Sad fact: Computers turn electrical energy into heat. Computation is a byproduct.

Energy and Performance

Air or water carries heat away, or chip melts.
This is how electric tea pots work...

Heats 1 gram of water
0.24 degree C

0.24 Calories per Second

1 Joule of Heat Energy per Second

1 Volt

1 Ohm Resistor

1 Ampere

1 Watt

20 W rating: Maximum power the package is able to transfer to the air. Exceed rating and resistor burns.
Cooling an iPod nano ...

Like resistor on last slide, iPod relies on passive transfer of heat from case to the air.

Why? Users don’t want fans in their pocket ...

To stay “cool to the touch” via passive cooling, power budget of 5 W.

If iPod nano used 5W all the time, its battery would last 15 minutes ...
Powering an iPod nano (2005 edition)

Battery has 1.2 W-hour rating: Can supply 1.2 W of power for 1 hour.

1.2 W / 5 W = 15 minutes.

More W-hours require bigger battery and thus bigger “form factor” -- it wouldn’t be “nano” anymore :-).

Real specs for iPod nano:
14 hours for music,
4 hours for slide shows.

85 mW for music.
300 mW for slides.
Finding the (2005) iPod nano CPU ...

A close relative ...

Two 80 MHz CPUs. One CPU used for audio, one for slides.

Low-power ARM roughly 1mW per MHz ... variable clock, sleep modes.

85 mW system power realistic ...
Year-to-year: continuous improvements

iPod nano 2005
14 hours battery life (audio playback)

iPod nano 2006
24 hours battery life (audio playback)

What changed inside?

Source: ifixit.com
iPod nano 2005 - a C-shaped PC board, with a battery in the “C” opening.

iPod nano 2006 - battery lies on top of PC board.

Source: ifixit.com
How? Small IC packages, fewer parts

Source: arstechnica.com
Aluminum permits thinner case ...
"Bill of Materials (BOM)" - Cost of all parts

Distinction: Costs that are borne for each unit sold (like BOM) vs once-per-new-product costs (like R&D).
iPod shuffle

- 0.55 ounces
- 12 hour battery life
- $49.00
- 1 GB
Notebooks ... now most of the PC market.

Performance: Must be “close enough” to desktop performance ... many people no longer own a desktop (including me!)

Size and Weight. Ideal: paper notebook.

Heat: No longer “laptops” -- top may get “warm”, bottom “hot”. Quiet fans OK.
Battery: Set by size and weight limits ...

Battery rating: 55 W-hour.

At 2.3 GHz, Intel Core Duo CPU consumes 31 W running a heavy load - under 2 hours battery life! And, just for CPU!

46x more energy than iPod nano battery. And iPod lets you listen to music for 14 hours!

Almost full 1 inch depth. Width and height set by available space, weight.

At 1 GHz, CPU consumes 13 Watts. "Energy saver" option uses this mode ...

46x more energy than iPod nano battery. And iPod lets you listen to music for 14 hours!
But CPU is only part of power budget!

2004-era notebook running a full workload.

“Other”
CPU
LCD
Backlight
LCD

“Amdahl’s Law for Power”

If our CPU took no power at all to run, that would only double battery life!

Data courtesy Mahesri et al., U of Illinois, 2004
55 W-hour battery stores the energy of 1/2 a stick of dynamite.

If battery short-circuits, catastrophe is possible ...
Machine rooms are expensive. Removing heat dictates how many servers to put in a machine room.

Electric bill adds up! Powering the servers + powering the air conditioners is a big part of TCO.

Reliability: running computers hot makes them fail more often.
Processors and Energy
Switching Energy: Fundamental Physics

Every logic transition dissipates energy.

\[
E_{0 \rightarrow 1} = \frac{1}{2} C V_{dd}^2 \quad E_{1 \rightarrow 0} = \frac{1}{2} C V_{dd}^2
\]

How can we limit switching energy?

1. Slow down clock (fewer transitions). But we like speed ...
2. Reduce Vdd. But lowering Vdd lowers the clock speed ...
3. Fewer circuits. But more transistors can do more work.
4. Reduce C per node. One reason why we scale processes.
Scaling switching energy per gate ...

Recall process scaling ("Moore’s Law")

Due to reducing V and C (length and width of Cs decrease, but plate distance gets smaller).
Recent slope more shallow because V is being scaled less aggressively.

