EECS 470
Power and Architecture

Many slides taken from Prof. David Brooks, Harvard University and modified by Mark Brehob . A couple of slides are also taken from Prof. Wenisch. Any errors are almost certainly Mark's.

Thanks to both!

Outline

- **Why is power a problem?**

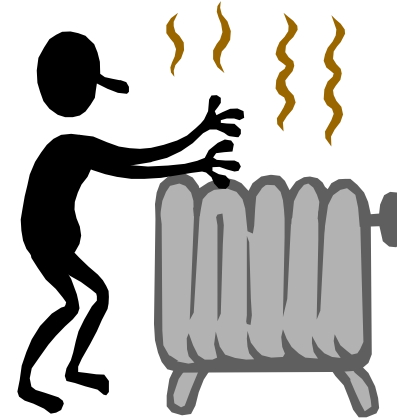- **What uses power in a chip?**

- **How can we reduce power?**

Outline

Introduction

Why is power a problem in a µP?

- **Power used by the µP, vs. system power**

- **Dissipating Heat**

  - **Melting (very bad)**

  - **Packaging (to cool → $)**

  - **Heat leads to poorer performance.**

- **Providing Energy**
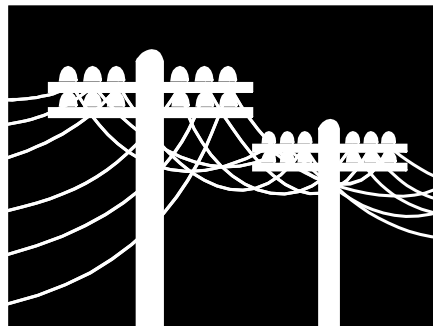
  - **Battery**

  - **Cost of electricity**

Why worry about power dissipation?

Battery life

Thermal issues: affect cooling, packaging, reliability, timing

Environment

Where does power go?

- **Obviously desktop and servers are going to be different.**
  - **But if CPUs are a small fraction of the power, maybe we don't care much?**

- **Let's take a look.**
  - **But first a few caveats**
    - **I can't find recent studies on this**
    - **I do find that different studies can get radically different results. I'm using some fairly well-discussed numbers.**
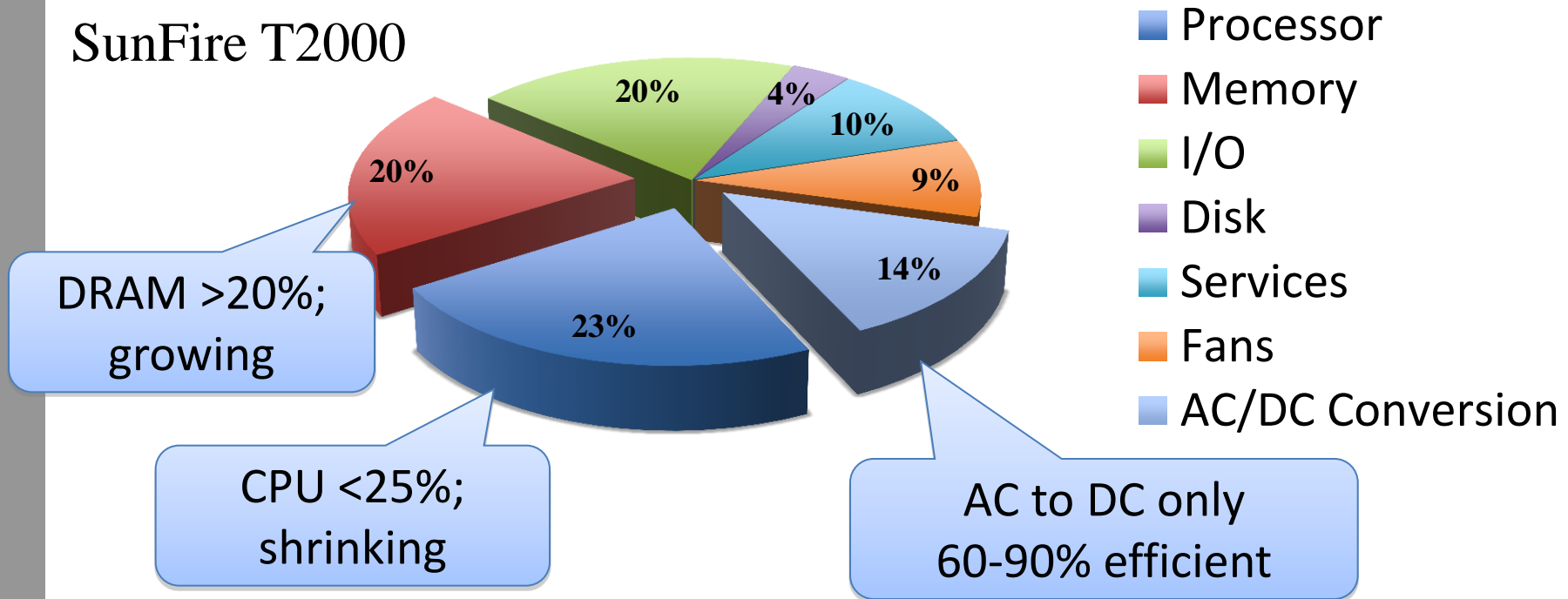
Where does the juice go in laptops?

Network, 4%

Hard Drive, 5%

Graphics, 8%

Processor, 9%

Processor, 9%
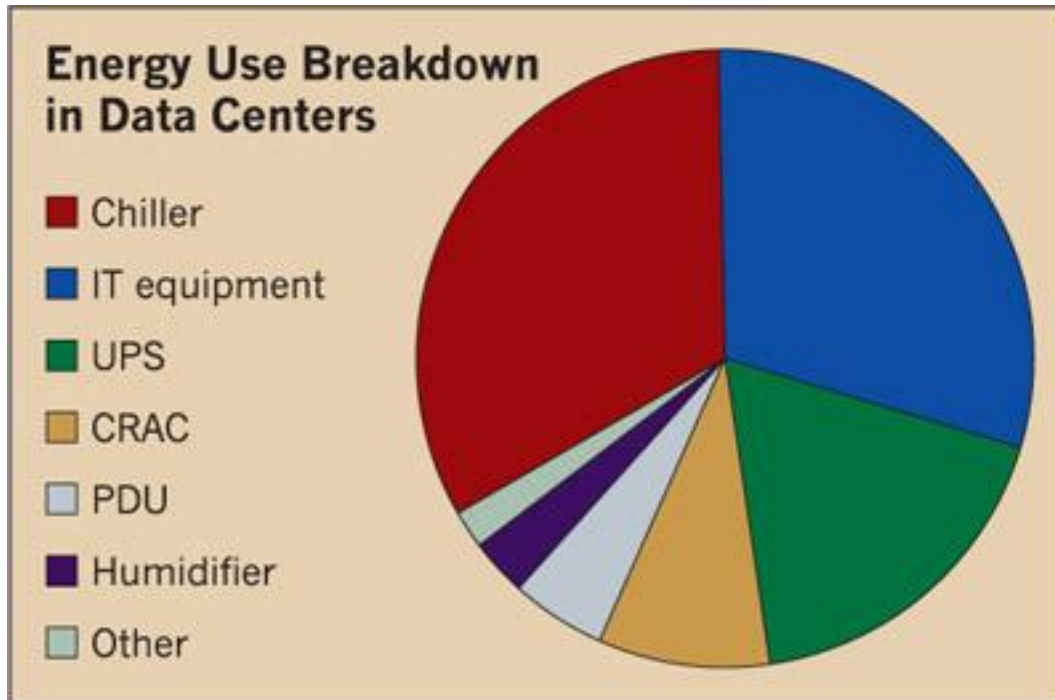
LCD Panel, 43%

Chipset, 21%
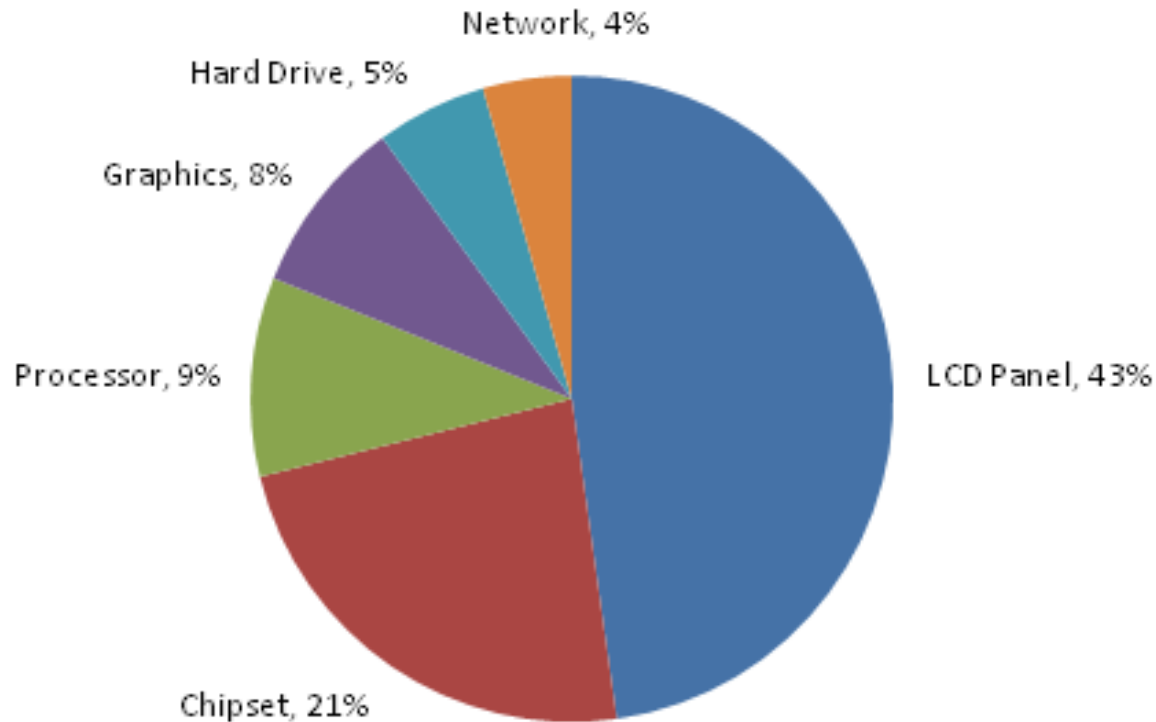
- **Microsoft, 2009.**
- **The processor power can be a lot higher depending on wha the laptop is doing. [Hsu+Kremer, 2002]**

What about servers?

