CS 250
VLSI System Design

Lecture 6 – Power and Energy

2011-9-19

John Wawrzynek and Krste Asanović

with John Lazzaro

TA: Brian Zimmer

www-inst.eecs.berkeley.edu/~cs250/

UC Regents Fall 2011 © UCB

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

Energy and Performance

Air or water carries heat away, or chip melts.

UC Regents Fall 2011 © UCB

The Joule: Unit of energy. Can also be expressed as **Watt-Seconds.** Burning 1 Watt for 100 seconds uses 100 Watt-Seconds of energy.

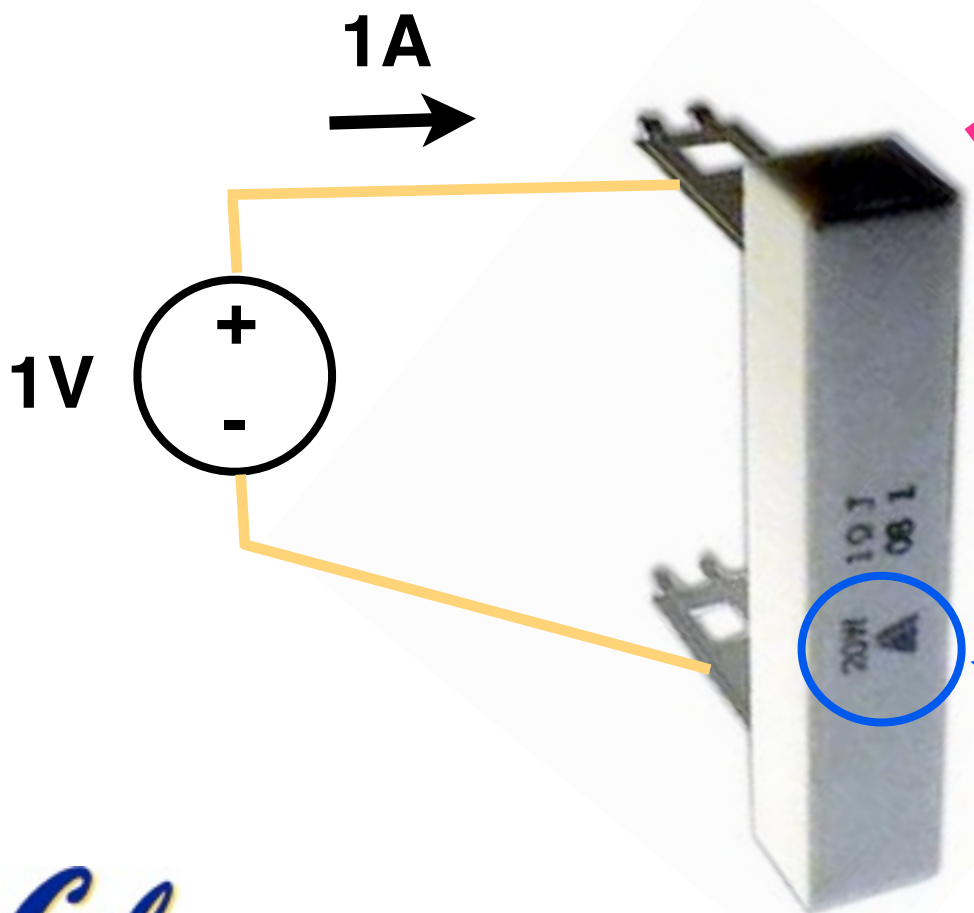
This is how electric tea pots work ...

1 Joule heats 1 gram of water 0.24 degree C

1 Joule of Heat Energy per Second

1 Watt

The Watt: Unit of power. The amount of energy burned in the resistor in 1 second.



20 W rating: Maximum power the package is able to transfer to the air. Exceed rating and resistor **burns.**

UC Regents Fall 2011 © UCB

Cooling an iPod nano ...

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

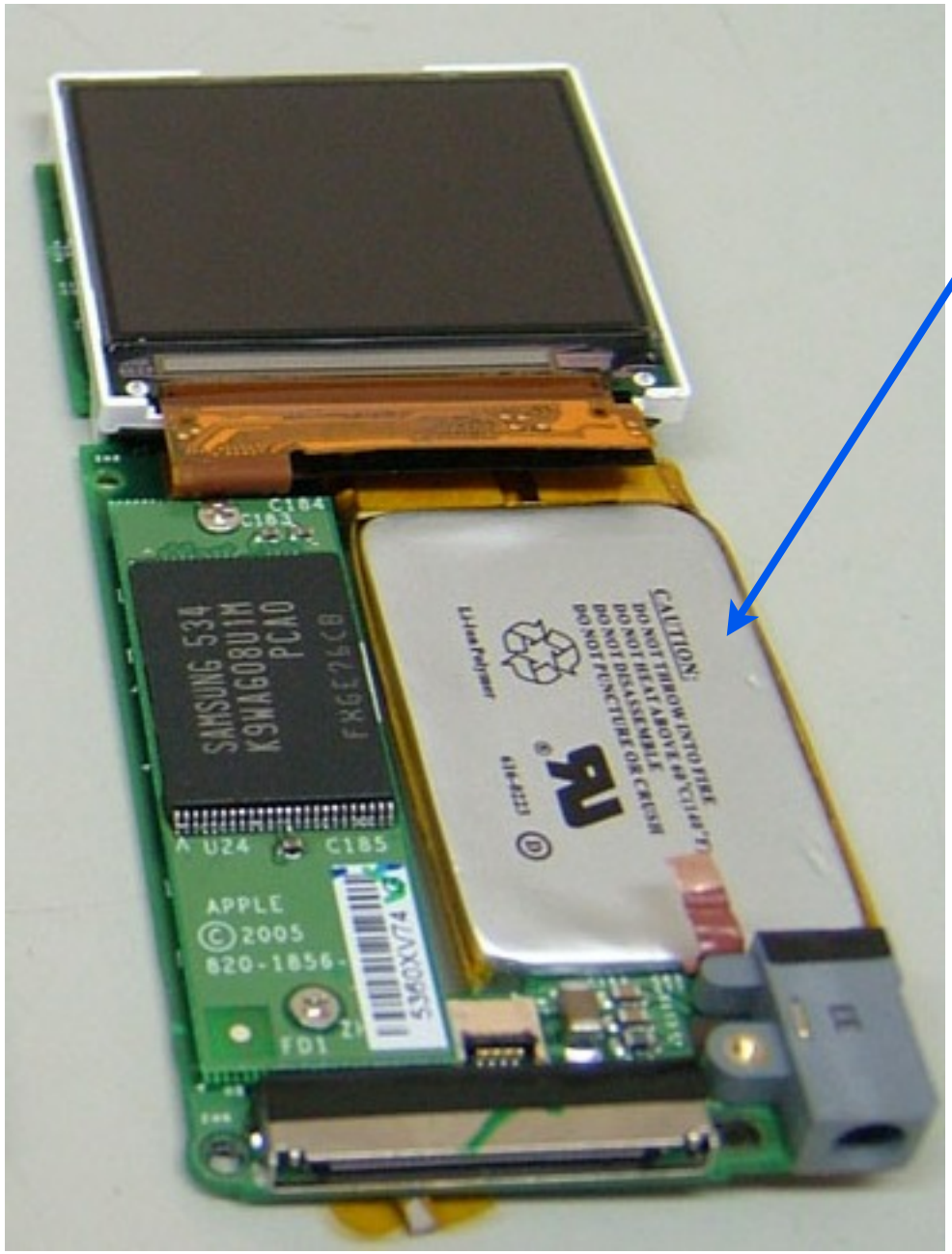
Why? Users don't want fans in their pocket ...

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 ...

UC Regents Fall 2011 © UCB

Powering an iPod nano (2005 edition)



1.2 W-hour battery:
Can supply 1.2 watts
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.

UC Regents Fall 2011 © UCB

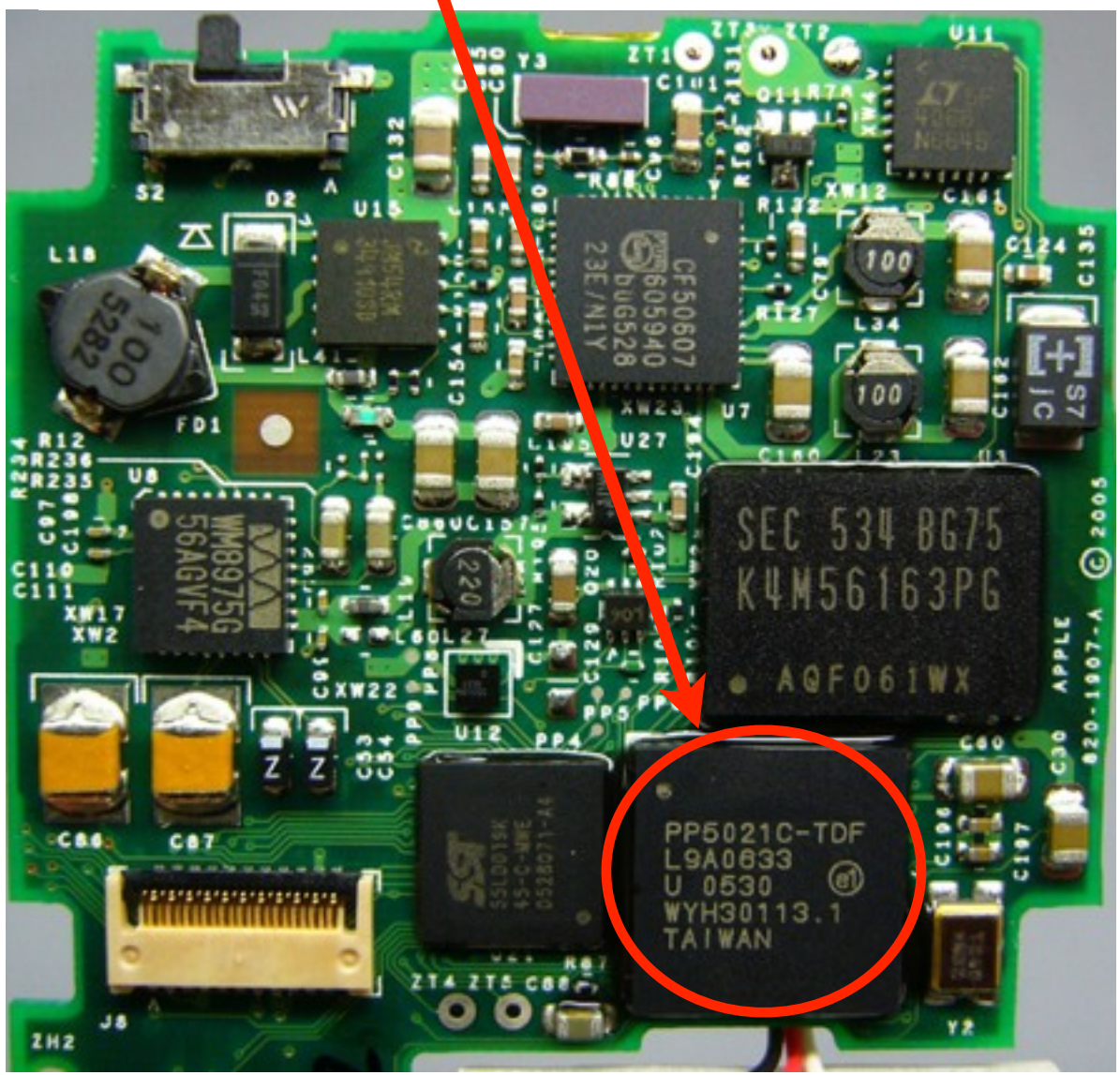
Finding the (2005) iPod nano CPU ...

portalplayer™

A close relative ...

