Announcements

- Project check #5 out
  - Due end of next week
  - Final report due end of RRR week

- Note on project grading
  - Individual check-offs treated like “homework”, graded on a curve
  - Most of credit will be assigned to final report
  - [I.e., plenty of opportunity to do well even if you are currently behind]
  - Have simplified the rest of the project to give everyone an excellent chance of completing
Announcements

• Today is Elad’s last “official” lecture
  – Out of town next week; possible project review lecture during RRR

• Next Tues. lecture will be given by John Lazzaro
  – Will discuss many practical examples of course material
  – E.g., iPad and MacBook Air - i.e., this will be a fun lecture!
  – Physical attendance required – HKN forms at end of lecture

• Dan’s office hours on Wed. cancelled

• Wed. lab will only be held from 11am – 2pm
  – Enjoy the holiday!

Where Does Power Go In CMOS?

• Two main components [focus for today]:
  – **Dynamic [switching] power**:
    • Charging/discharging capacitors
  – **Leakage power**:
    • Transistors aren’t perfect switches

• Additional components:
  – Crowbar [short-circuit]:
    • Associated with transitions
  – Static
    • Biasing currents – should be zero in static CMOS logic
Side Note: Power vs. Energy

- Power can (in principle) be made arbitrarily low...
  - How?

- Energy per operation is often a more useful metric
  - $E_{\text{op}} = P_{\text{average}} \times T_{\text{average}}$
  - No "cheating"
Dynamic Power/Energy

- Power = Energy/transition \cdot (Transition rate/2)
  = Energy/transition \cdot (Rising transition rate)

  = C_{\text{load}} V_{dd}^2 \cdot f_{0 \rightarrow 1}
  = C_{\text{load}} V_{dd}^2 \cdot f_{\text{clk}} \cdot \alpha_{0 \rightarrow 1}

- \alpha_{0 \rightarrow 1} \text{ is called "activity factor" } \rightarrow \text{ average probability of 0 to 1 transition}
  - Will sometimes see P = (C_{\text{load}} V_{dd}^2 \cdot f \cdot \alpha), where \alpha = 2\alpha_{0 \rightarrow 1}

- Power dissipation is data and logic dependent

Example Activity Factor Calculations (1)

- Assume all primary inputs are equally likely to be 0 or 1
  - Not necessarily true in practice, but easy to handle a different assumption

- Activity factor of an inverter:
Example Activity Factor Calculations (2)

- Activity factor of a NAND gate:

Notes on Activity Factor

- Watch out for correlations between signals in complex circuits
  - Need to look at activity factor for whole circuit together, not just one gate at a time
  - (Verilog simulation is good for this)

- Watch out for glitches $\Rightarrow$ extra power...

- What is the highest activity factor signal you are likely to have?
Leakage Power

- Transistors that are supposed to be off actually leak:

\[ I_{\text{leak}} \]

Sub-Threshold Conduction

- Drain current doesn’t immediately drop to 0 at \( V_T \):

\[ I_D (V) \]

\[ V_{GS} (V) \]

- Leakage drops **exponentially** as increase \( V_T \)
**Leakage and Total Power**

- Assume that know average $I_{\text{leak}}$
  - (Note that $I_{\text{leak}}$ is actually also state dependent – NMOS vs. PMOS leakage not identical)

- $P_{\text{leak}} = I_{\text{leak}} \cdot V_{dd}$

- $P_{\text{total}} = P_{\text{dyn}} + P_{\text{leak}}$

  $$P_{\text{total}} = \alpha \cdot C_{\text{load}} \cdot V_{dd}^2 \cdot f_{\text{clk}} + I_{\text{leak}} \cdot V_{dd}$$

---

**Key Implications**

- Once you’ve done a “good” job with the design
  - i.e., gotten rid of all of the capacitance that doesn’t need to switch, used only the hardware you really needed, etc.

- Biggest knob for reducing power is reducing $V_{dd}$
  - $E/\text{op}$ proportional to $V_{dd}^2$

- But, reducing $V_{dd}$ reduces performance unless reduce $V_t$
  - And reducing $V_t$ increases $I_{\text{leak}}$ exponentially
  - How to find the best balance?
Energy-Performance Space

- Plot all designs on a 2-D plane
  - No matter what you do, can’t get below/to the right of solid line

- This line is called the “Pareto Optimal Curve”
  - Usually (always) follows law of diminishing returns

Optimization Perspective

- Instead of looking at an “arbitrary” metric like “Energy*Delay”, this curve can answer questions like:
  - What is minimum energy for X performance?
  - Over what range of performance is a particular technique (dotted line) actually better?
Key Observation

• For optimal designs, “sensitivity” to all free parameters should be equal

\[ S_{V_{dd}} = \frac{\partial \text{Energy} / \partial V_{dd}}{\partial \text{Perf} / \partial V_{dd}} \quad S_{V_{T}} = \frac{\partial \text{Energy} / \partial V_{T}}{\partial \text{Perf} / \partial V_{T}} \]

- Must equal slope of the Pareto curve – otherwise could trade one parameter for the other

Implications: Optimal \( V_{dd} \) and \( V_{T} \)

• No single pair of optimal \( V_{dd}, V_{T} \)
  - Depends on performance, power target
  - High performance: high \( V_{dd} \), low \( V_{T} \)
  - Low power: low \( V_{dd} \), high \( V_{T} \)

• However, optimal \( P_{\text{leak}} / P_{\text{dyn}} \) is roughly constant
  - Typically, \( P_{\text{leak}} \sim 30 - 50\% \) of \( P_{\text{dyn}} \)
Pipelining

Parallelism