SunFire T2000

Legend:
- Processor
- Memory
- I/O
- Disk
- Services
- Fans
- AC/DC Conversion

Pie chart values: 20%, 20%, 4%, 10%, 9%, 14%, 23%

DRAM >20%; growing

CPU <25%; shrinking

AC to DC only 60-90% efficient

Need whole-system approaches to save energy

Energy Use Breakdown in Data Centers

- Chiller
- IT equipment
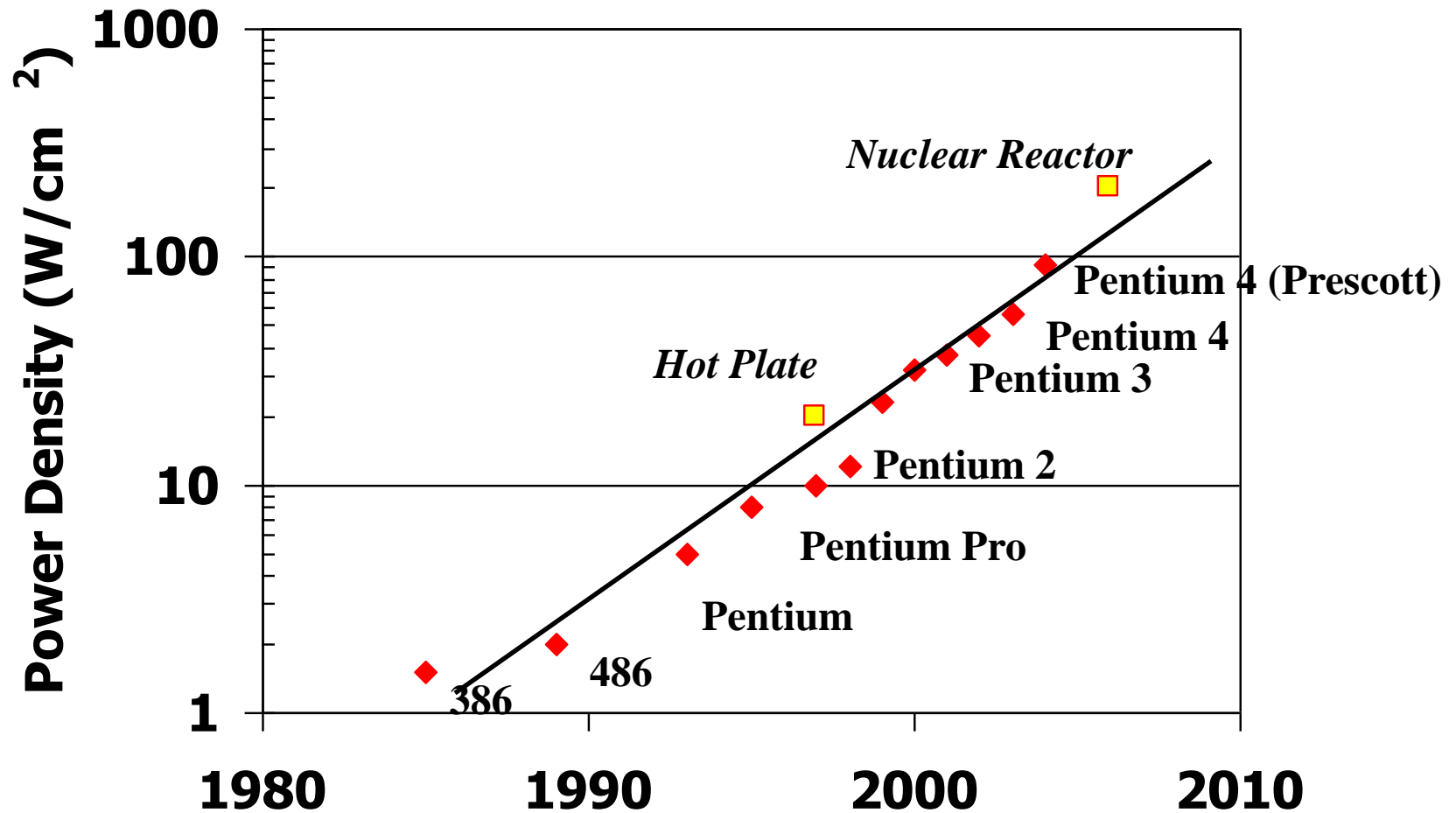- UPS
- CRAC
- PDU
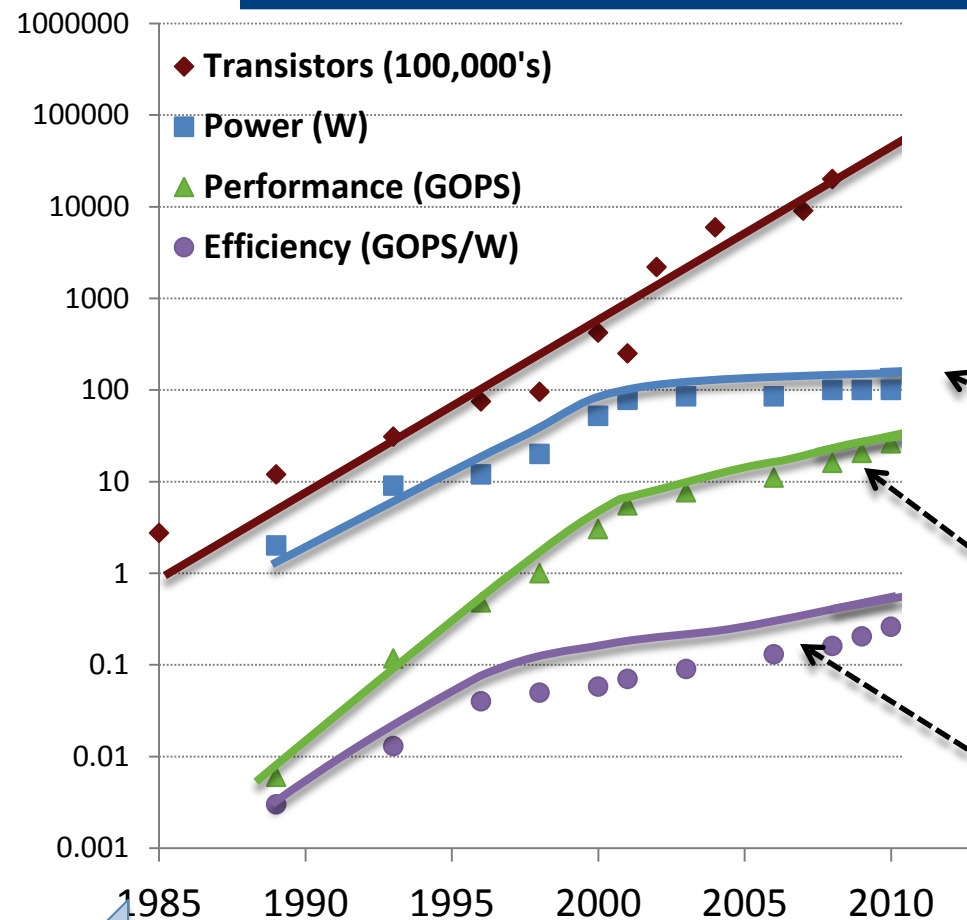- Humidifier
- Other

Power usage effectiveness (PUE)
A PUE of 2.0 means that for every 2W of power supplied to the data center, only 1W is consumed by computing equipment. Values of around 2.9 are pretty common

Total Power Dissipation Trends

Why is power a problem?

A Paradigm Shift In Computing

Spot Heat Issues in Microprocessors

Why is power a problem?

Temperature (°C)

82.007
80.1173
78.2277
76.338
74.4483
72.5587
70.669
68.7793
66.8897

Data center energy use

Installed base grows 11%/yr.