PP5020

digital media management system-on-chip



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 ...

UC Regents Fall 2011 © UCB

Year-to-year: continuous improvements

iPod nano
2006
24 hours
battery
life
(audio
playback)

iPod nano
2005
14 hours
battery
life
(audio
playback)

What
changed
inside ?

Source: ifixit.com

Source: ifixit.com

Source: ifixit.com

iPod nano 2005 - a C-shaped PC board, with a battery in the "C" opening.

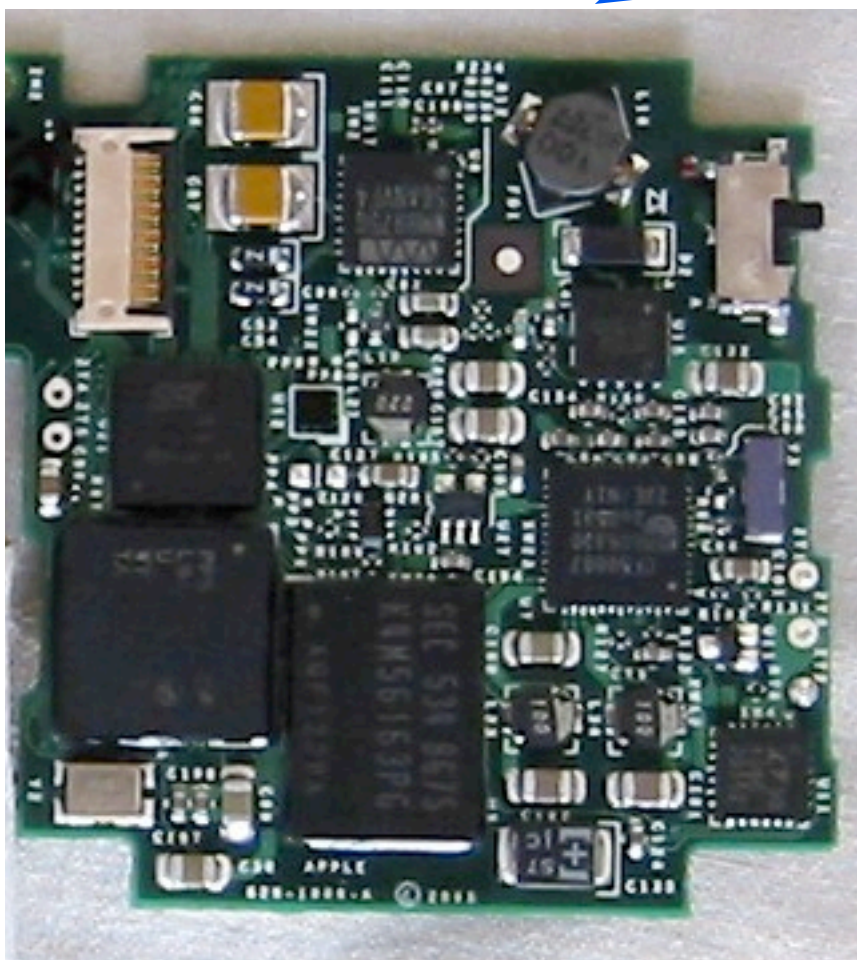
iPod nano 2006 - battery lies on top of PC board.

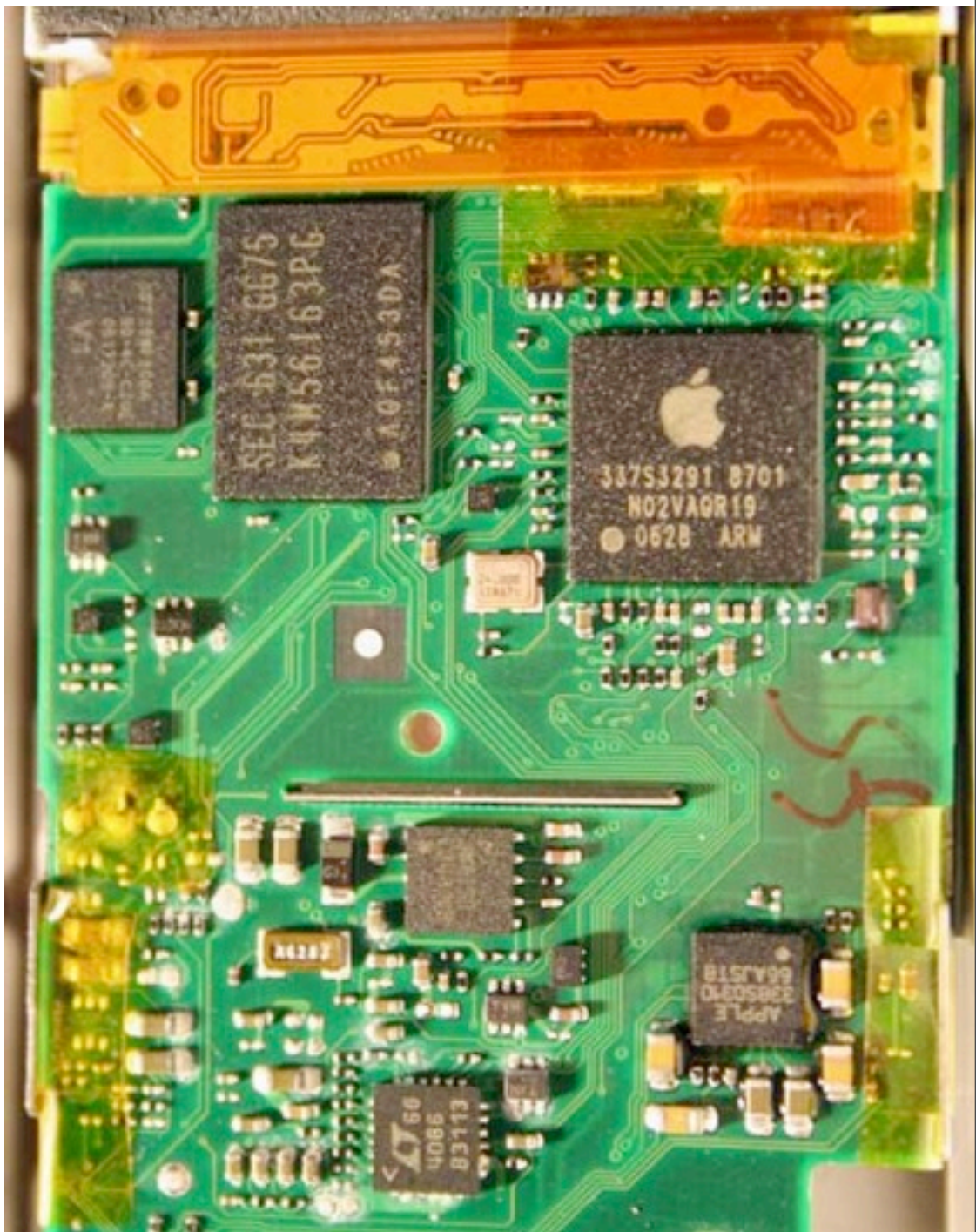
Source: ifixit.com

How? Small IC packages, fewer parts

iPod nano 2006 →

iPod nano 2005 →





Source: arstechnica.com

Aluminum permits thinner case ...

Source: ilounge.com

2010 Nano

2010 Shuffle

nearly the same depth

0.74 ounces

0.44 ounces

2010 Nano:
"up to" 24 hours
audio playback

2010 Shuffle:
"up to" 15 hours
audio playback

0.39 W Hr
(33% of 2005 Nano)

0.19 W Hr

Sources: iFixit, Apple

UC Regents Fall 2011 © UCB

Notebooks ... as designed in 2006 ...

2006 Apple MacBook -- 5.2 lbs

- **Performance:** Must be "close enough" to desktop performance ... most people no longer used a desktop (even in 2006).
- **Size and Weight.** Ideal: paper notebook.
- **Heat:** No longer "laptops" -- top may get "warm", bottom "hot". Quiet fans OK.

UC Regents Fall 2011 © UCB

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!

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 ...

UC Regents Fall 2011 © UCB

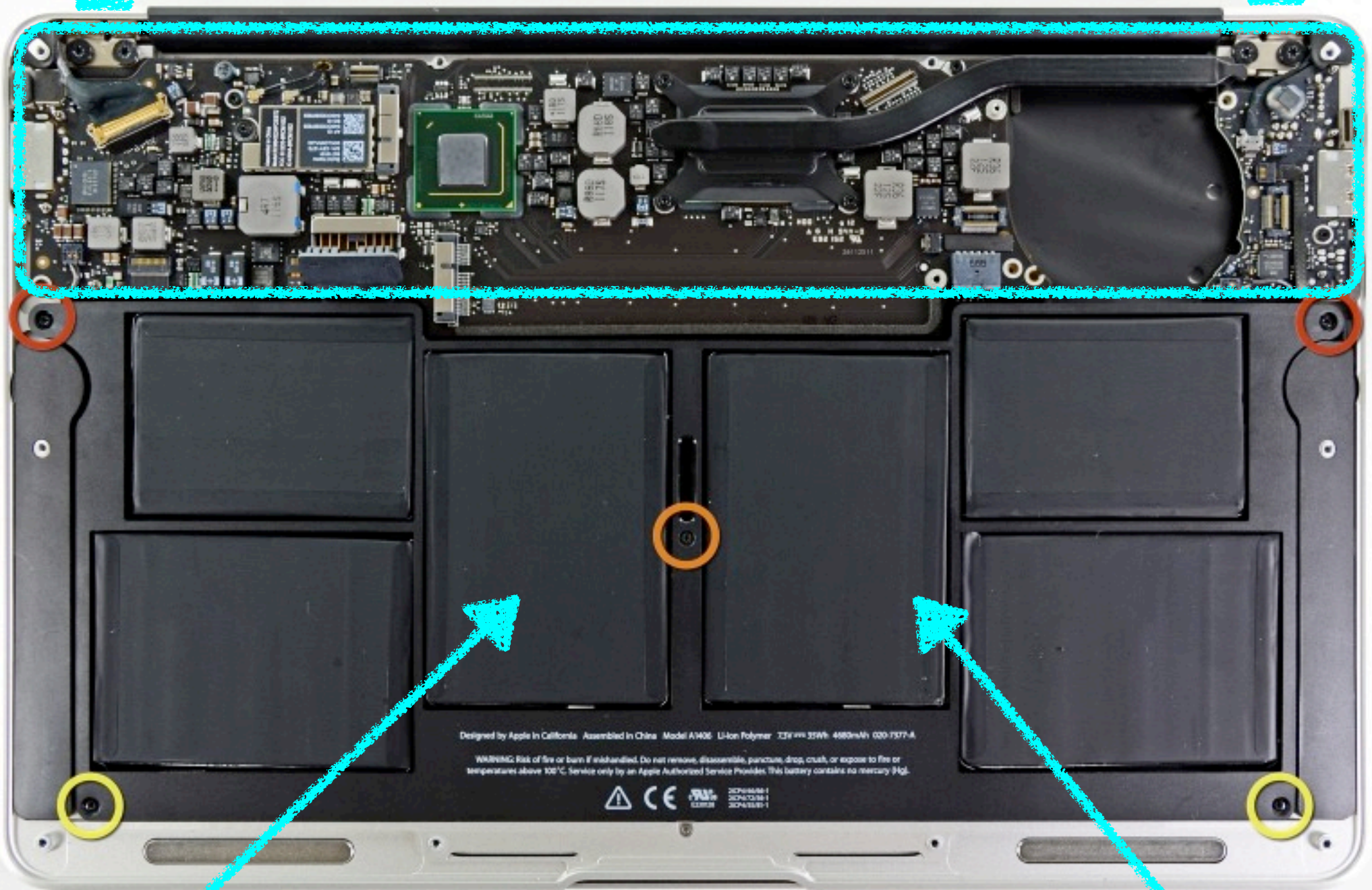
MacBook Air … design the laptop like an iPod

2011 Air: 11.8 in x 7.56 in x 0.68 in; 2.38 lbs

2006 Macbook: 12.8 in x 8.9 in x 1 in; 5.2 lbs

UC Regents Fall 2011 © UCB

Mainboard: fills about 25% of the laptop



35 W-h battery: 63% of 2006 MacBook's 55 W-h

Condé Nast: The Evil Empire
703
$2.99
iPad

2011 Air: 35 W-h battery, 5 hour battery life*

iPad 2: 25 W-h battery, 10 hour battery life*

*For a content-consumption workload.

