

ECE 574 – Cluster Computing

Lecture 22

Vince Weaver

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

vincent.weaver@maine.edu

24 November 2015

Announcements

- Project groups status report due:
 1. Short summary of project
 2. Are things going well? Will you finish on time?
 3. Are you willing to present on Tuesday rather than Thursday, or do you need to otherwise present early?



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, kWh (3.6MJ), Therm (105.5MJ), 1 Ton TNT (4.2GJ), eV (1.6×10^{-19} J), BTU (1055 J), horsepower-hour (2.68 MJ), calorie (4.184 J)
- Power – Energy/Time – Watts (1 J/s), Horsepower (746W), Ton of Refrigeration (12,000 Btu/h)
- Volt-Amps (for A/C) – same units as Watts, but not same thing
- Charge – mAh (batteries) – need voltage to convert to Energy



CPU Power and Energy



CMOS Dynamic Power

- $P = C\Delta VV_{dd}\alpha f$

Charging and discharging capacitors big factor
($C\Delta VV_{dd}$) from V_{dd} to ground

α is activity factor, transitions per clock cycle

f is frequency

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

- 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



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: $P_{static} = I_{leakage}V_{dd}$



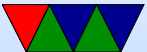
Leakage Mitigation

- SOI – Silicon on Insulator (AMD, IBM but not Intel)
- High-k dielectric – instead of SiO_2 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



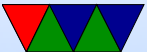
Total Energy

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



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



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.



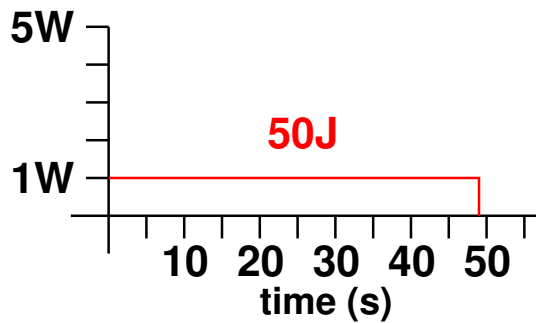
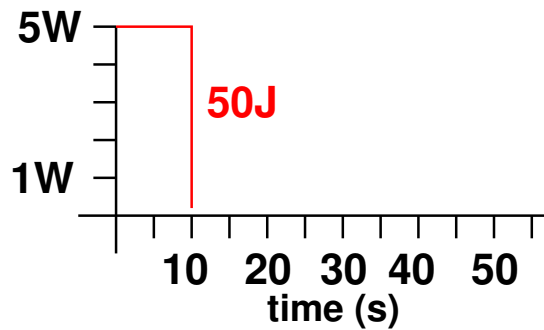
Energy Optimization

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



Energy Optimization

Which is better?



Energy Delay – $\text{Watt}/t * t$

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

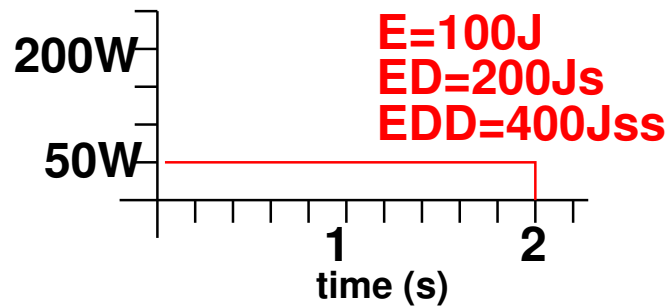
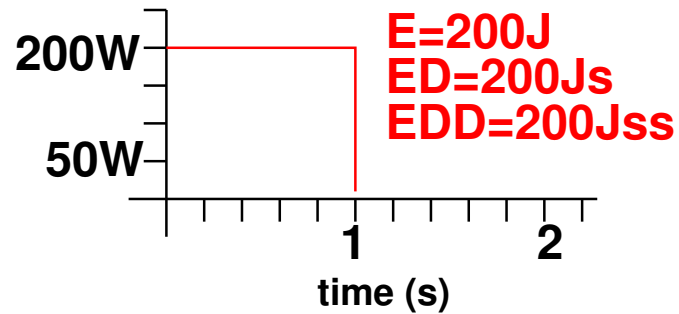


- Technology Scaling – reduces V and power.
- Transition Reduction – better logic design, have fewer transitions
Get rid of clocks? Asynchronous? Clock-gating?