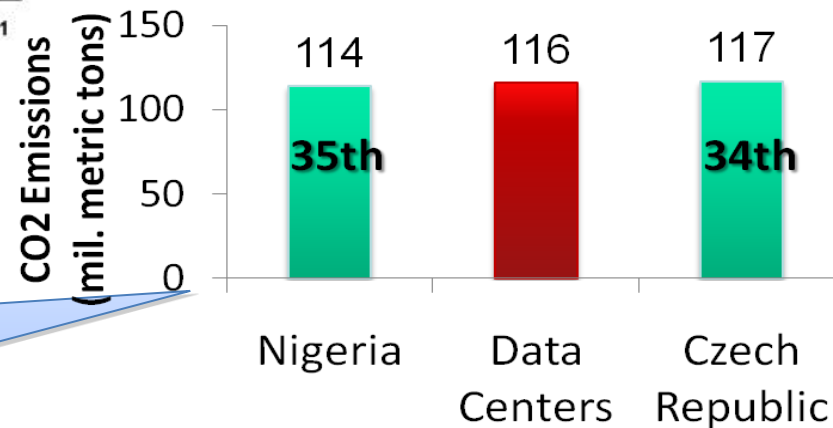Future energy use projections →
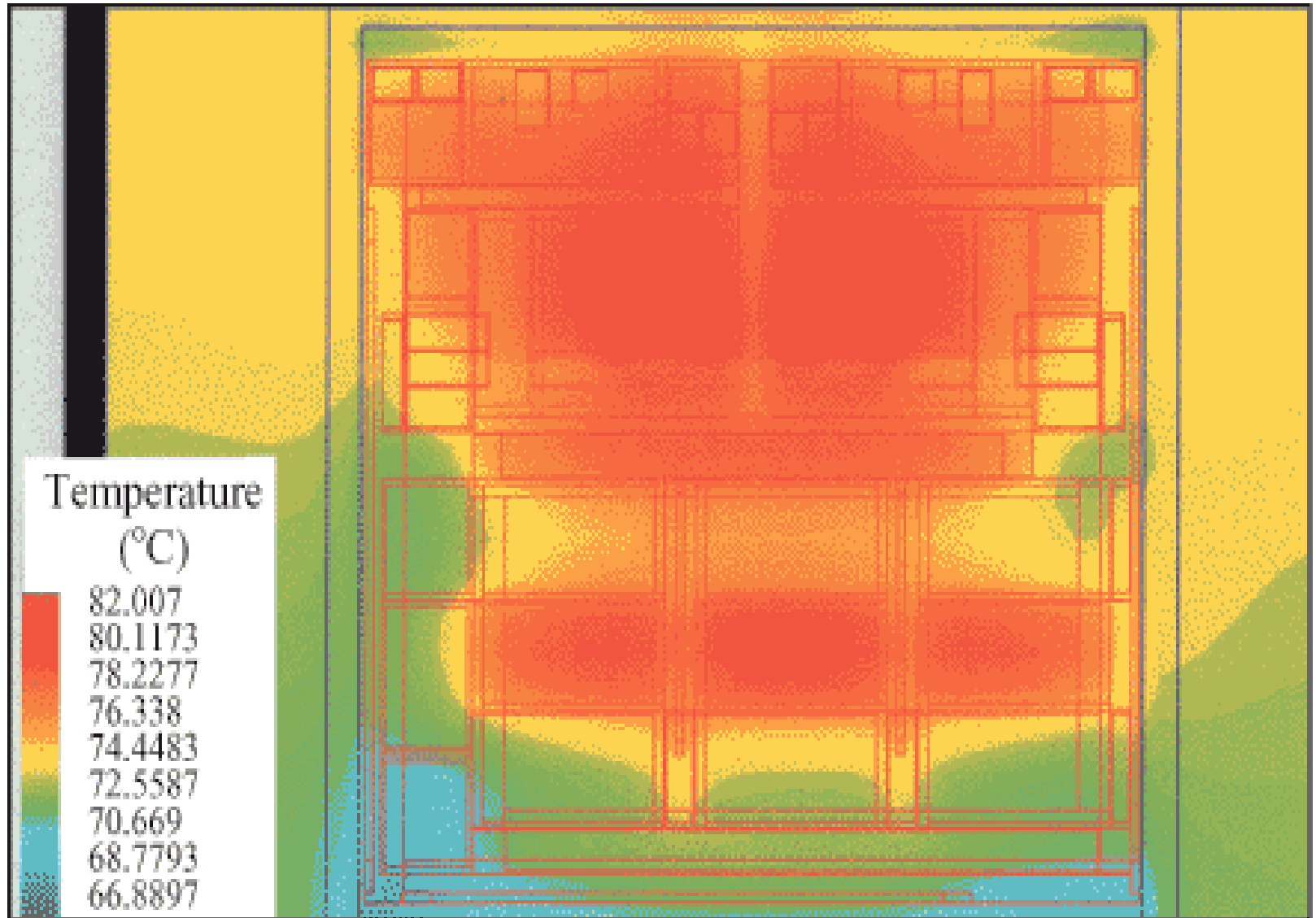
← Historical energy use
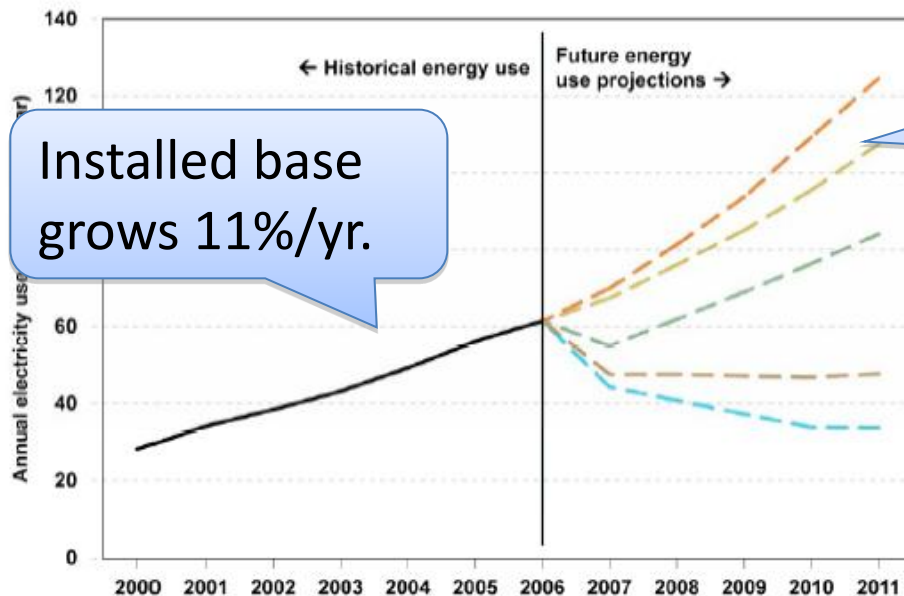
Source: US EPA 2007—Newest I can find

In 2012,
~2.0% of US energy
~$7 billion/yr.
(Rich Miller, 2012)

Source: Mankoff et al, IEEE Computer 2008
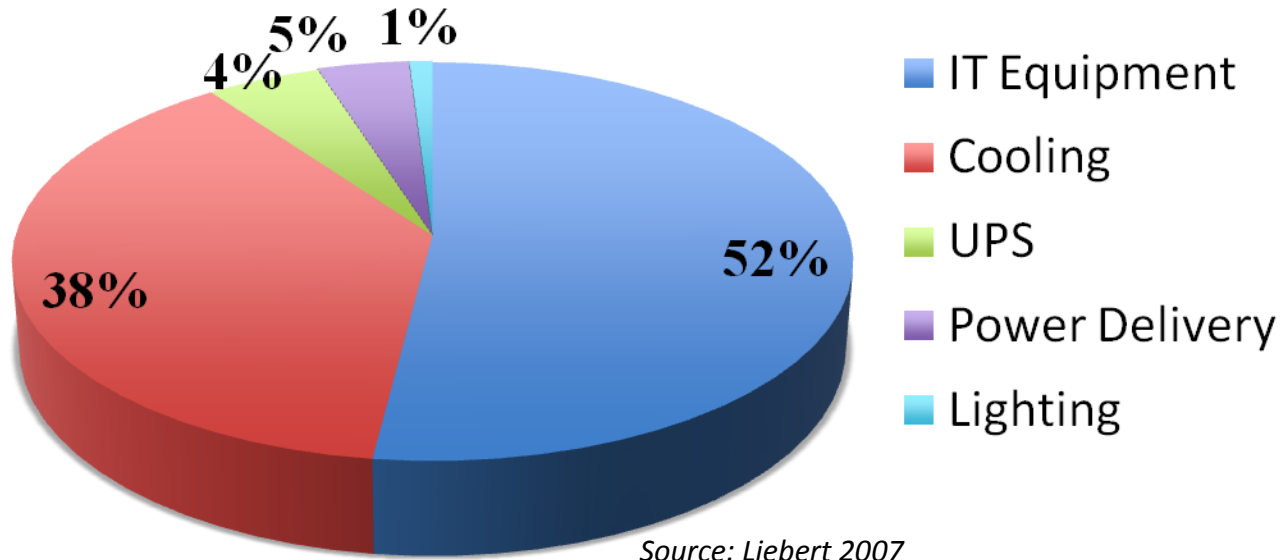
0.5% of world CO_2 emissions; rivals entire Czech Republic

CO2 Emissions (mil. metric tons)

114 — Nigeria — 35th
116 — Data Centers
117 — Czech Republic — 34th

Improving energy efficiency is a critical challenge

13

Where does all the power go?

Pie chart legend:
- IT Equipment
- Cooling
- UPS
- Power Delivery
- Lighting

52% IT Equipment, 38% Cooling, 4% UPS, 5% Power Delivery, 1% Lighting
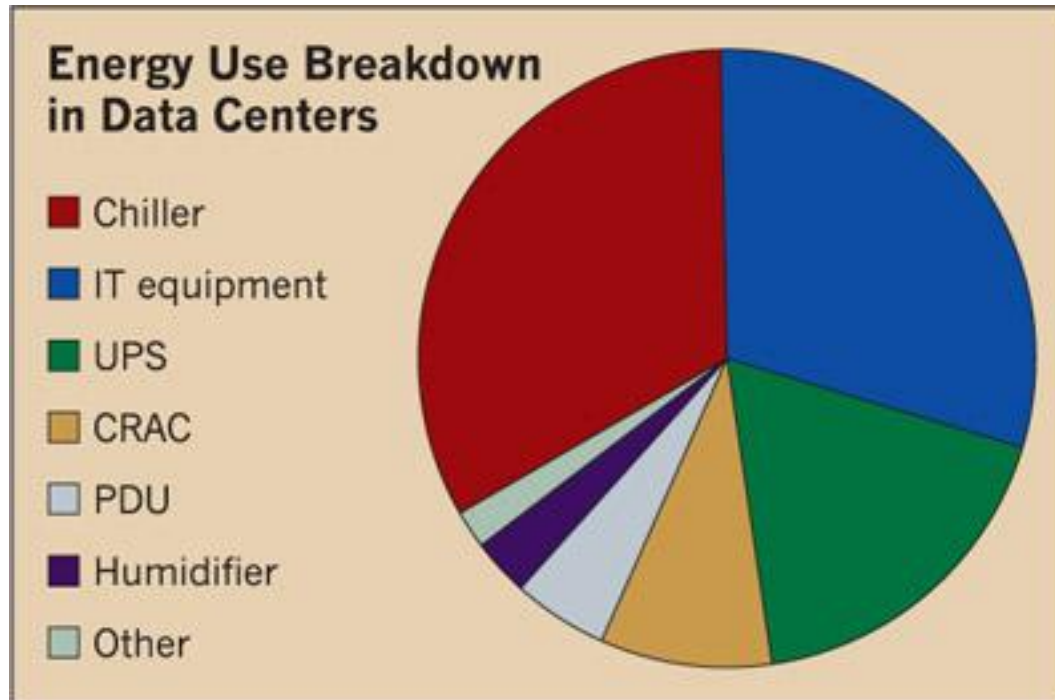
Source: Liebert 2007

Servers account for barely half of power

- 1W of cooling per 1.5W of IT load
- 10MW data center: cooling costs $4M to $8M / yr.