Battery-Life-Hour/W-h: 2.8x iPad advantage

UC Regents Fall 2011 © UCB

"Content Creation vs. Content Consumption"

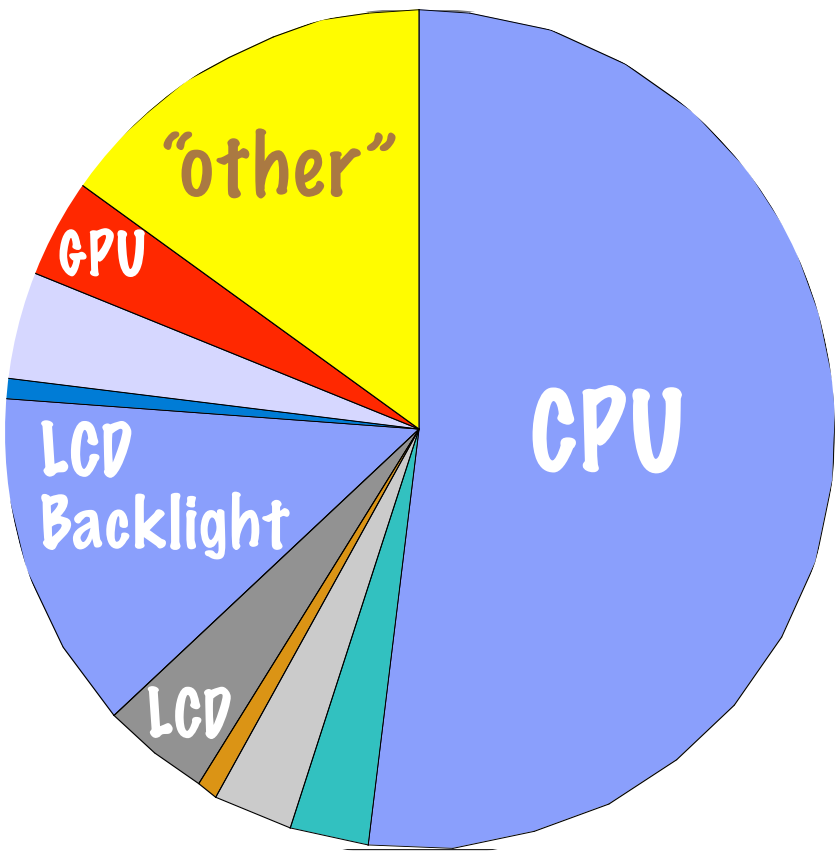
2011 Air: $999 -- 64 GB SSD, 2 GB RAM, x86

iPad 2: $699 -- 64 GB SSD, 512 MB RAM, ARM

UC Regents Fall 2011 © UCB

The CPU is only part of power budget!

2004-era notebook running a full workload.



"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

UC Regents Fall 2011 © UCB

Servers: Total Cost of Ownership (TCO)

Machine rooms are expensive. Removing heat dictates how many servers to put in a machine room.

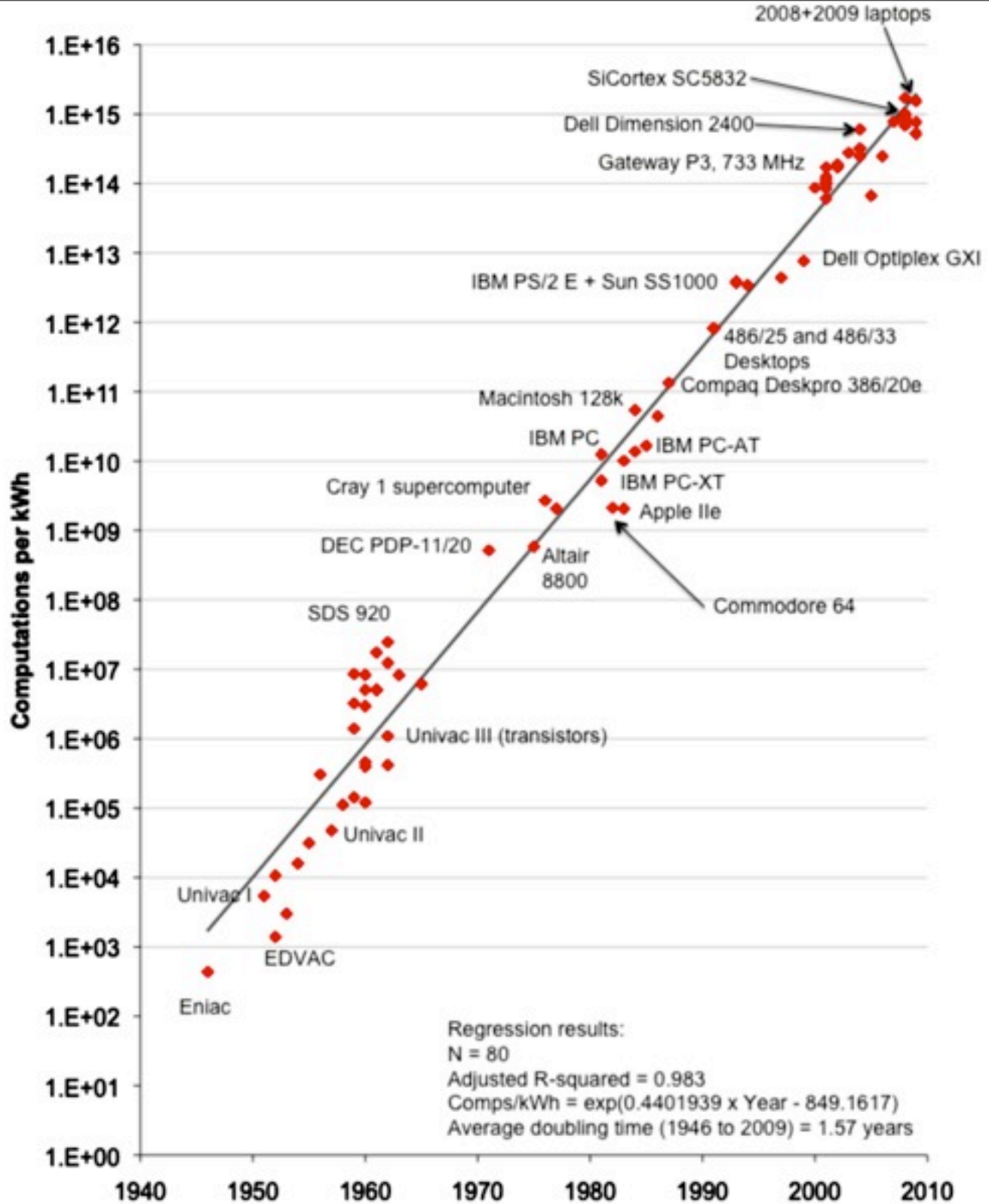
Electric bill adds up! Powering the servers + powering the air conditioners is a big part of TCO.

Reliability: running computers hot makes them fail more often.

UC Regents Fall 2011 © UCB

Computations per W-h doubles every 1.6 years, going back to the first computer.

(Jonathan Koomey, Stanford).

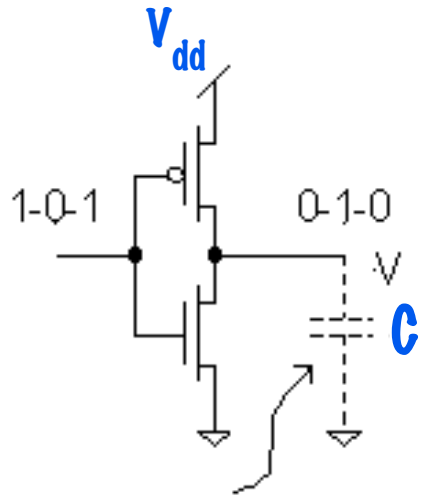


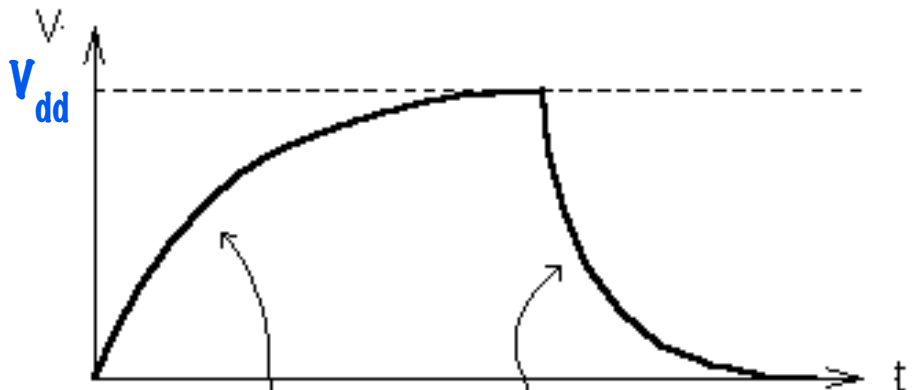
Processors and Energy

UC Regents Fall 2011 © UCB

Switching Energy: Fundamental Physics

Every logic transition dissipates energy.





$$E_{0\to1} = \frac{1}{2} C V_{dd}^2 \qquad E_{1\to0} = \frac{1}{2} C V_{dd}^2$$

Strong result: Independent of technology.

How can we limit switching energy?

(1) Reduce # of clock transitions. But we have work to do ...

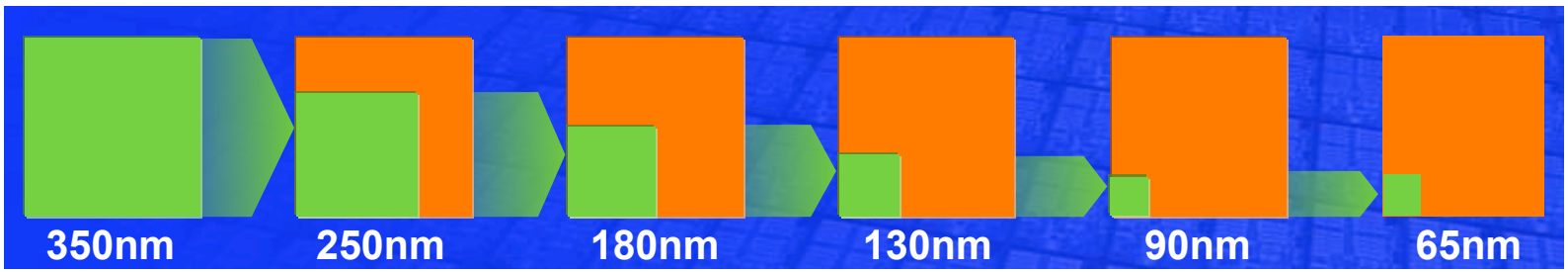
(2) Reduce Vdd. But lowering Vdd limits the clock speed ...

(3) Fewer circuits. But more transistors can do more work.