Second Factor: Leakage Currents

Even when a logic gate isn’t switching, it burns power.

\[ 0V = V_{IN} \]

\[ I_{Gate} \]

\[ I_{Sub} \]

\[ V_{OUT} \]

\[ C_L \]

**I\text{sub}:** Even when this nFet is off, it passes an \text{Ioff} leakage current.

We can engineer any \text{Ioff} we like, but a lower \text{Ioff} also results in a lower \text{Ion}, and thus the lower the clock speed.

**I\text{gate}:** Ideal capacitors have zero DC current. But modern transistor gates are a few atoms thick, and are not ideal.

---

Intel’s 2006 processor designs, leakage vs switching power

Bill Holt, Intel, Hot Chips 17.

A lot of work was done to get a ratio this good ... 50/50 is common.
Engineering “On” Current at 25 nm ...

We can increase $I_{on}$ by raising $V_{dd}$ and/or lowering $V_t$.

$I_{ds} = 1.2 \text{ mA} = I_{on}$

$0.25 = V_t$

$0.7 = V_{dd}$

$I_{off} = 0$ ???
We can decrease $I_{\text{off}}$ by raising $V_t$ - but that lowers $I_{\text{on}}$.

$I_{\text{off}} \approx 10 \text{ nA}$

0.25 $\approx V_t$

1.2 mA $= I_{\text{on}}$

0.7 $= V_{dd}$
Device engineers trade speed and power

We can reduce $CV^2$ ($P_{\text{active}}$) by lowering $V_{dd}$.

We can increase speed by raising $V_{dd}$ and lowering $V_{t}$.

We can reduce leakage ($P_{\text{standby}}$) by raising $V_{t}$.

---

From: Silicon Device Scaling to the Sub-10-nm Regime
Meikei Ieong, Bruce Doris, Jakub Kedzierski, Ken Rim, Min Yang
Customize processes for product types ...

Effective Process and Design Collaboration Succeeds in Power Improvements

Next generation product if 130nm process only

Process, design efficiency, and frequency increase

Find enough tricks, and you can afford to raise Vdd a little so that you can raise the clock speed!

Clock speed unchanged ...

Lower Vdd, lower C, but more leakage.

Design tricks: architecture & circuits.

Relative Power

130nm
Next generation product if 130nm
90nm process only
Process plus design efficiency
Process, design efficiency, and frequency increase

Intel: Comparing 2 CPU generations ...
Trading Hardware for Power
Gate delay roughly linear with Vdd

Block processes stereo audio. 1/2 of clocks for “left”, 1/2 for “right”.

Into this:

CV^2 power only

Ex: Top block processes audio channel 1, bottom block processes audio channel 2.
Multiple Cores for Low Power

Trade hardware for power, on a large scale ...
Cell:
The PS3 chip
Cell: Conventional CPU + 8 "SPUs"

- L2 Cache
  - 512 KB
- 8 Synergistic Processing Units (SPUs)
- PowerPC
One Synergistic Processing Unit (SPU)

256 KB Local Store -- 128 128-bit Registers
SPU issues 2 inst/cycle (in order) to 7 execution units
SPU fills Local Store using DMA to DRAM and network
A “Schmoo” plot for a Cell SPU ...

The lower Vdd, the less dynamic energy consumption.

\[ E_{0 \rightarrow 1} = \frac{1}{2} CV_{dd}^2 \quad E_{1 \rightarrow 0} = \frac{1}{2} CV_{dd}^2 \]

The lower Vdd, the longer the maximum clock period, the slower the clock frequency.

\[ V_{dd} \]

\[ E_{0 \rightarrow 1} \]

\[ E_{1 \rightarrow 0} \]
Clock speed alone doesn’t help E/op ...

But, lowering clock frequency while keeping voltage constant spreads the same amount of work over a longer time, so chip stays cooler ...

\[ E_{0 \rightarrow 1} = \frac{1}{2} C V_{dd}^2 \quad E_{1 \rightarrow 0} = \frac{1}{2} C V_{dd}^2 \]
Scaling V and f does lower energy/op

1 W to get 2.2 GHz performance. 26 C die temp.

7 W to reliably get 4.4 GHz performance. 47 C die temp.

If a program that needs a 4.4 GHz CPU can be recoded to use two 2.2 GHz CPUs ... big win.
Intel’s dual-core analysis ...

But only if your app(s) can put 2 cores to use!

In the same process technology...

Voltage = 1  Voltage = -15%
Freq     = 1  Freq     = -15%
Area     = 1  Area     = 2
Power    = 1  Power    = 1
Perf     = 1  Perf     = ~1.8

How iPod puts its 2 cores to use ...