System designers must think about cooling

This may be getting worse.

Energy Use Breakdown in Data Centers

- Chiller
- IT equipment
- UPS
- CRAC
- PDU
- Humidifier
- Other

- **UPS**: Power supplies/converters
- **CRAC**: Computer Room Air Conditioner.
- **PDU**: Power distribution unit

Power Usage Effectiveness (PUE)
A PUE of 2.0 means that for every 2W of power supplied to the data center, only 1W is consumed by computing equipment. Values of around 2.8 are pretty common these days, while 1.9 was a common number in 2007.
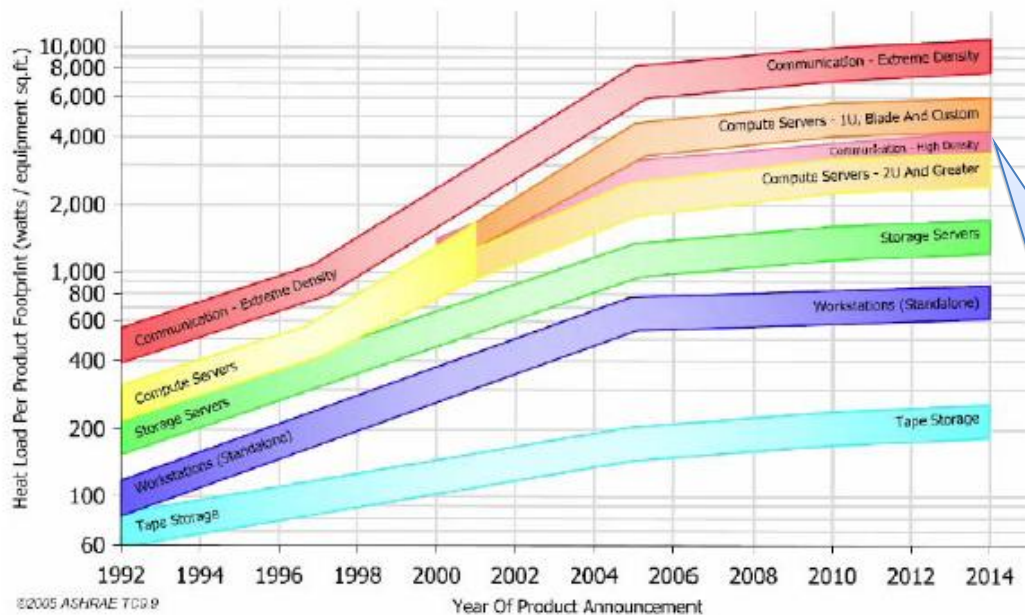- Not clear why.

http://ecmweb.com/energy-efficiency/data-center-efficiency-trends

Server density increasing

- **Integration, disaggregation reduce hardware costs**

- **Need for high-BW interconnect**
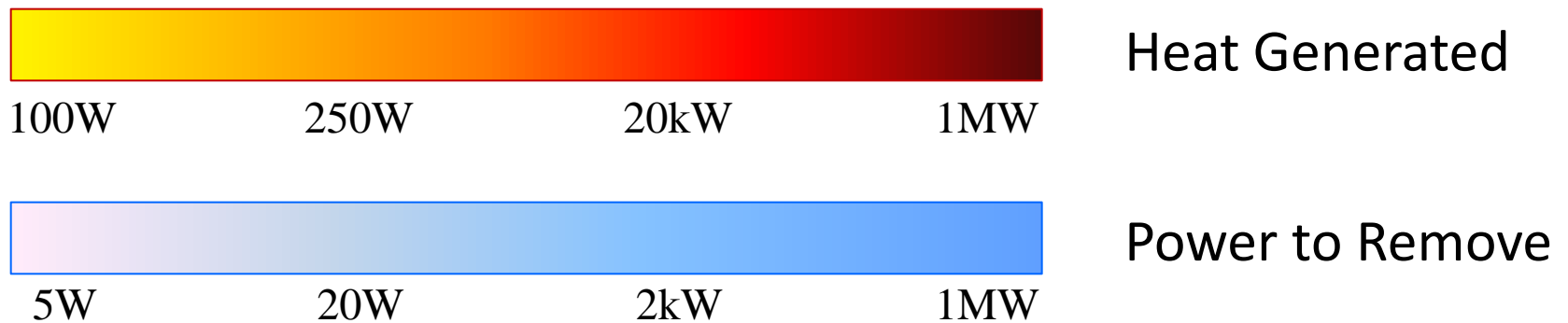
- **Data center floor space costs up to $15,000 /m^2**

Heat flux up to
500W per ft^2 floor space;
Racks draw 4 to 20kW each

Source: AHRAE 2006

Heat density drives cooling cost

Cooling power grows super-linearly with thermal load

Heat Generated

100W   250W   20kW   1MW
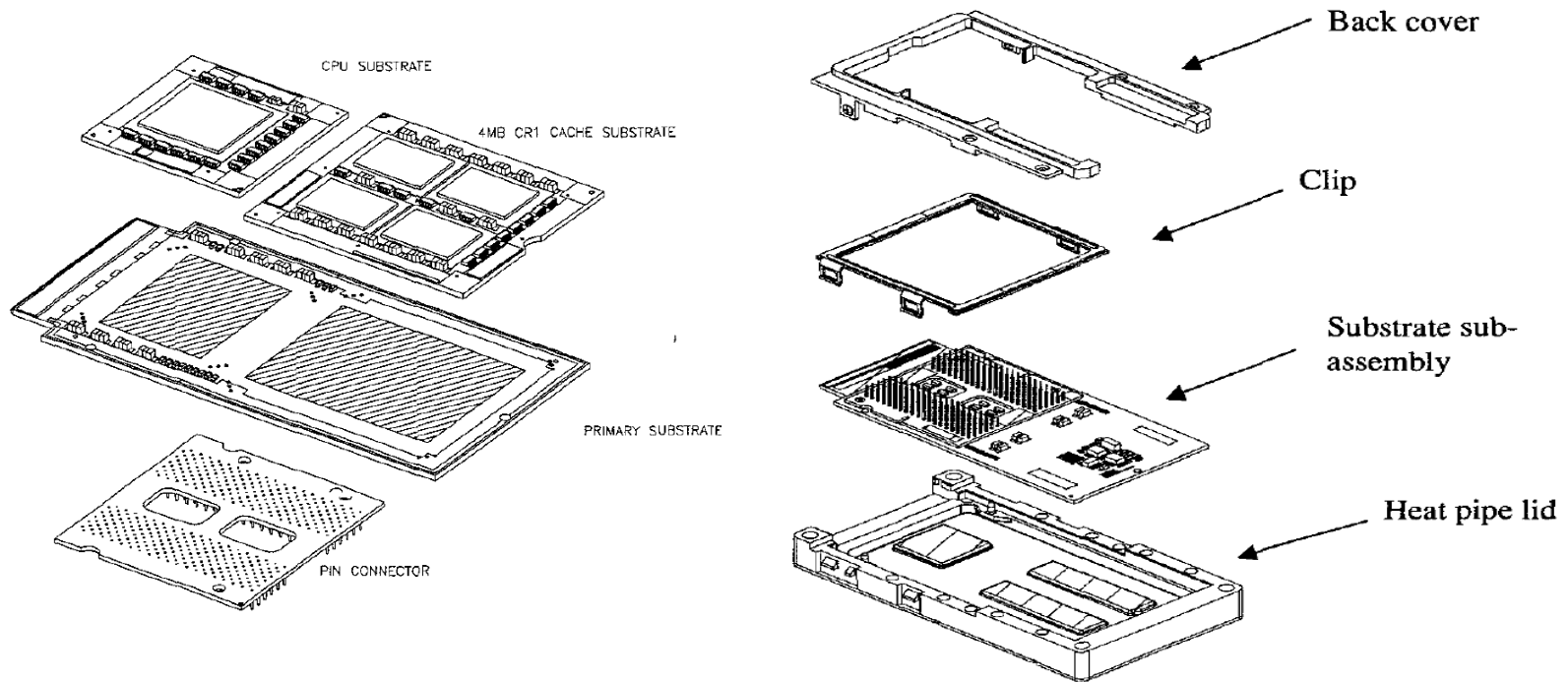
Power to Remove

5W   20W   2kW   1MW

Source: C. Patel, HP Labs

Servers' 3-year power & cooling costs nearing their purchase price

Intel Itanium packaging
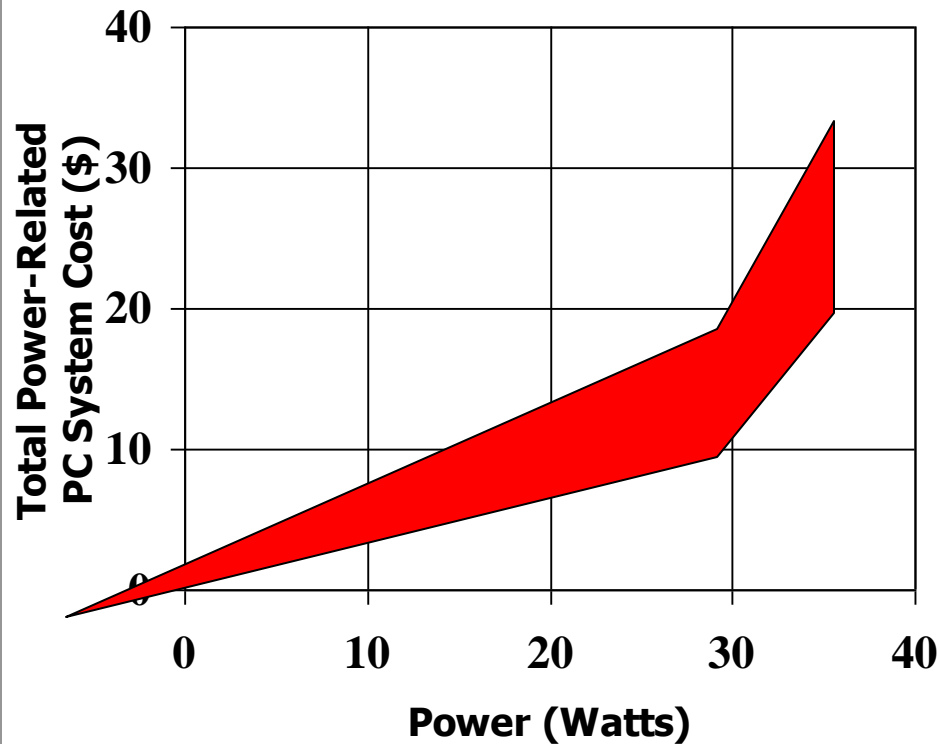
Complex and expensive (note heatpipe)