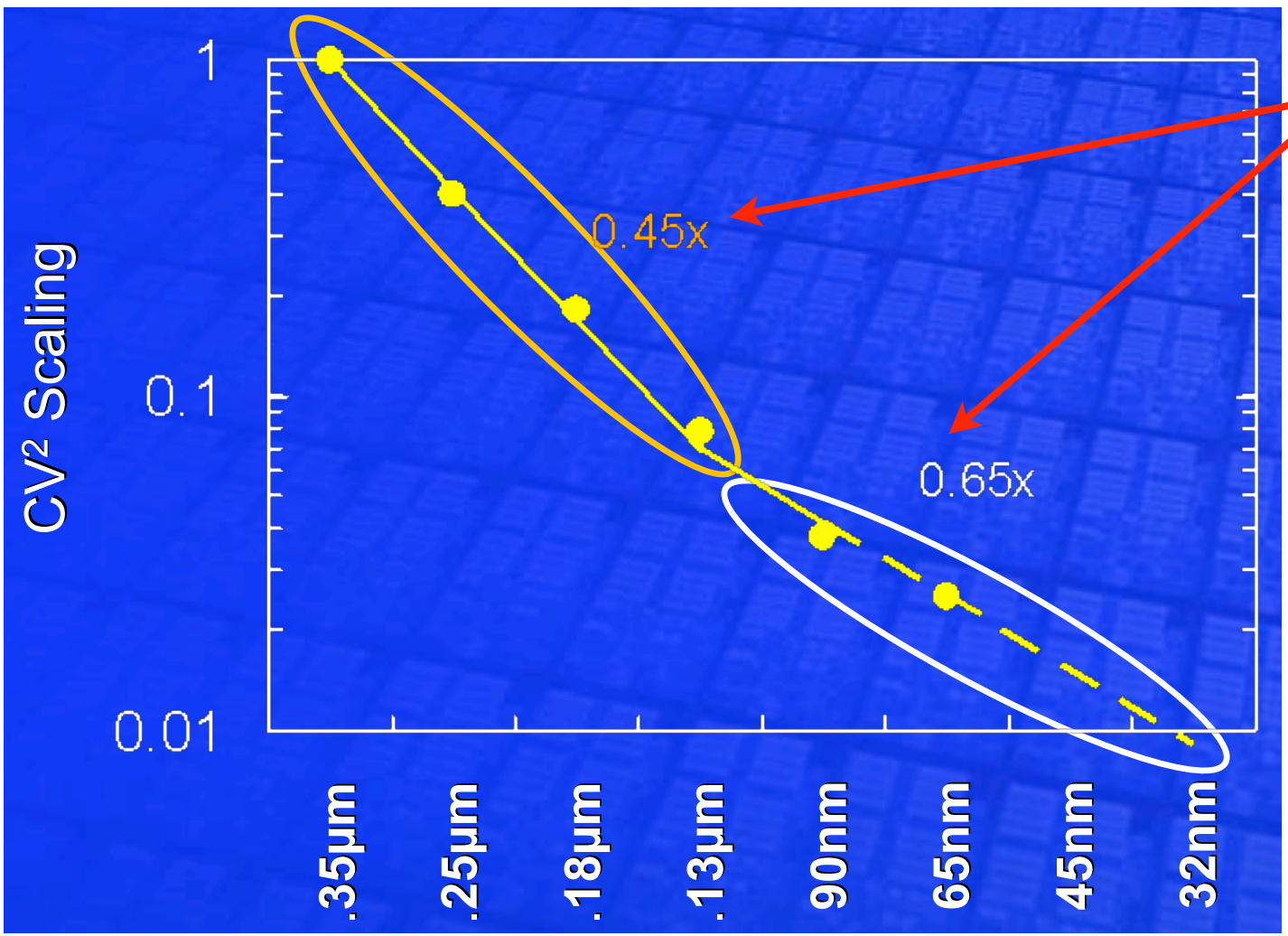
(4) Reduce C per node. One reason why we scale processes.

UC Regents Fall 2011 © UCB

Scaling switching energy per gate ...



IC process scaling ("Moore's Law")



Due to reducing V and C (length and width of Cs decrease, but plate distance gets smaller).

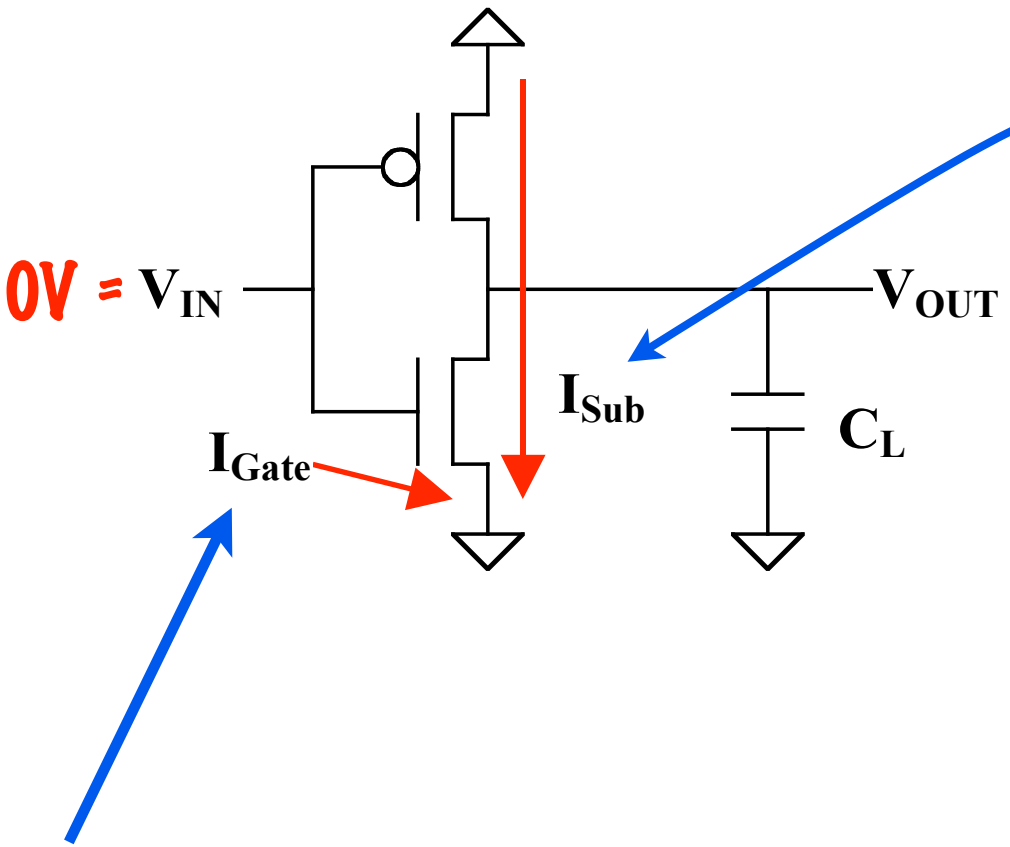
Recent slope more shallow because V is being scaled less aggressively.

From: "Facing the Hot Chips Challenge Again", Bill Holt, Intel, presented at Hot Chips 17, 2005.

UC Regents Fall 2011 © UCB

Second Factor: Leakage Currents

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

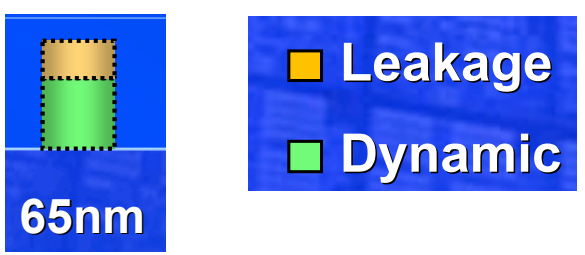


Isub: Even when this nFet is off, it passes an Ioff leakage current.

We can engineer any Ioff we like, but a lower Ioff also results in a lower Ion, and thus a lower maximum clock speed.

Igate: 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

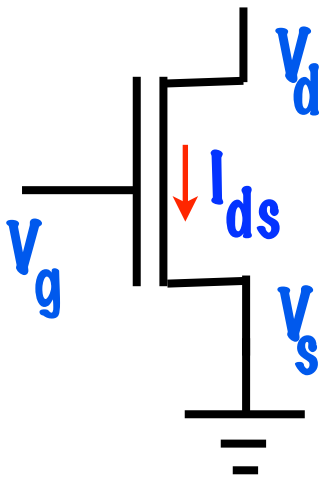


A lot of work was done to get a ratio this good ... 50/50 is common.

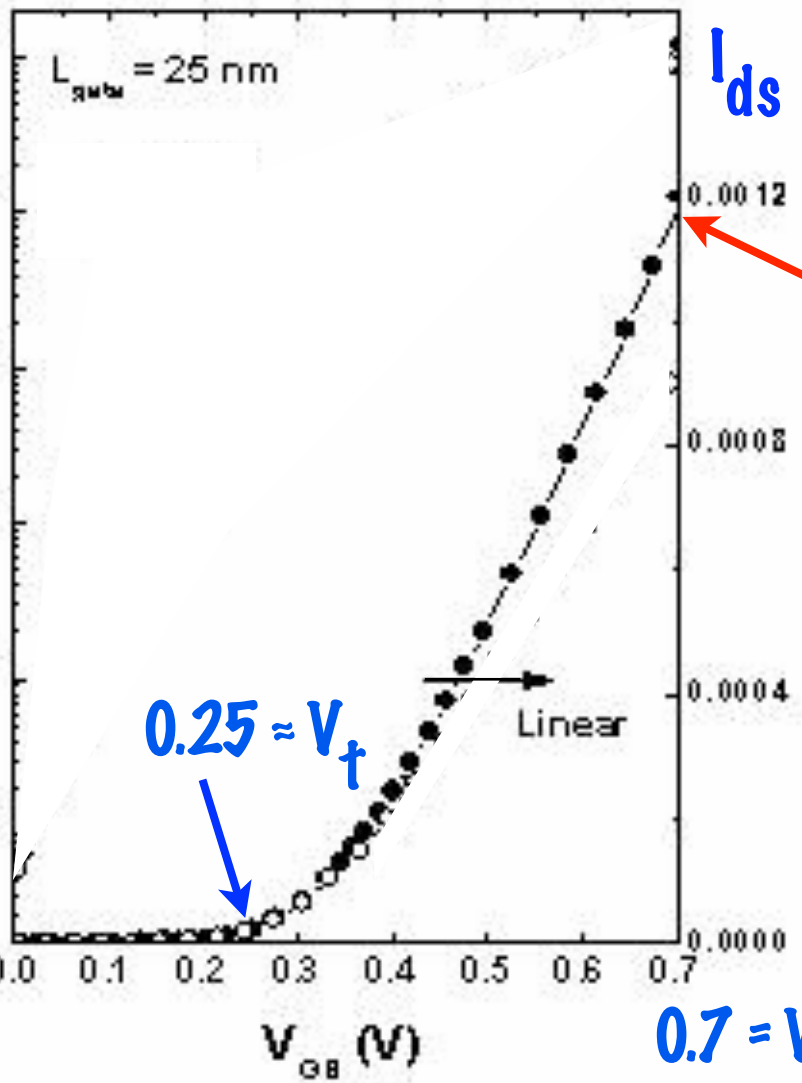
Bill Holt, Intel, Hot Chips 17.

UC Regents Fall 2011 © UCB

Engineering "On" Current at 25 nm …



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



1.2 mA = I_{on}

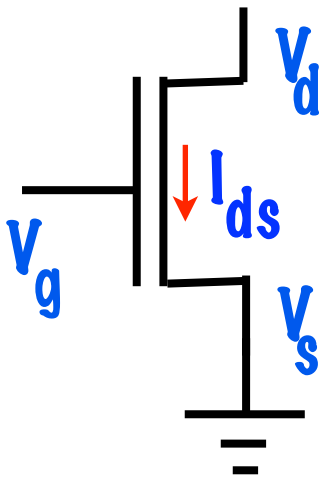
0.25 ≈ V_t

I_{off} = 0 ???

