ECE 471 – Embedded Systems
Lecture 21

Vince Weaver

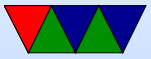
http://www.eece.maine.edu/~vweaver

vincent.weaver@maine.edu

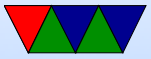
20 November 2014

Announcements

- Project groups and topic!



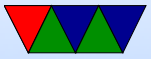
Power and Energy



Definitions

People often say Power when they mean Energy

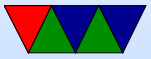
- Dynamic Power – only consumed while computing
- Static Power – consumed all the time.
  Sets the lower limit of optimization



Units

- Energy – Joules, BTU (1055J), HorsePower (2.6845MJ), kWH (3.6MJ), Therm (105.5MJ), Calorie (4.184J), 1g TNT (4184J), eV $= 1.6 \times 10^{-19}$J
  Related but not energy: mAh (batteries)

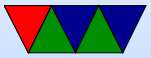
- Power – Energy/Time – Watts, Volt-Amps (for A/C)



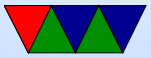
Power and Energy in a Computer System

Mahersi and Vardhan, PACS'04.

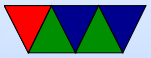
- Measured V and Current. P=IIR. V=IR P=IV, subtractive
- Thinkpad Laptop, Pentium M, 256M, 14" display
- Total System Power 14-30W
- Hard Drive 0.5-2W (Flash/SSD less)
- LCD 1W (slightly more black than white)



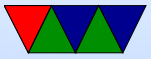
- Backlight Inverter (this is before LED) 1-4W depending on brightness
- CPU 2-15W (with scaling)
- GPU 1-5W
- Memory 0.45 - 1.5W
- Power Supply Loss - 0.65W
- Wireless 0.1 - 3W (wifi on cellphones)
- CDROM 3-5W



- (USB 2.0 – 5V, can draw 5 units of 100mA each, 2.5W)



CPU Power and Energy

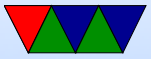


CMOS Dynamic Power

- $P = C\Delta V V_{dd} \alpha f$
  Charging and discharging capacitors big factor ($C\Delta V V_{dd}$) from V_{dd} to ground
  α is activity factor, transitions per clock cycle
  f is frequency

- α often approximated as $\frac{1}{2}$, $\Delta V V_{dd}$ as V_{dd}^2 leading to $P \approx \frac{1}{2} C V_{dd}^2 f$

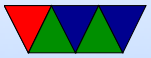
- Some pass-through loss (V momentarily shorted)



CMOS Dynamic Power Reduction

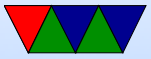
How can you reduce Dynamic Power?

- Reduce C – scaling
- Reduce V_{dd} – eventually hit transistor limit
- Reduce α (design level)
- Reduce f – makes processor slower



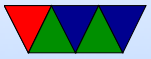
Metrics to Optimize

- Power
- Energy
- MIPS/W, FLOPS/W (don't handle quadratic V well)
- $Energy * Delay$
- $Energy * Delay^2$



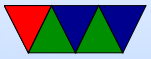
Power Optimization

- Does not take into account time. Lowering power does no good if it increases runtime.



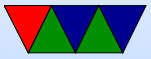
Energy Optimization

- Lowering energy can affect time too, as parts can run slower at lower voltages



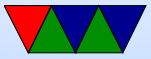
Energy Delay – Watt/t*t

- Horowitz, Indermaur, Gonzalez (Low Power Electronics, 1994)
- Need to account for delay, so that lowering Energy does not made delay (time) worse
- Voltage Scaling – in general scaling low makes transistors slower
- Transistor Sizing – reduces Capacitance, also makes transistors slower



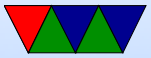
- Technology Scaling – reduces V and power.
- Transition Reduction – better logic design, have fewer transitions
  Get rid of clocks? Asynchronous? Clock-gating?
- Example with inverse ED (higher better):

Alpha 21064	SPEC=155	Power=30W	SPEC*SPEC/W=800
PPC603	SPEC=80	Power=3W	SPEC*SPEC/W=2100

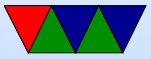


Energy Delay Squared– E*t*t

- Martin, Nyström, Pénzes – Power Aware Computing, 2002
- Independent of Voltage in CMOS
- Et can be misleading
  Ea=2Eb, ta=tB/2
  Reduce voltage by half, Ea=Ea/4, ta=2ta, Ea=Eb/2, ta=tb
- Can have arbitrary large number of delay terms in Energy



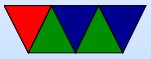
product, squared seems to be good enough