Source: H. Xie et al. "Packaging the Itanium Microprocessor"
Electronic Components and Technology Conference 2002

P4 packaging

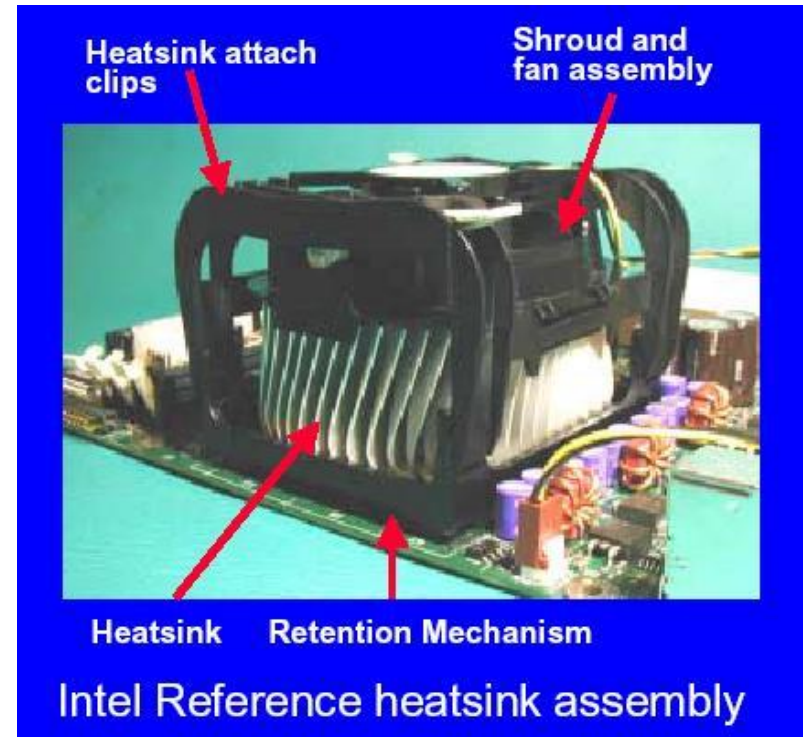- ## Simpler, but still…

Chart: Power (Watts) vs. Total Power-Related PC System Cost ($)

From Tiwari, et al., DAC98

Intel Reference heatsink assembly
- Heatsink attach clips
- Shroud and fan assembly
- Heatsink
- Retention Mechanism

Source: Intel web site

Power-Aware Computing Applications

Temperature/di-dt-Constrained

Energy-Constrained Computing

Environment

- **Environment Protection Agency (EPA): computers consume 10% of commercial electricity consumption**
  - **This incl. peripherals, possibly also manufacturing**
  - **A DOE report suggested this percentage is much lower (3.0-3.5%)**
  - **No consensus, but it's still a lot**
  - **Interesting to look at the numbers:**
    - **http://enduse.lbl.gov/projects/infotech.html**
- **Data center growth was cited as a contribution to the 2000/2001 California Energy Crisis**
- **Equivalent power (with only 30% efficiency) for AC**
- **CFCs used for refrigeration**
- **Lap burn**
- **Fan noise**

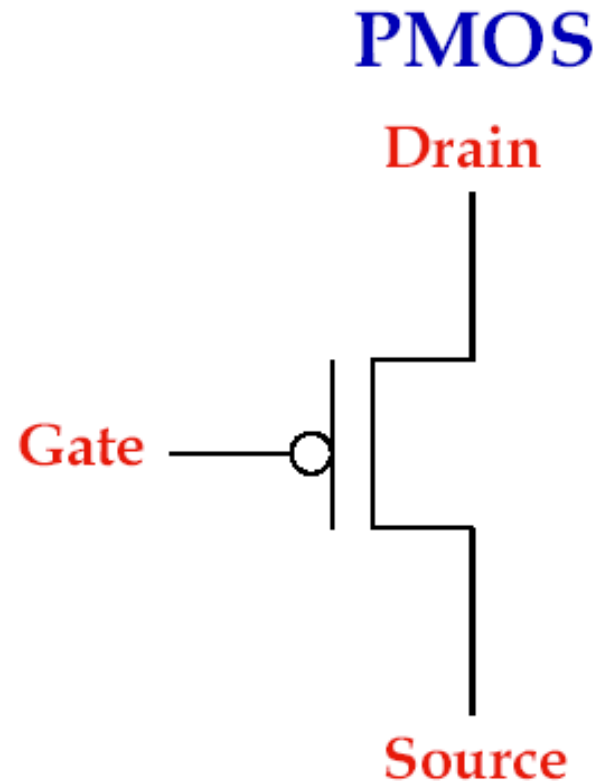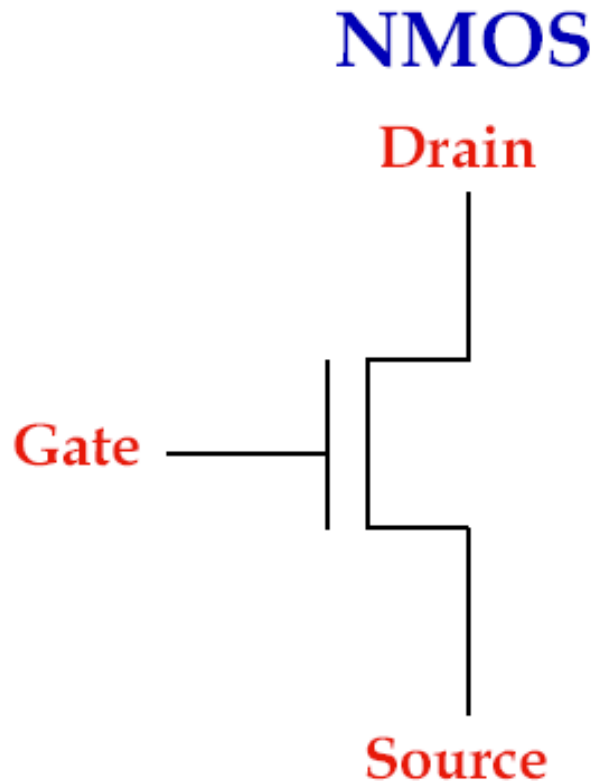# Power-Aware Needed across all computing platforms

- **Mobile/portable (cell phones, laptops, PDA)**

  - **Battery life is critical**

- **Desktops/Set-Top (PCs and game machines)**

  - **Packaging cost is critical**

- **Servers (Mainframes and compute-farms)**

  - **Packaging limits**
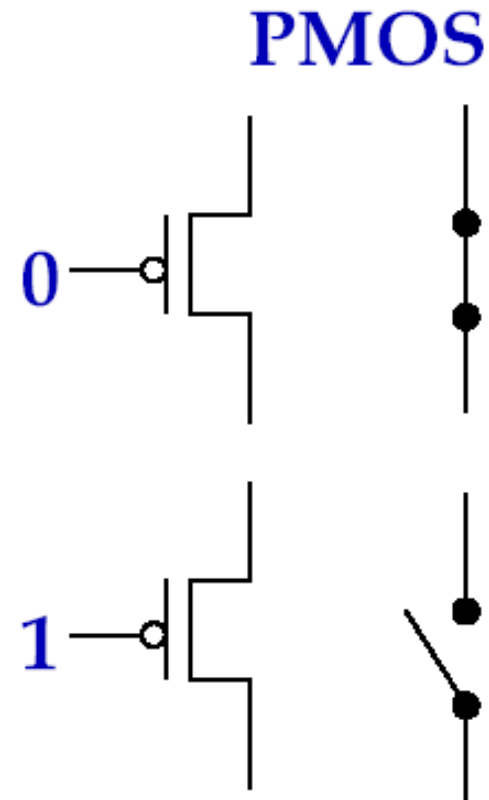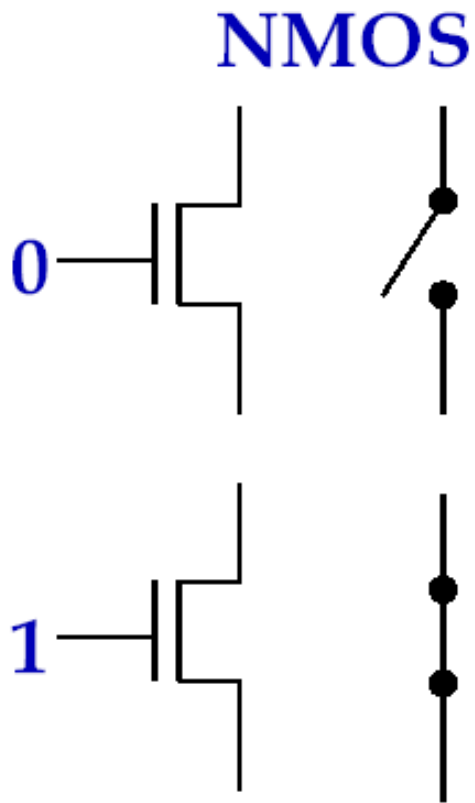 - **Volumetric (performance density)**

What uses power in a chip?

How CMOS Transistors Work

What uses power in a chip?

NMOS

Drain

Gate

Source

PMOS

Drain

Gate

Source

What uses power in a chip?