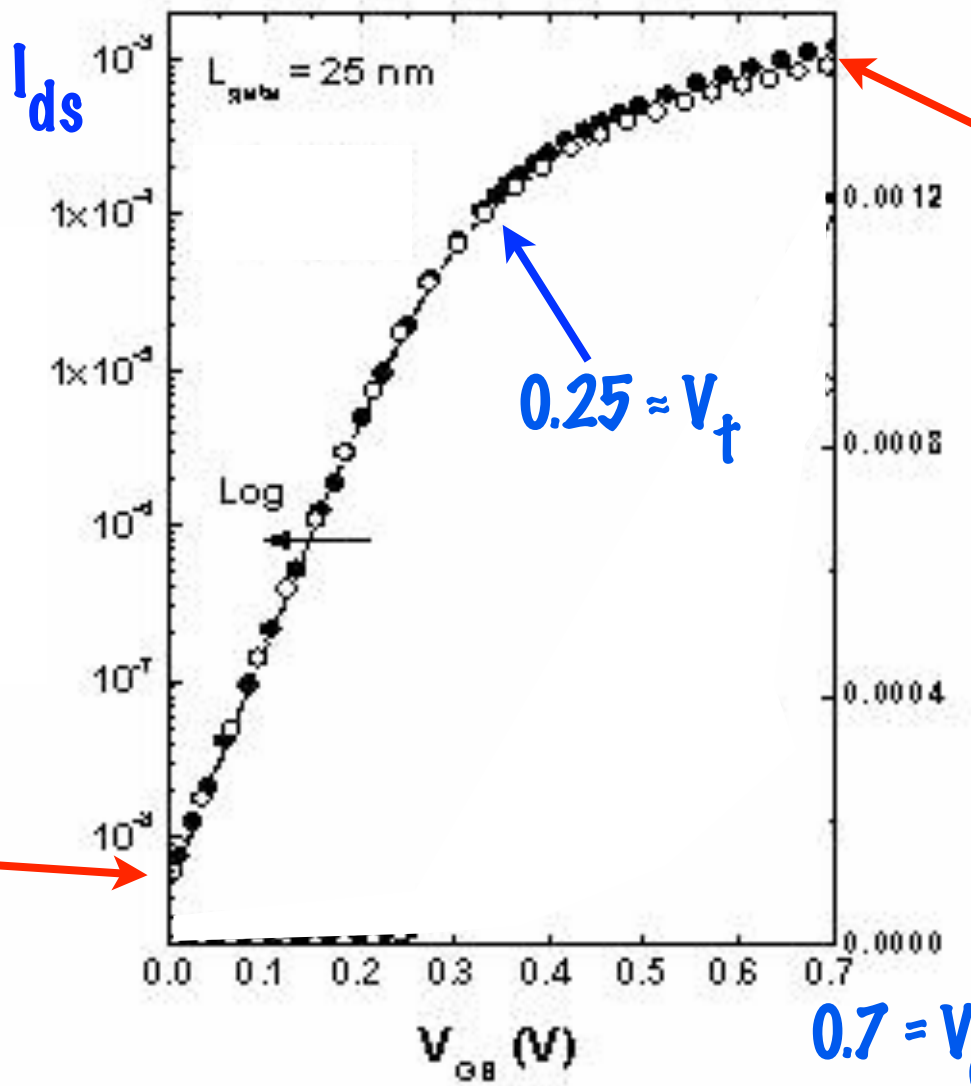
0.7 = V_{dd}

UC Regents Fall 2011 © UCB

Plot on a "Log" Scale to See "Off" Current



We can decrease I_{off} by raising V_t - but that lowers I_{on}.



I_{off} ≈ 10 nA

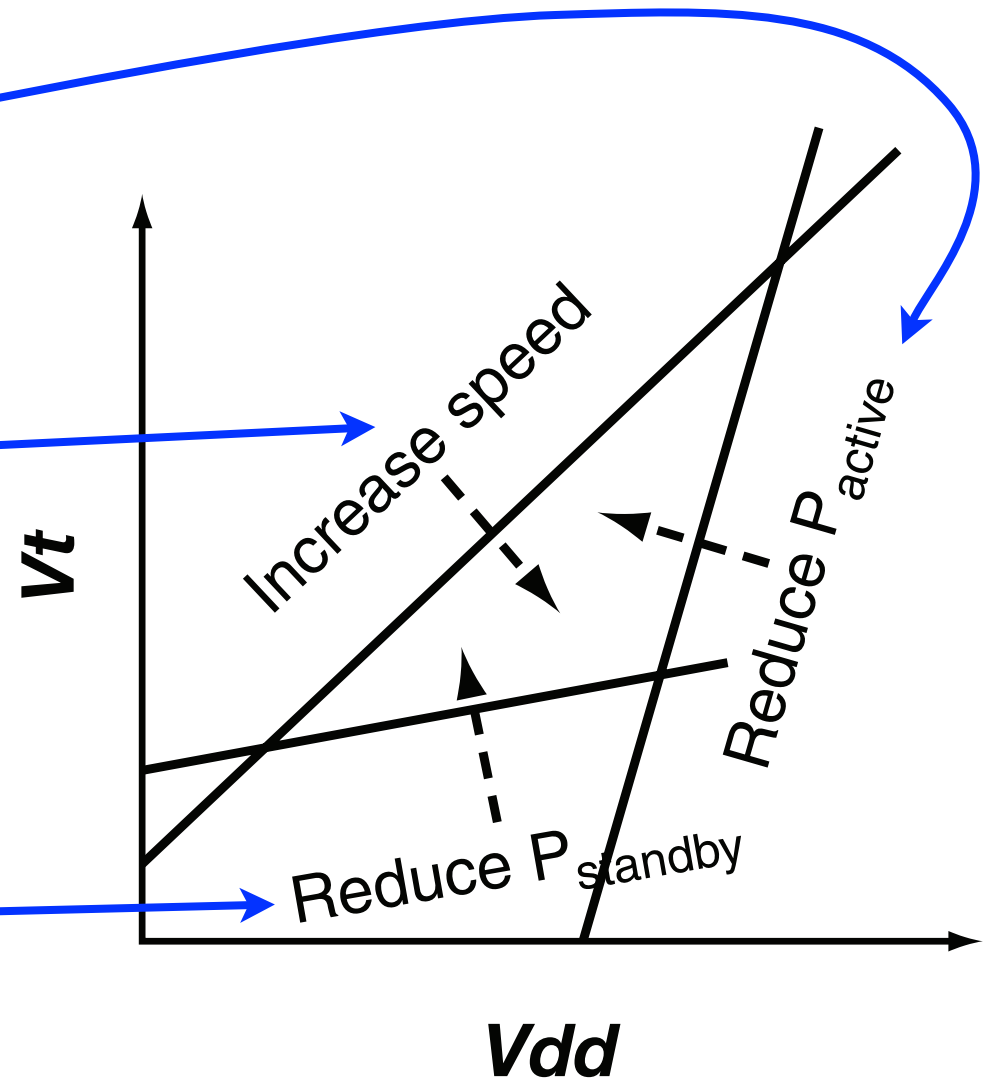
UC Regents Fall 2011 © UCB

Device engineers trade speed and power

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

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

We can reduce leakage ($P_{standby}$) by raising V_t.

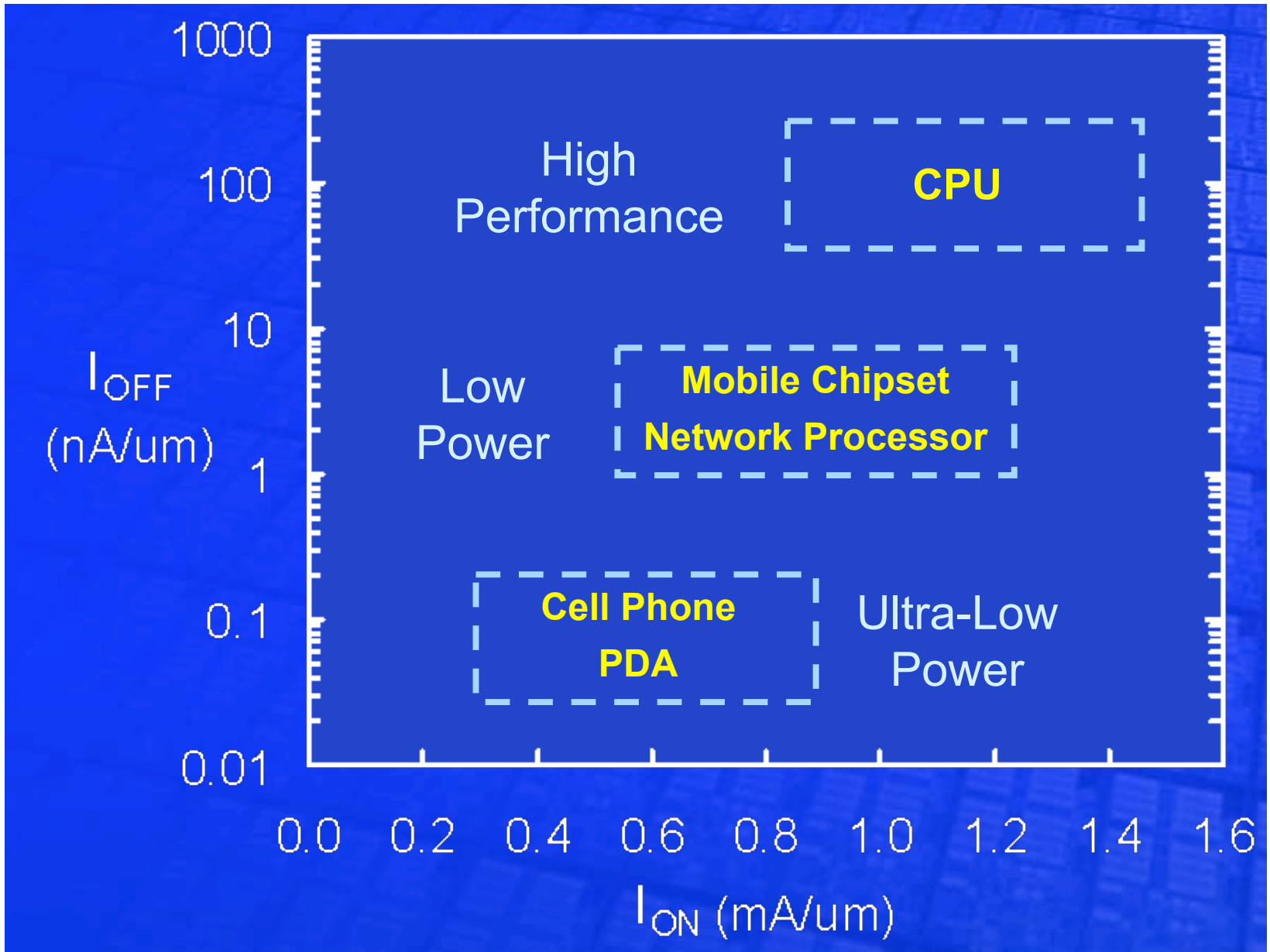


From: Silicon Device Scaling to the Sub-10-nm Regime

Meikei Ieong,[1*] Bruce Doris,[2] Jakub Kedzierski,[1] Ken Rim,[1] Min Yang[1]

