profiles use the same amount of energy at different rates or power

Same Energy = area under the curve
Dynamic Power vs. Dynamic Energy

- Dynamic power: consider the rate at which switching (energy dissipation) takes place

\[ P_{\text{dynamic}} = \alpha \left( \frac{C_L}{2} \right) \cdot V_{dd} \cdot V_{dd} \cdot F \]

Delay = \( k \cdot C \cdot \frac{V_{dd}}{(V_{dd} - V_t)^2} \)

activity factor = fraction of total capacitance that switches each cycle
Charge as a State Variable

For computation we should be able to identify if each of the variable (a, b, c, x, y) is in a ‘1’ or a ‘0’ state.

We could have used any physical quantity to do that
- Voltage
- Current
- Electron spin
- Orientation of magnetic field
- ........

All nodes have some capacitance associated with them

We choose voltage distinguish between a ‘0’ and a ‘1’.

Logic 1: Cap is charged
Logic 0: Cap is discharged
Higher Level Blocks

![Diagram of higher level blocks]
Gate Power Dissipation

- Switching activity depends on the input pattern and combinational logic
- Consider a $0 \rightarrow 1$ transition on the output of a gate

\[ p_0 \times p_1 \]

- Probability gate output was 0
- Probability gate output is 1

\[ p_0 = \frac{N_0}{2^n} \quad p_1 = \frac{N_1}{2^n} \]

\[ N_0 = \text{number of 0's in the truth table} \]

Example:

\[(16)\]
ALU Energy Consumption

- Can we count the number of transitions in each 1-bit ALU for an operation?
- Can we estimate static power?
- Computing per operation energy

(17)
Closer Look: A 4-bit Ripple Adder

- Critical Path = $D_{XOR} + 4(D_{AND} + D_{OR})$ for 4-bit ripple adder (9 gate levels)
- For an $N$-bit ripple adder
  - Critical Path Delay $\sim 2(N-1)+3 = (2N+1)$ Gate delays
- Activity (and therefore power) is a function of the input data values!
Implications

• What if I halved the frequency?
• What if I lower the voltage?
• How can I reduce the capacitance?

\[ P_{\text{dynamic}} = \alpha \left( \frac{C_L}{2} \right) \cdot V_{dd} \cdot V_{dd} \cdot F \]

We will see later that these two are interrelated!
Technology scaling has caused transistors to become smaller and smaller. As a result, static power has become a substantial portion of the total power.

\[ P_{static} = V_{dd} \cdot I_{static} \]
Energy-Delay Interaction

- Delay decreases with supply voltage but energy/power increases

\[ P_{\text{dynamic}} = \alpha \left( \frac{C_L}{2} \right) \cdot V_{dd} \cdot V_{dd} \cdot F \]

\[ \text{Delay} = k \cdot C \frac{V_{dd}}{(V_{dd} - V_t)^2} \]

(21)
Static Energy-Delay Interaction

- Static energy increases exponentially with decrease in threshold voltage
- Delay increases with threshold voltage

\[ \text{Delay} = k \cdot C \frac{V_{dd}}{(V_{dd} - V_t)^2} \]
Temperature Dependence

- As temperature increases static power increases\(^1\)

\[
P_{\text{static}} = V_{dd} \cdot N \cdot K_{\text{design}} \cdot I_{\text{leakage}}
\]

\[
I_{\text{leakage}} = F(\text{Temp})
\]

\(^1\)J. Butts and G. Sohi, "A Static Power Model for Architects, MICRO 2000
The World Today

- Yesterday → scaling to minimize time (max $F$)

\[ P_{\text{dynamic}} = \alpha \left( \frac{C_L}{2} \right) \cdot V_{dd} \cdot V_{dd} \cdot F \quad \text{Delay} = k \cdot C \frac{V_{dd}}{(V_{dd} - V_t)^2} \]

- Maximum performance (minimum time) is too expensive in terms of power
  - Imaging scaling voltage by 0.7 and frequency by 1.5
  - → how does dynamic power scale?