NMOS

PMOS

Static CMOS

V_{DD}

In1
In2
In3

PUN

PMOS Only

$F = \overline{G}$

In_1
In_2
In_3

PDN

NMOS Only

V_{SS}
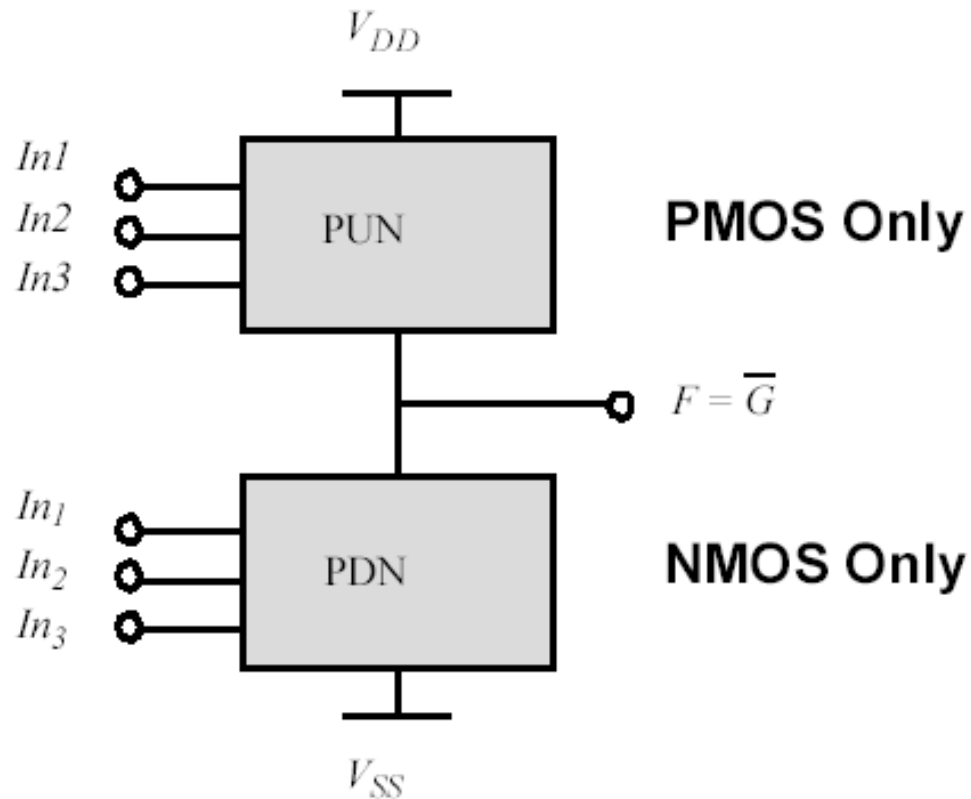
PUN and PDN are Dual Networks

Basic Logic Gates

OUT = $\overline{A}$

V_{DD}

OUT = $\overline{A \cdot B}$

OUT = $\overline{A + B}$
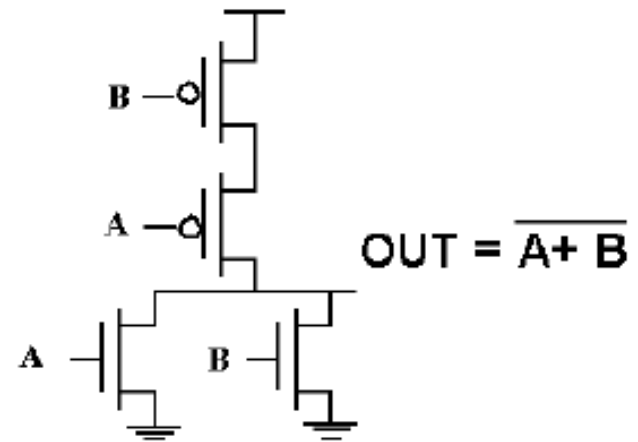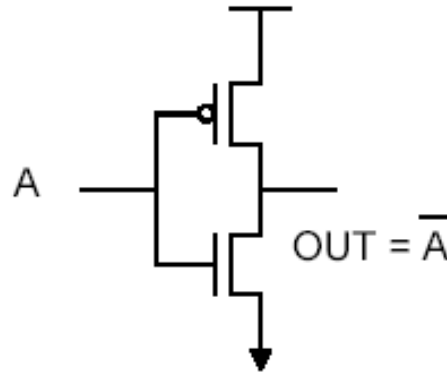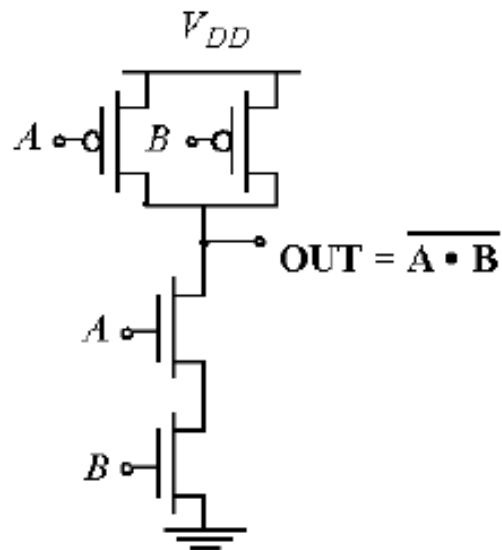
CMOS Water Analogy

Electron: water molecule

Charge: weight of water

Voltage: height

Current: flow rate

Capacitance: container cross-section

(Think of power-plants that store energy in water towers)

Liquid Inverter

Vin = 0   Vout = 1

switching threshold
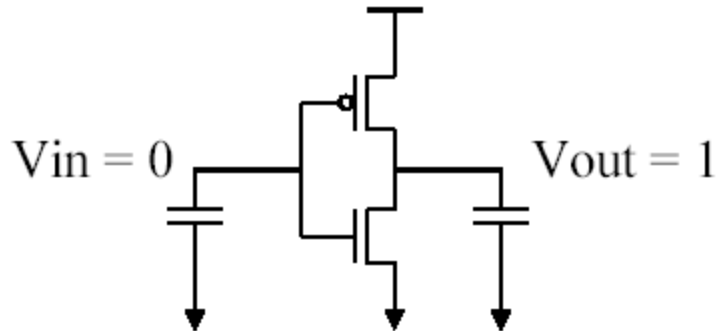
- **Capacitance at input**
  - **Gates of NMOS, PMOS**
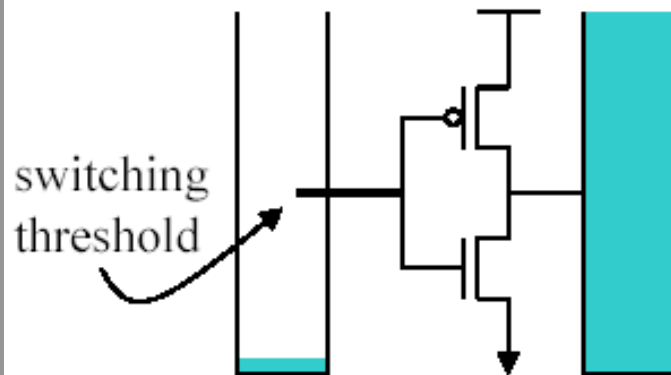 - **Metal interconnect**
- **Capacitance at output**
  - **Fanout (# connections) to other gates**
  - **Metal Interconnect**

NMOS conducts when water level is above switching threshold

PMOS conducts below

No conduction after container full

Slide courtesy D. Brooks, Harvard
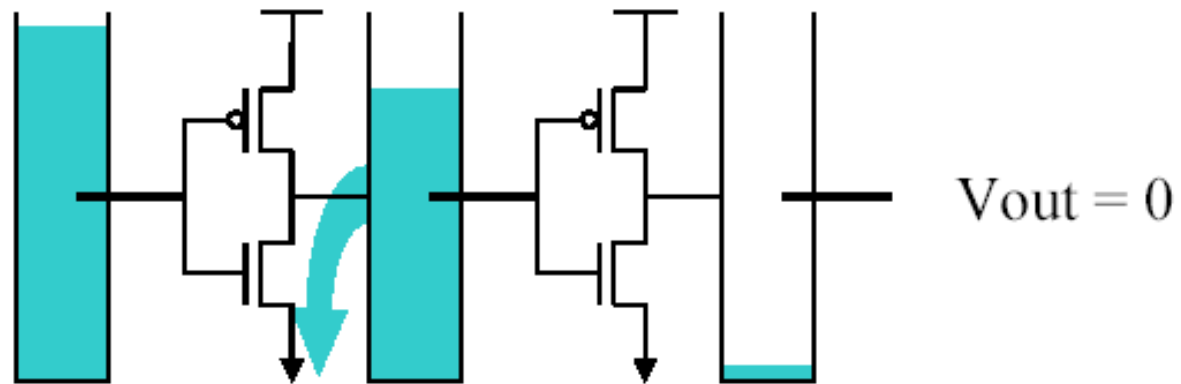
t < 0
Vin = 0