UC Regents Fall 2011 © UCB

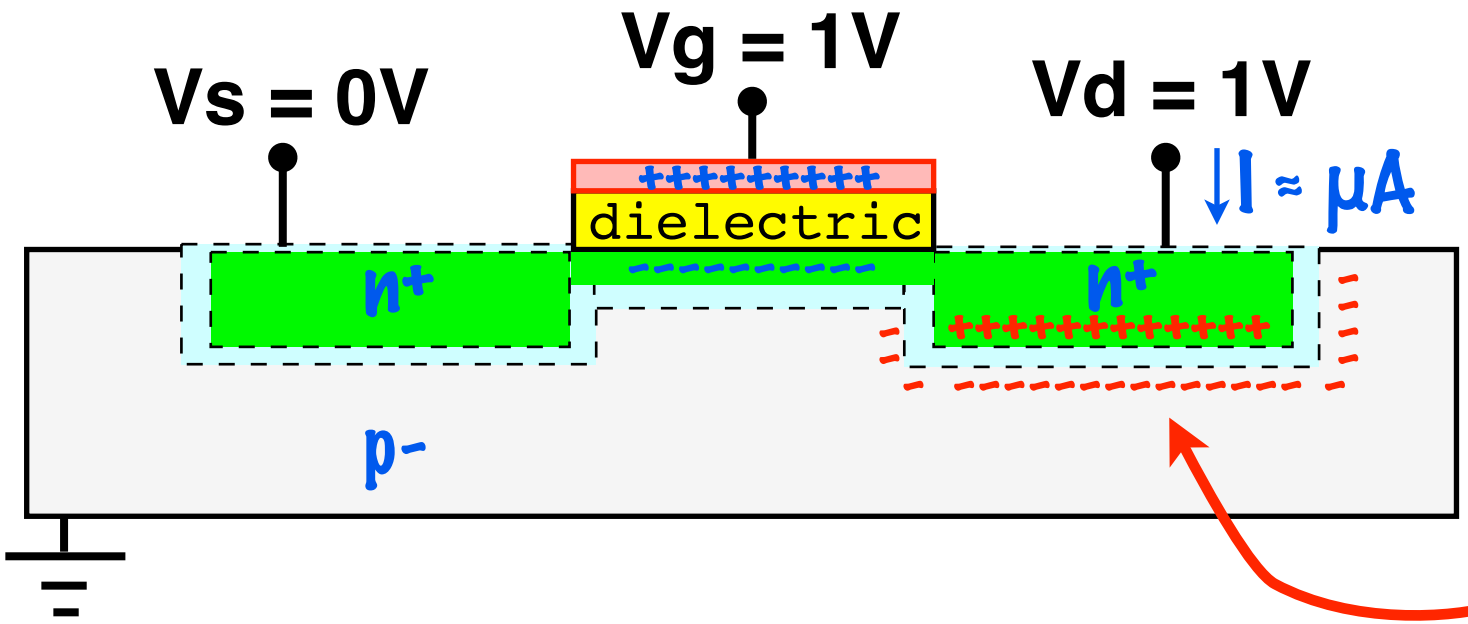
Customize processes for product types ...



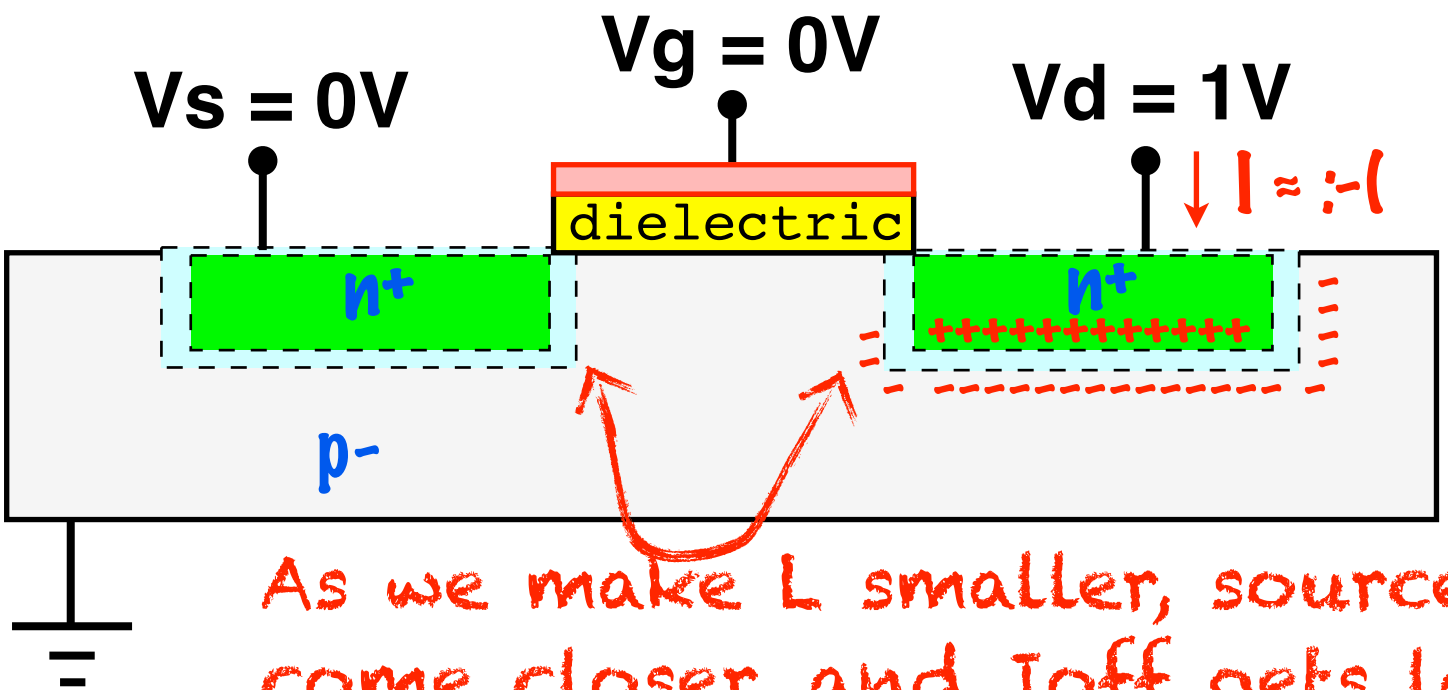
From: "Facing the Hot Chips Challenge Again", Bill Holt, Intel, presented at Hot Chips 17, 2005.

UC Regents Fall 2011 © UCB

Transistor physics revisited ...



The drain junction is also a capacitor, and puts - charges in the substrate.

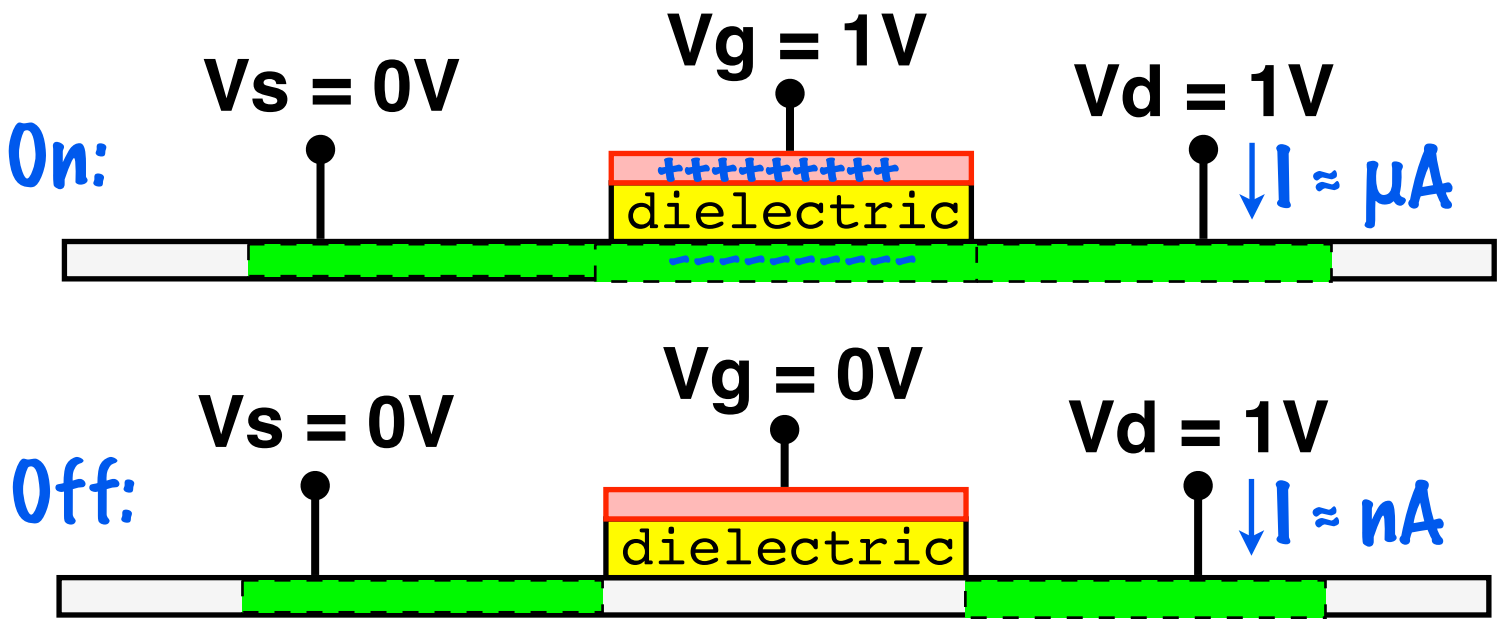


Away from the surface, the drain-induced charges remain even when the gate is off!

As we make L smaller, source and drain come closer, and Ioff gets larger!

UC Regents Fall 2011 © UCB

Solution concept: Fully-depleted channel



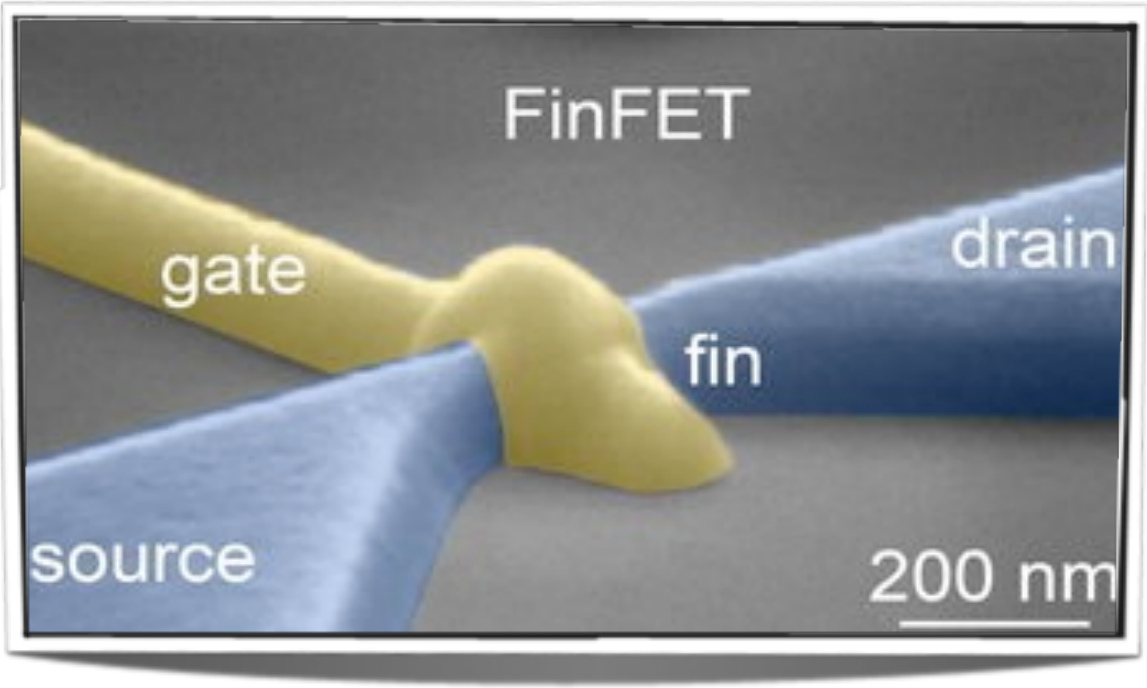
We limit the depth of the channel so that the gate voltage "wins" over the drain voltage.