- Today: trade/balance performance for power efficiency
Factors Affecting Power

- Transistor size
  - Affects capacitance ($C_L$)

- Rise times and fall times (delay)
  - Affects short circuit power (not in this course)

- Threshold voltage
  - Affects leakage power

- Temperature
  - Affects leakage power

- Switching activity
  - Frequency ($F$) and number of switching transistors ($\alpha$)

\[
P_{\text{dynamic}} = \alpha \left( \frac{C_L}{2} \right) \cdot V_{dd} \cdot V_{dd} \cdot F
\]
\[
\text{Delay} = k \cdot C \frac{V_{dd}}{(V_{dd} - V_t)^2}
\]
Low Power Design: Options?

\[ P_{\text{dynamic}} = \alpha \left( \frac{C_L}{2} \right) \cdot V_{dd} \cdot V_{dd} \cdot F \]
\[ \text{Delay} = k \cdot C \frac{V_{dd}}{(V_{dd} - V_t)^2} \]

- **Reduce \( V_{dd} \)**
  - Increases gate delay
  - Note that this means it reduces the frequency of operation of the processor!

- **Compensate by reducing threshold voltage?**
  - Increase in leakage power

- **Reduce frequency**
  - Computation takes longer to complete
  - Consumes more energy (but less power) if voltage is not scaled
Example

AMD Trinity A10-5800 APU: 100W TDP

<table>
<thead>
<tr>
<th>HW Only (Boost)</th>
<th>CPU P-state</th>
<th>Voltage (V)</th>
<th>Freq (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pb0</td>
<td>1</td>
<td>2400</td>
<td></td>
</tr>
<tr>
<td>Pb1</td>
<td>0.875</td>
<td>1800</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SW-Visible</th>
<th>CPU P-state</th>
<th>Voltage (V)</th>
<th>Freq (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P0</td>
<td>0.825</td>
<td>1600</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>0.812</td>
<td>1400</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>0.787</td>
<td>1300</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>0.762</td>
<td>1100</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>0.75</td>
<td>900</td>
<td></td>
</tr>
</tbody>
</table>
Optimizing Power vs. Energy

Thermal envelopes → minimize peak power

Maximize battery life → minimize energy

Example:
Modeling Component Energy

- **Per-use** energies can be estimated from:
  - Gate level designs and analyses
  - Circuit-level designs and analyses
  - Implementation and measurement
- There are various open-source tools for analysis:
  - Mentor, Cadence, Synopsys, etc.

![Diagram showing flow from Hardware Design to Technology Parameters through Circuit-level Estimation Tool, leading to Estimation Results: Area, Energy, Timing, etc.](image)
Datapath Elements

- Can we measure (offline) the average energy consumed by each component?
- Can we measure (offline) the average energy consumed by each component?
A Simple Power Model for Processors

- Per instruction energy measurements
  - Permits a software model of energy consumption of a program
  - Execution time use to assess power requirements

- A first order model of energy consumption for software
  - A table of energy consumption per instruction
  - More on this later!
What About Wires?

• We will not directly address delay or energy expended in the interconnect in this class
  ❖ Simple architecture model: lump the energy/power with the source component
Summary

- Two major classes of energy/power dissipation – static and dynamic
- Managing energy is different from managing power → leads to different solutions
- Technology plays a major role in determining relative costs
- Energy of components are often estimated using approximate models of switching activity
• Explain the difference between energy dissipation and power dissipation

• Distinguish between static power dissipation and dynamic power dissipation

• What is the impact of threshold voltage on the delay and energy dissipation?

• As you increase the supply voltage what is the behavior of the delay of logic elements? Why?

• As you increase the supply voltage what is the behavior of static and dynamic energy and static and dynamic power of logic elements?
• Do you expect the 0-1 and 1-0 transitions at the output of a gate to dissipate the same amount of energy?

• For a mobile device, would you optimize power or energy? Why? What are the consequences of trying to optimize one or the other?

• Why does the energy dissipation of a 32-bit integer adder depend on the input values?

• If I double the processor clock frequency and run the same program will it take less or more energy?
• When the chip gets hotter, does it dissipate more or less energy? Why?

• How can you reduce dynamic energy of a combinational logic circuit?

• How can you reduce static energy of a combinational logic circuit?
Glossary

- Dynamic Energy
- Dynamic Power
- Load capacitance
- Static Energy
- Static Power
- Time constant
- Threshold voltage
- Switching delay
- Switching energy