Vout = 0

t = 0
Vin = 1

Vout = 0

Slide courtesy D. Brooks, Harvard

t = 1
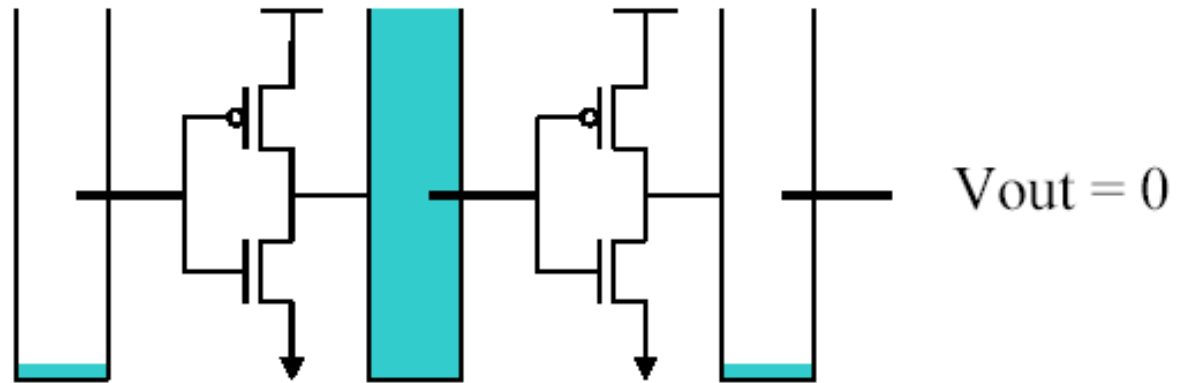Vin = 0

Vout = 0

t = 2
Vin = 1
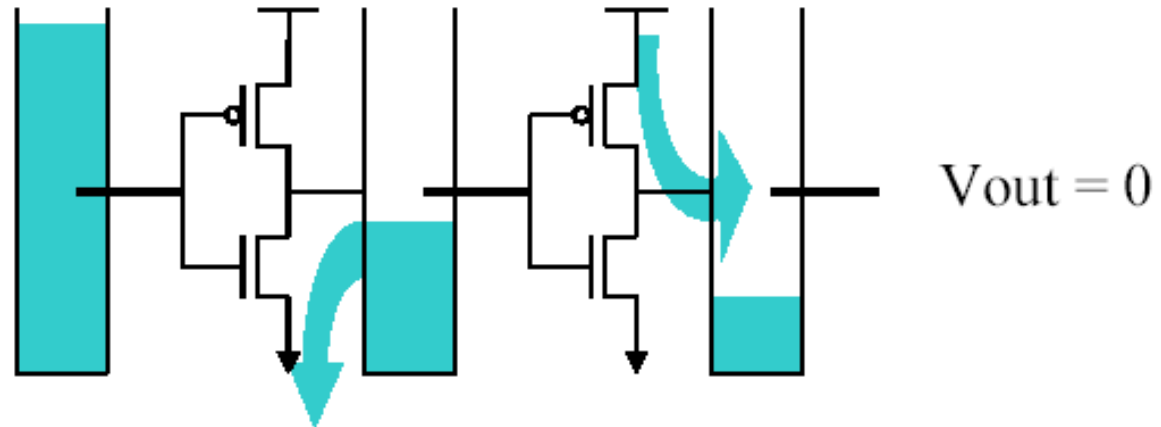
Vout = 1
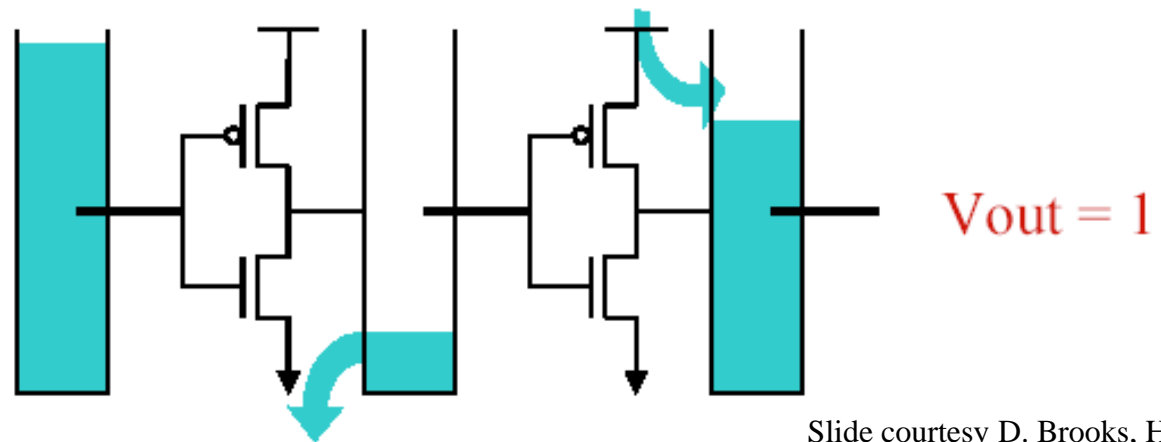
Slide courtesy D. Brooks, Harvard

Delay and Power Observations

- **Load capacitance increases delay**
  - **High fanout (gates attached to output)**
  - **Interconnection**

- **Higher current can increase speed**
  - **Increasing transistor width raises currents but also raises capacitance**

- **Energy per switching event independent of current**
  - **Depends on amount of charge moved, not rate**

What uses power in a chip?

Power: The Basics

- **Dynamic power vs. Static power**
  - Dynamic: "switching" power
  - Static: "leakage" power
  - Dynamic power dominates, but static power increasing in importance

- **Static power: steady, per-cycle energy cost**
- **Dynamic power: capacitive and short-circuit**
- **Capacitive power: charging/discharging at transitions from 0→1 and 1→0**
- **Short-circuit power: power due to brief short-circuit current during transitions.**

What uses power in a chip?

Dynamic (Capacitive) Power Dissipation
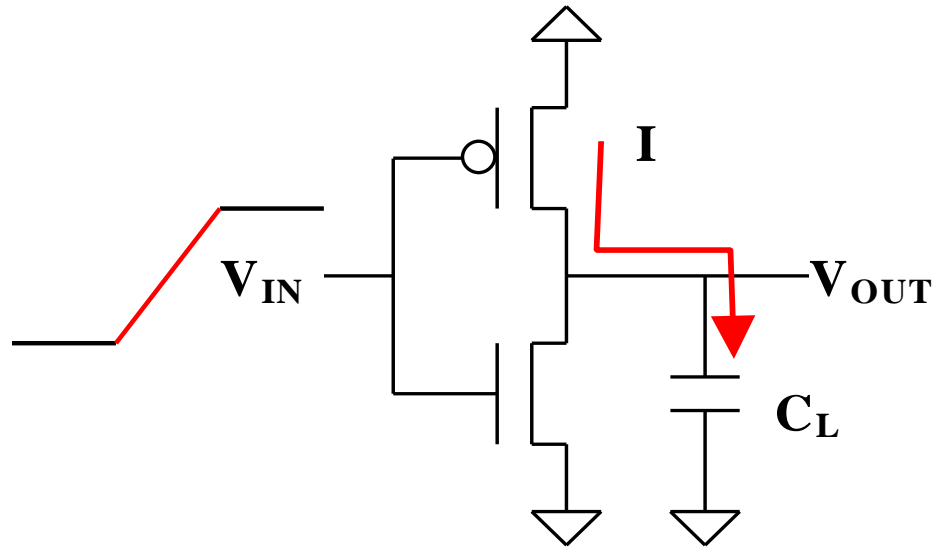
- **Data dependent – a function of switching activity**

Capacitive Power dissipation

Capacitance:
Function of wire
length, transistor size

Supply Voltage:
Has been dropping
with successive fab
generations

$$\text{Power} \sim \tfrac{1}{2}\, CV^2 Af$$

Activity factor:
How often, on average,
do wires switch?

Clock frequency:
Increasing…

Lowering Dynamic Power

- **Reducing Vdd has a quadratic effect**
    - **Has a negative (~linear) effect on performance however**

- **Lowering C_L**
    - **May improve performance as well**
    - **Keep transistors small (keeps intrinsic capacitance (gate and diffusion) small)**

- **Reduce switching activity**
    - **A function of signal transition stats and clock rate**
    - **Clock Gating idle units**
    - **Impacted by logic and architecture decisions**

Static Power: Leakage Currents
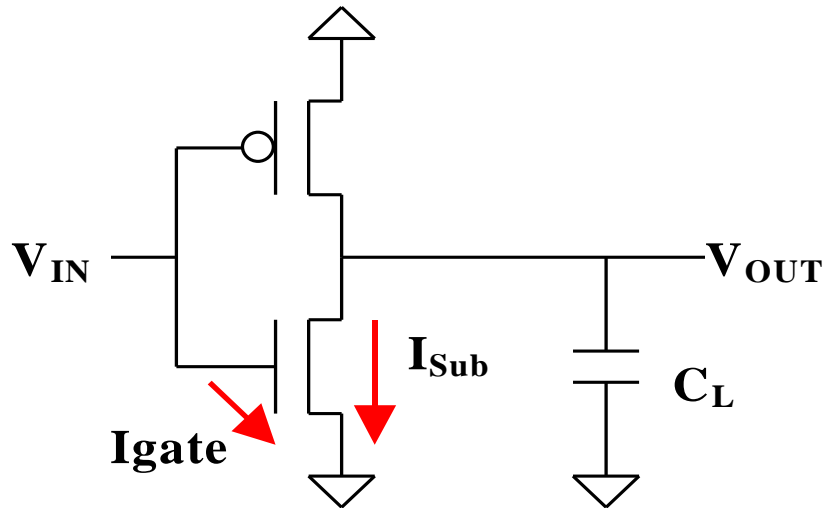
$$I_{DSub} = k \cdot e^{\frac{-q \cdot V_T}{a \cdot ka \cdot T}}$$

- **Subthreshold currents grow exponentially with increases in temperature, decreases in threshold voltage**
  - **But threshold voltage scaling is key to circuit performance!**
- **Gate leakage primarily dependent on gate oxide thickness, biases**
- **Both type of leakage heavily dependent on stacking and input pattern**

Lowering Static Power

- **Design-time Decisions**
  - **Use fewer, smaller transistors -- stack when possible to minimize contacts with Vdd/Gnd**
  - **Multithreshold process technology (multiple oxides too!)**
    - **Use "high-Vt" slow transistors whenever possible**
- **Dynamic Techniques**
  - **Reverse-Body Bias (dynamically adjust threshold)**
    - **Low-leakage sleep mode (maintain state), e.g. XScale**
  - **Vdd-gating (Cut voltage/gnd connection to circuits)**
    - **Zero-leakage sleep mode**
    - **Lose state, overheads to enable/disable**