Done as shown, 5 to 7 nm depth for a 20 nm transistor.

Requires expensive wafers

"FD-SOI" -- Fully-Depleted Silicon-On-Insulator

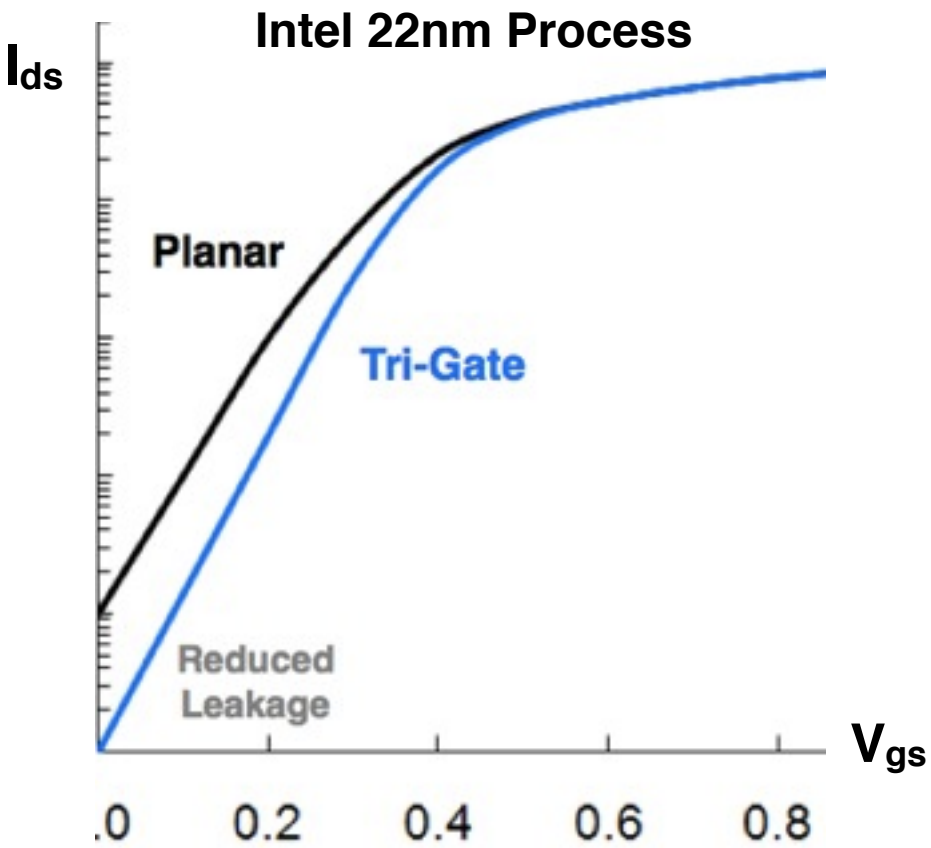
UC Regents Fall 2011 © UCB



Transistor channel is a raised fin.

Gate controls channel from sides and top.

Channel depth is fin width. 12-15nm for L=22nm.



(12) **United States Patent**
Hu et al. Filed: **Oct. 23, 2000**

(54) **FINFET TRANSISTOR STRUCTURES HAVING A DOUBLE GATE CHANNEL EXTENDING VERTICALLY FROM A SUBSTRATE AND METHODS OF MANUFACTURE**

(75) Inventors: **Chenming Hu**, Alamo; **Tsu-Jae King**, Fremont; **Vivek Subramanian**, Redwood City; **Leland Chang**, Berkeley; **Xuejue Huang**; **Yang-Kyu Choi**, both of Albany; **Jakub Tadeusz Kedzierski**, Hayward; **Nick Lindert**, Berkeley; **Jeffrey Bokor**, Oakland, all of CA (US); **Wen-Chin Lee**, Beaverton, OR (US)

UC Regents Fall 2011 © UCB

Five low-power design techniques

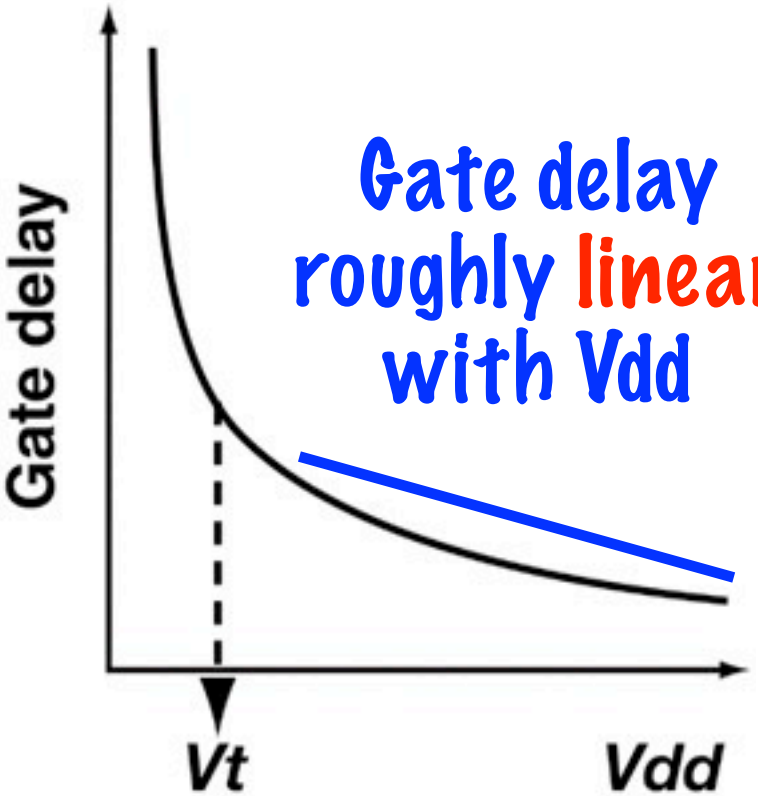
- Parallelism and pipelining
- Power-down idle transistors
- Slow down non-critical paths
- Clock gating
- Thermal management

UC Regents Fall 2011 © UCB

Trading Hardware for Power

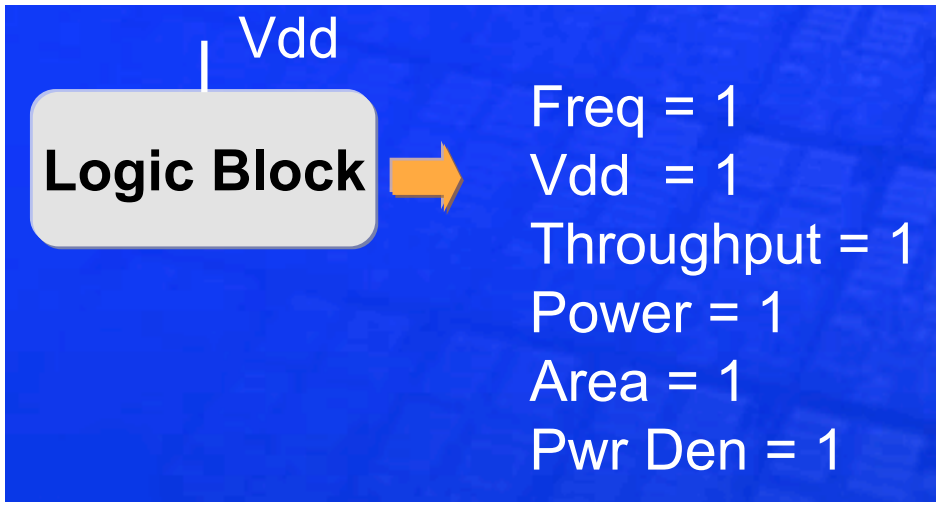
via Parallelism and Pipelining ...

UC Regents Fall 2011 © UCB



Gate delay roughly linear with Vdd

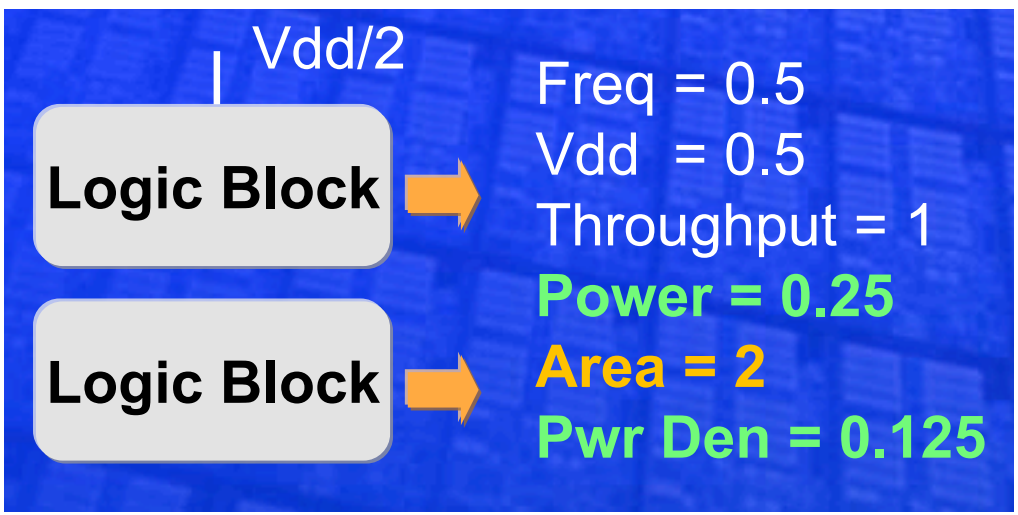
And so, we can transform this:



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.