ED Optimization

Which is better?



Energy Delay Squared– $E \cdot t \cdot t$

- Martin, Nyström, Péntzes – Power Aware Computing, 2002

- Independent of Voltage in CMOS

- ED can be misleading

$$E_a = 2E_b, t_a = \frac{t_b}{2}$$

Reduce voltage by half, $E_a = \frac{E_a}{4}, t_a = 2t_a, E_a = \frac{E_b}{2},$

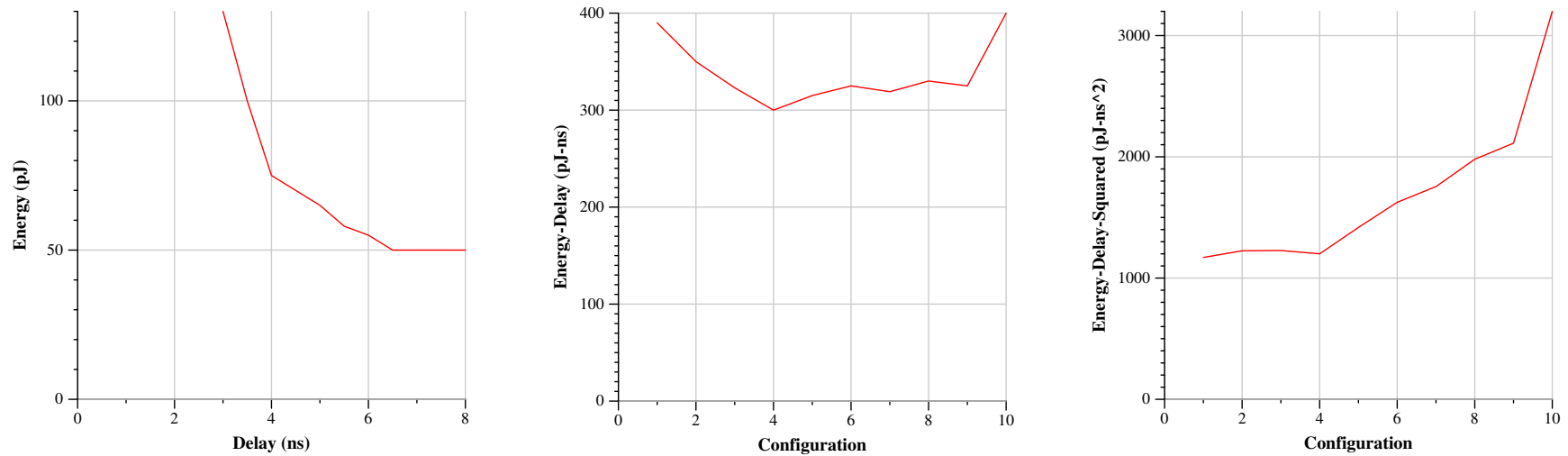
$$t_a = t_b$$



- Can have arbitrary large number of delay terms in Energy product, squared seems to be good enough



Energy-Delay Product Redux



Roughly based on data from “Energy-Delay Tradeoffs in CMOS Multipliers” by Brown et al.



Raw Data

Delay	Energy	ED	ED^2
3	130	390	1170
3.5	100	350	1225
3.8	85	323	1227
4	75	300	1200
4.5	70	315	1418
5	65	325	1625
5.5	58	319	1755
6	55	330	1980
6.5	50	390	2535
8	50	400	3200



Other Metrics

- $Energy - Delay^n$ – choose appropriate factor
- $Energy - Delay - Area^2$ – takes into account cost (die area) [McPAT]
- Power-Delay – units of Energy – used to measure switching
- Energy Delay Diagram – [SWEEP]



Measuring Power and Energy



Why?

- New, massive, HPC machines use impressive amounts of power
- When you have 100k+ cores, saving a few Joules per core quickly adds up
- To improve power/energy draw, you need some way of measuring it



Energy/Power Measurement is Already Possible

Three common ways of doing this:

- Hand-instrumenting a system by tapping all power inputs to CPU, memory, disk, etc., and using a data logger
- Using a pass-through power meter that you plug your server into. Often these will log over USB
- Estimating power/energy with a software model based on system behavior