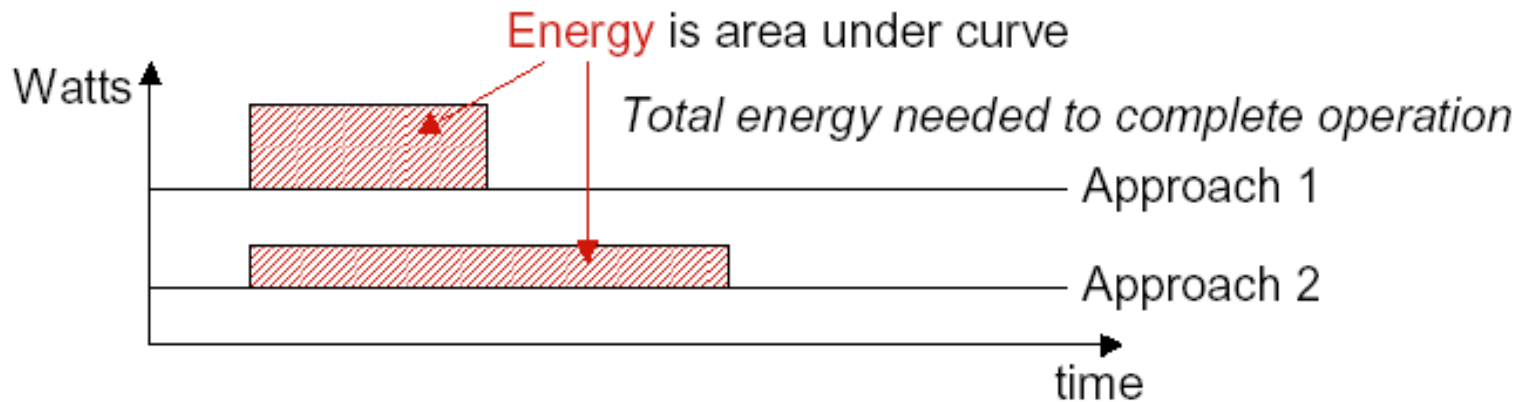# Power vs. Energy

- **Power** consumption in **Watts**
  - Determines battery life in hours
  - Sets packaging limits
- **Energy** efficiency in **joules**
  - Rate at which energy is consumed over time
  - Energy = power * delay (joules = watts * seconds)
  - Lower energy number means less power to perform a computation at same frequency

Power vs. Energy

Power is height of curve

Lower power design could simply be slower

Watts / time — Approach 1, Approach 2

Energy is area under curve

Total energy needed to complete operation

Watts / time — Approach 1, Approach 2
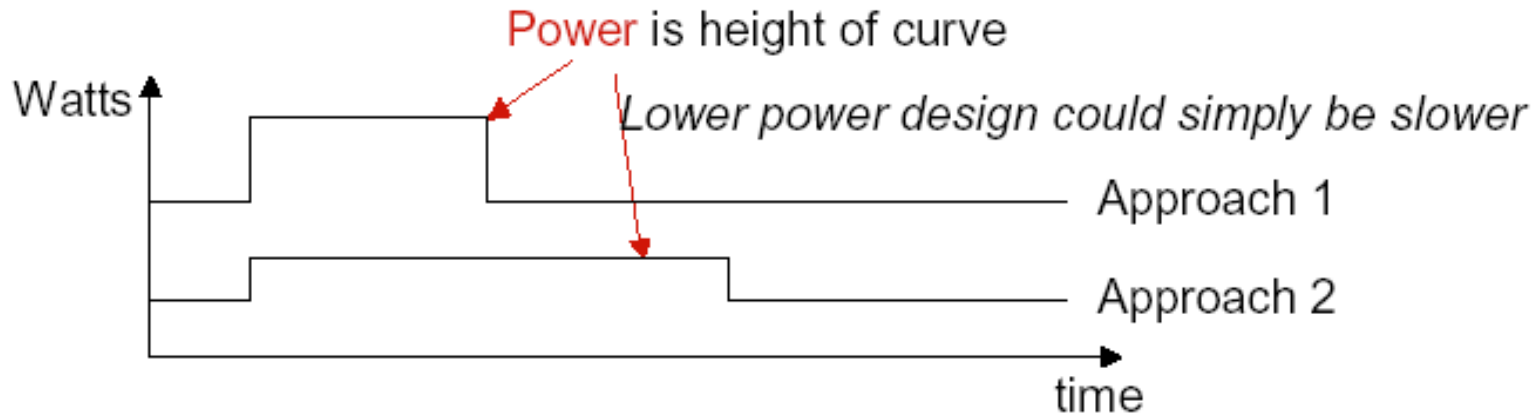
Power vs. Energy

- **Power-delay Product (PDP) = P_{avg} * t**
  - **PDP is the average energy consumed per switching event**

- **Energy-delay Product (EDP) = PDP * t**
  - **Takes into account that one can trade increased delay for lower energy/operation**

- **Energy-delay^2 Product (EDDP) = EDP * t**
  - **Why do we need so many formulas?!!?**
  - **We want a voltage-invariant efficiency metric! Why?**
  - **Power ~ ½ CV^2Af, Performance ~ f (and V)**

E vs. EDP vs. ED²P

- Power $\sim CV^2f \sim V^3$ (fixed microarch/design)

- Performance $\sim f \sim V$ (fixed microarch/design)

- (For the nominal voltage range, f varies linearly with V)

- Comparing processors that can only use freq/voltage scaling as the primary method of power control:

  - $(perf)^3$ / power, or $MIPS^3$ / W is a fair metric to compare energy efficiencies.

  - This is an $ED^2 P$ metric. We could also use: $(CPI)^3$ * W for a given application

E vs. EDP vs. ED²P

- **Currently have a processor design:**
  - **80W, 1 BIPS, 1.5V, 1GHz**
  - **Want to reduce power, willing to lose some performance**
  - **Cache Optimization:**
    - **IPC decreases by 10%, reduces power by 20% =>**
      Final Processor: 900 MIPS, 64W
    - **Relative E = MIPS/W (higher is better) = 14/12.5 = 1.125x**
  - **Energy is better, but is this a "better" processor?**

Not necessarily

- **80W, 1 BIPS, 1.5V, 1GHz**
  - **Cache Optimization:**
    - **IPC decreases by 10%, reduces power by 20% => Final Processor: 900 MIPS, 64W**
    - **Relative E = MIPS/W (higher is better) = 14/12.5 = 1.125x**
    - **Relative EDP = MIPS^2/W = 1.01x**
    - **Relative ED^{2}P = MIPS^3/W = .911x**
- **What if we just adjust frequency/voltage on processor?**
  - **How to reduce power by 20%?**
  - **P = CV^{2}F = CV^3 => Drop voltage by 7% (and also Freq) => .93*.93*.93 = .8x**
  - **So for equal power (64W)**
    - **Cache Optimization = 900MIPS**
    - **Simple Voltage/Frequency Scaling = 930MIPS**

What uses power in a chip?

What do we mean by Power?
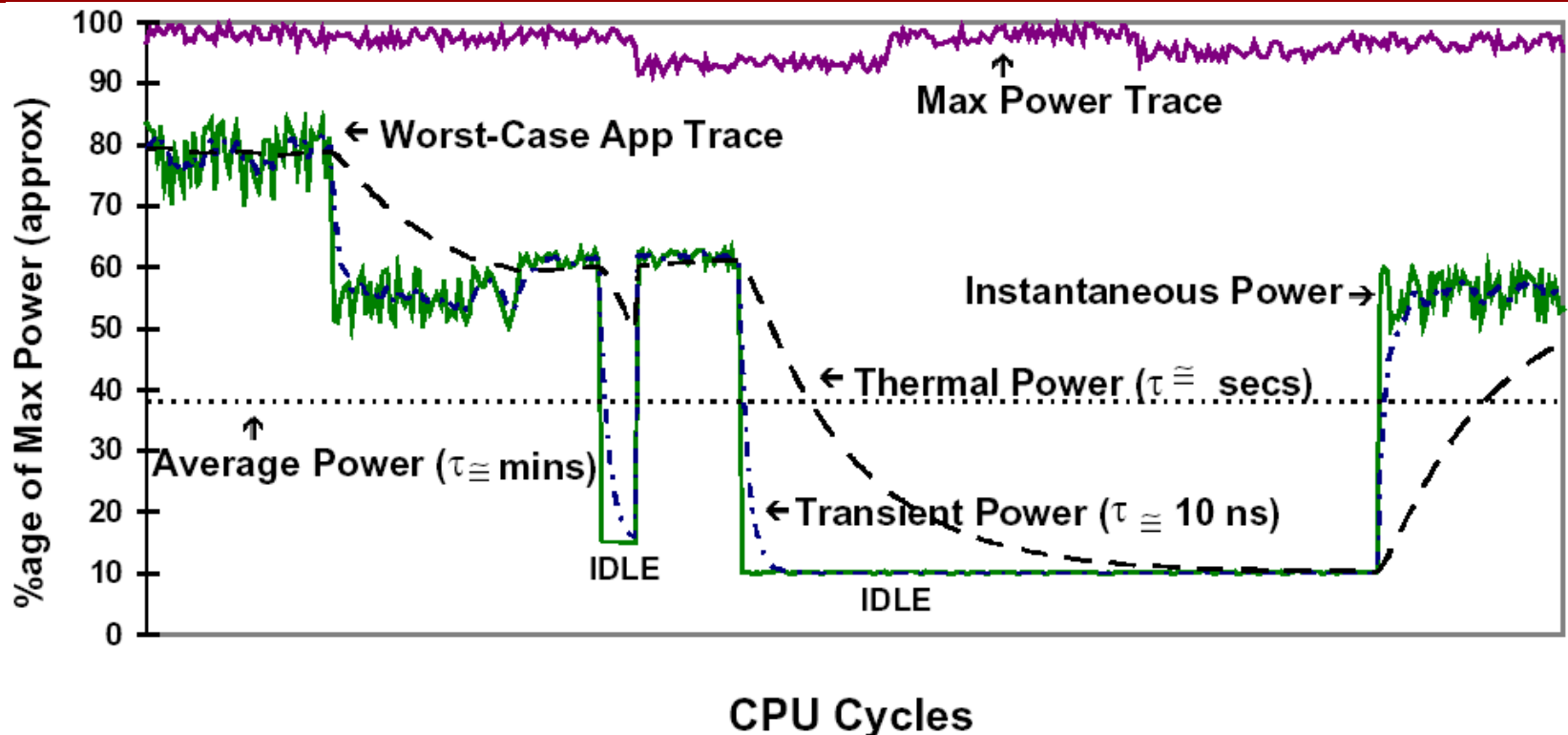
- **Max Power: Artificial code generating max CPU activity**
- **Worst-case App Trace:** *Practical* **applications worst-case**
- **Thermal Power: Running average of worst-case app power over a time period corresponding to thermal time constant**
- **Average Power: Long-term average of typical apps (minutes)**
- **Transient Power: Variability in power consumption for supply net**