Two 80 MHz CPUs. This chip is used in the most iPods now, with one CPU doing audio decoding, the other doing photos, etc.
Other Low-Power Techniques
Add “sleep” transistors to logic ...

Example: Floating point unit logic.

When running fixed-point instructions, put logic “to sleep”.

+++ When “asleep”, leakage power is dramatically reduced.

--- Presence of sleep transistors slows down the clock rate when the logic block is in use.
Intel example: Sleeping cache blocks

70 Mbit SRAM leakage current map

>3x SRAM leakage reduction on inactive blocks

Recall: Most logic on a chip “too fast”

The critical path

Most wires have hundreds of picoseconds to spare.

Use several supply voltages on a chip ...

Why? We could use the low-power Vdd for logic well off the critical path.

On a CPU, where does the heat go?

Half of the power goes to latches (Flip-Flops).

Most of the time, the latches don’t change state.

So (gasp) gated clocks are a big win.
But, done with CAD tools in a disciplined way.

From: Bose, Martonosi, Brooks: Sigmetrics-2001 Tutorial
IBM Power 4: How does die heat up?

4 dies on a multi-chip module

2 CPUs per die
115 Watts: Concentrated in “hot spots”

66.8°C == 152°F  82°C == 179.6°F
Device physics ...

NAND Flash Memory
The physics of non-volatile memory

1. Electrons “placed” on floating gate stay there for many years (ideally).
2. 10,000 electrons on floating gate shift transistor threshold by 2V.
3. In a memory array, shifted transistors hold “0”, unshifted hold “1”.

Two gates. But the middle one is not connected.
1. Hot electron injection and tunneling produce tiny currents, thus writes are slow.

2. High voltages damage the floating gate. Too many writes and a bit goes “bad”.

A high drain voltage injects “hot electrons” onto floating gate.

A high gate voltage “tunnels” electrons off of floating gate.
Architecture ...

NAND Flash Memory
Flash: Disk Replacement

Chip “remembers” for 10 years.

Presents memory to the CPU as a set of **pages**.

**Page format:**

<table>
<thead>
<tr>
<th>2048 Bytes</th>
<th>+</th>
<th>64 Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(user data)</td>
<td></td>
<td>(meta data)</td>
</tr>
</tbody>
</table>

1GB Flash: 512K pages
2GB Flash: 1M pages
4GB Flash: 2M pages
Reading a Page ...

33 MB/s Read Bandwidth

Bus Control

Flash Memory

8-bit data or address (bi-directional)

Samsung K9WAG08U1A

Page address in: 175 ns

Clock out page bytes: 52,800 ns

First byte out: 10,000 ns

UC Regents Fall 2008 © UCB
First byte out: 10,000 ns
Clock out page bytes: 52,800 ns

Where Time Goes

Figure 1. K9K8G08U0A Functional Block Diagram

Page address in: 175 ns

CS 194-6 L11: The Power Wall + Flash
Writing a Page ...

A page lives in a block of 64 pages:
- 1GB Flash: 8K blocks
- 2GB Flash: 16K blocks
- 4GB Flash: 32K blocks

To write a page:

1. Erase all pages in the block (cannot erase just one page).
   Time: 1,500,000 ns

2. May program each page individually, exactly once.
   Time: 200,000 ns per page.

Block lifetime: 100,000 erase/program cycles.
Block Failure

Even when new, not all blocks work!

1GB: 8K blocks, 160 may be bad.
2GB: 16K blocks, 220 may be bad.
4GB: 32K blocks, 640 may be bad.

During factory testing, Samsung writes good/bad info for each block in the meta data bytes.

<table>
<thead>
<tr>
<th>Block 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page 0</td>
</tr>
<tr>
<td>Page 1</td>
</tr>
<tr>
<td>:</td>
</tr>
<tr>
<td>Page 62</td>
</tr>
<tr>
<td>Page 63</td>
</tr>
</tbody>
</table>

2048 Bytes + 64 Bytes

(user data) (meta data)

After an erase/program, chip can say “write failed”, and block is now “bad”. OS must recover (migrate bad block data to a new block). Bits can also go bad “silently” (!!!).
Flash controllers: Chips or Verilog IP ...

Flash memory controller manages write lifetime management, block failures, silent bit errors ...

Software sees a “perfect” disk-like storage device.
Recall: iPod 2005 ...

Flash memory

Flash controller
This Friday:

<table>
<thead>
<tr>
<th>F</th>
<th>Student Presentations: Final Project Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/5</td>
<td>10 AM to Noon, 125 Cory</td>
</tr>
</tbody>
</table>

No Lecture on Dec 8!