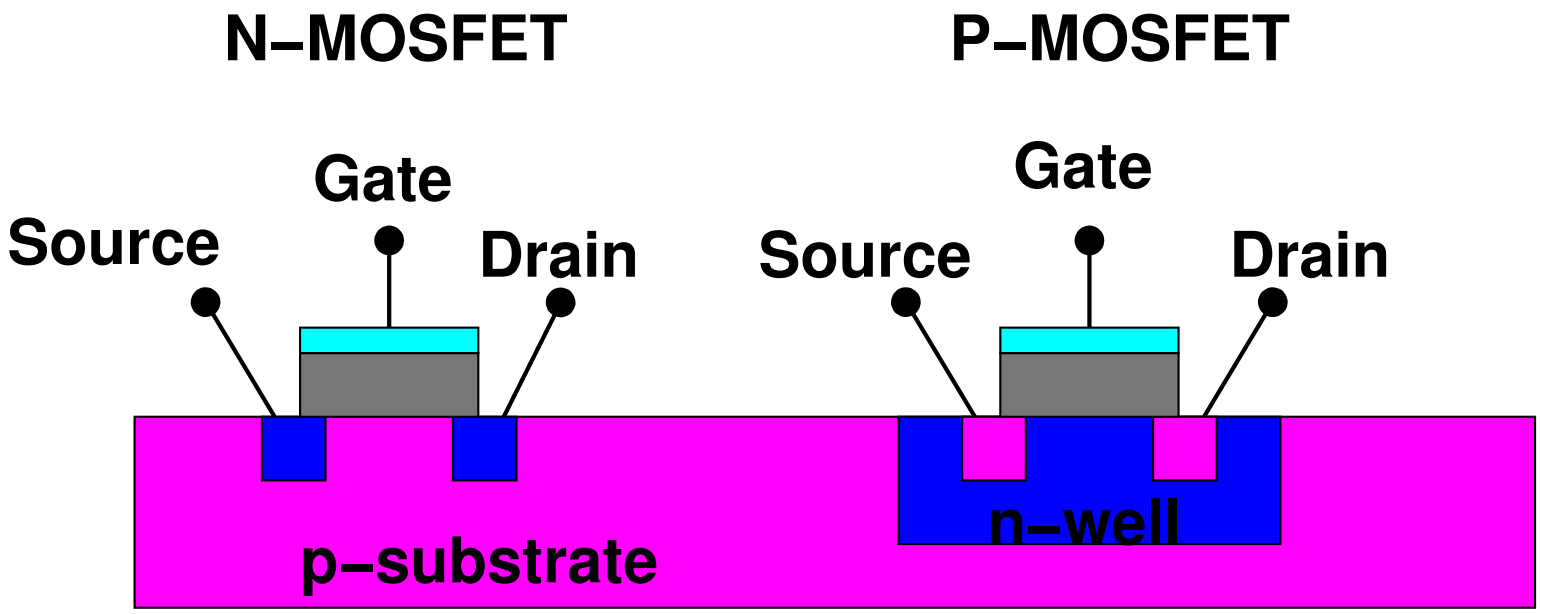
CMOS Dynamic Power Review

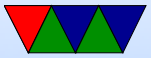
- $P = C\Delta V V_{dd} \alpha f$
  Current is from charging/discharging capacitors ($C\Delta V V_{dd}$)
  So $P = IV = C\Delta V V_{dd}$
  The number of transitions is α (activity factor), times clock frequency (f).

- α is often approximated as 1/2, $\Delta V V_{dd}$ as V_{dd}^2
  $P = \frac{1}{2} C V_{dd}^2 f$



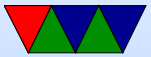
CMOS Transistors





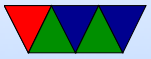
CMOS Static Power

- Leakage Current – bigger issue as scaling smaller. Forecast at one point to be 20-50% of all chip power before mitigations were taken.
- Various kinds of leakage (Substrate, Gate, etc)
- Linear with Voltage: $P_{static} = I_{leakage} V_{dd}$



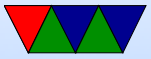
Leakage Mitigation

- SOI – Silicon on Insulator (AMD, IBM but not Intel)
- High-k dielectric – instead of SO2 use some other material for gate oxide (Hafnium)
- Transistor sizing – make only critical transistors fast; non-critical can be made slower and less leakage prone
- Body-biasing
- Sleep transistors



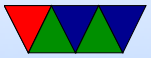
Total Energy

- $E_{tot} = [P_{dyanmic} + P_{static}]t$
- $E_{tot} = [(C_{tot}V_{dd}^2\alpha f) + (N_{tot}I_{leakage}V_{dd})]t$



Delay

- $T_d = \frac{C_L V_{dd}}{\mu C_{ox}(\frac{W}{L})(V_{dd}-V_t)}$
- Simplifies to $f_{MAX} \sim \frac{(V_{dd}-V_t)^2}{V_{dd}}$
- If you lower f, you can lower V_{dd}



Thermal Issues

- Temperature and Heat Dissipation are closely related to Power
- If thermal issues, need heatsinks, fans, cooling