THIS MAGIC TRICK BROUGHT TO YOU BY CORY HALL …

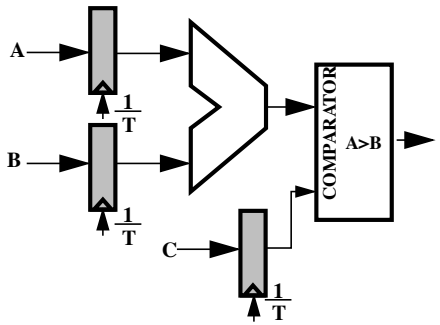
UC Regents Fall 2011 © UCB

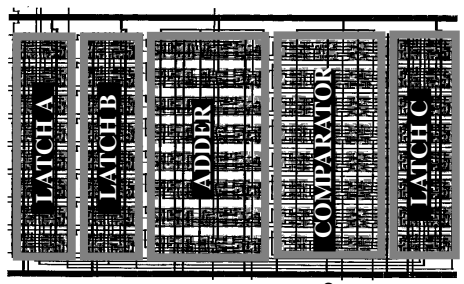
Chandrakasan & Brodersen (UCB EECS)

Architecture	Power (normalized)
Simple	1
Parallel	0.36
Pipelined	0.39
Pipelined-Parallel	0.2

Architecture	Area (normalized)
Simple	1
Parallel	3.4
Pipelined	1.3
Pipelined-Parallel	3.7

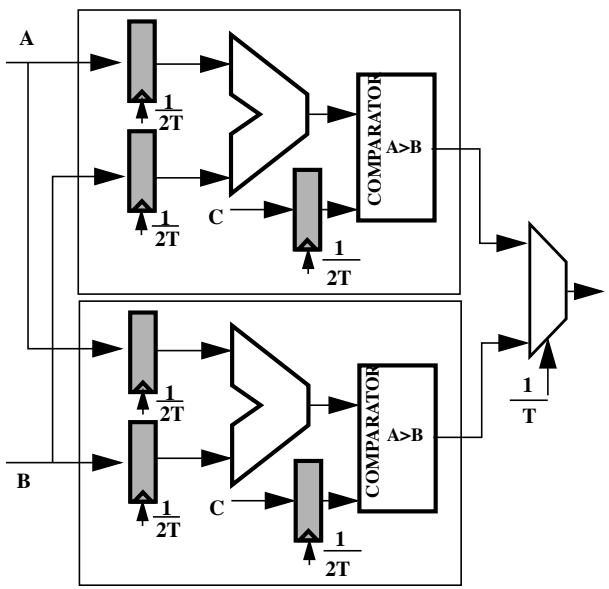
Architecture	Voltage
Simple	5V
Parallel	2.9V
Pipelined	2.9V
Pipelined-Parallel	2.0

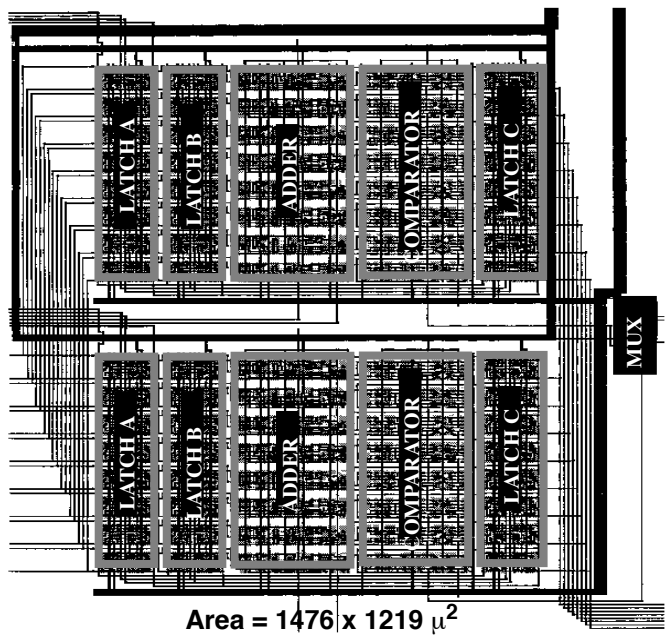




Area = 636 x 833 μ^2

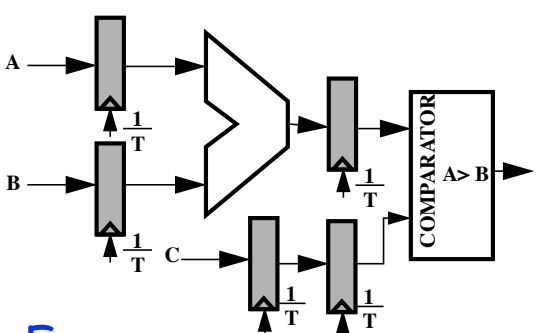
Simple

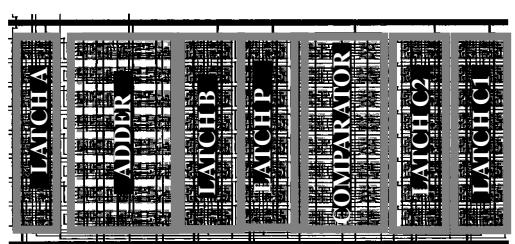




Area = 1476 x 1219 μ^2

Parallel





Area = 640 x 1081 μ^2

Pipelined

From:

Minimizing Power Consumption in CMOS Circuits Anantha P. Chandrakasan, Robert W. Brodersen

JC Regents Fall 2011 © UCB

Multiple Cores for Low Power

Trade hardware for power,
on a large scale ...

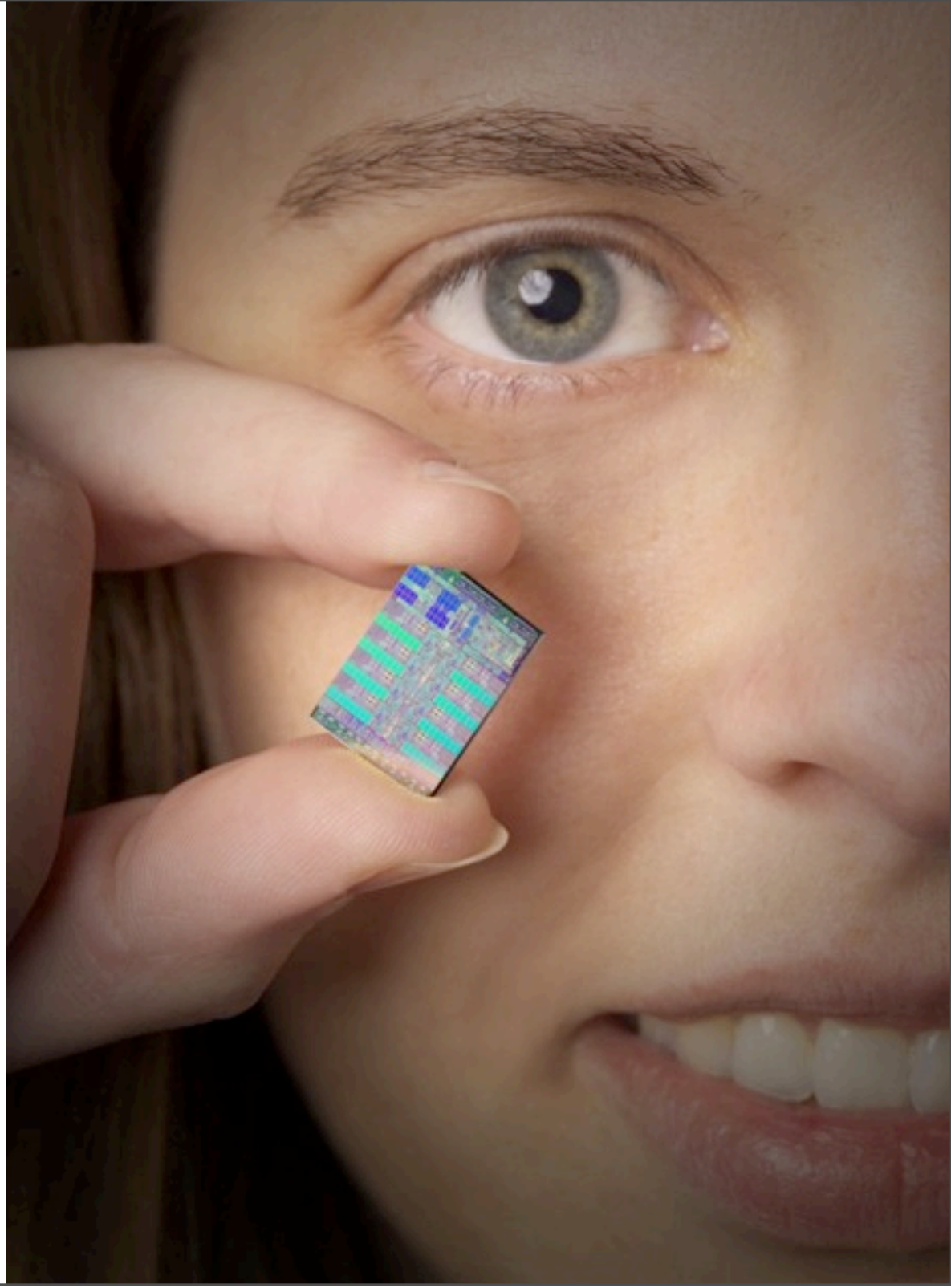
UC Regents Fall 2011 © UCB

Cell: The PS3 chip

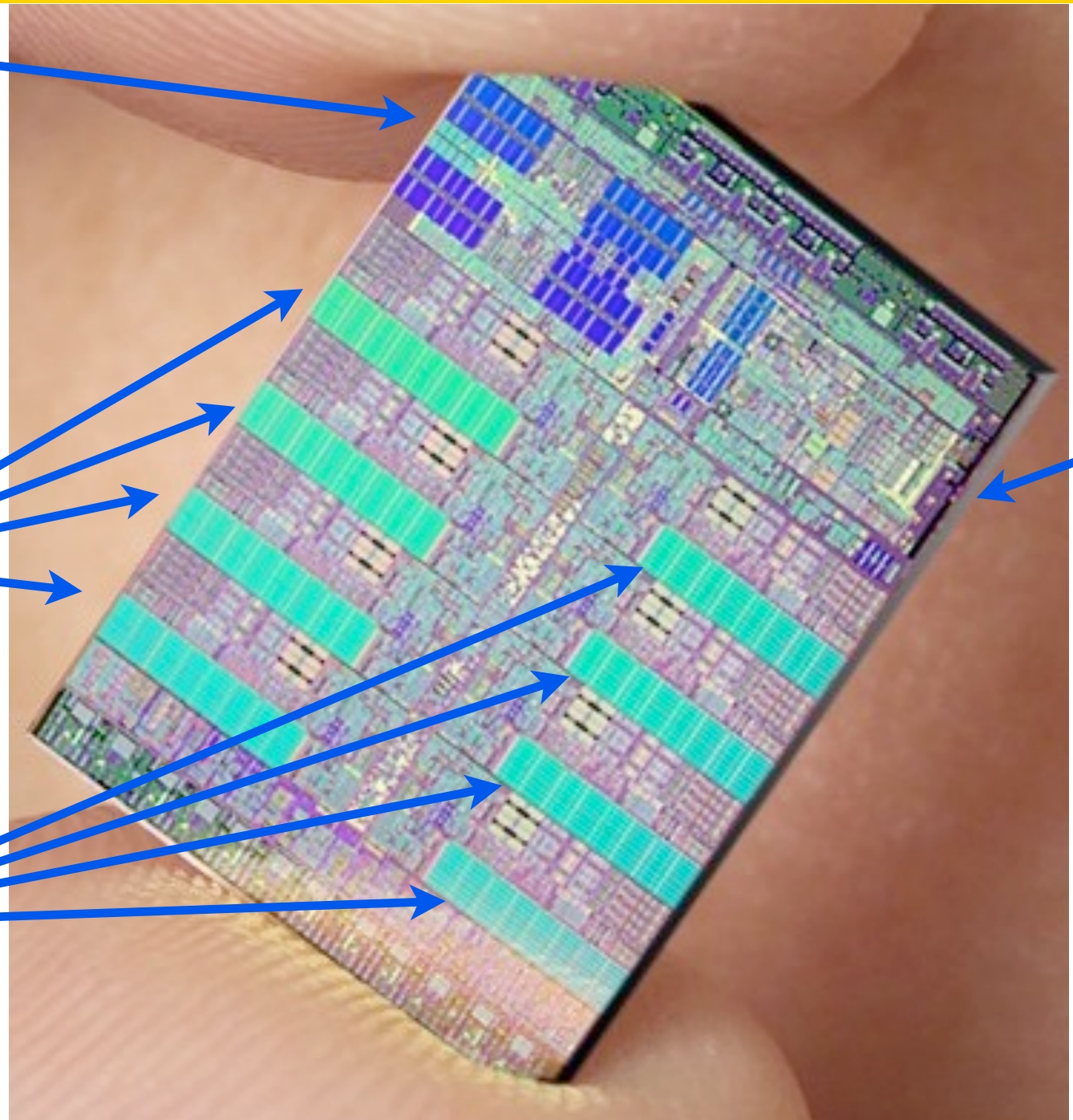
IBM

SONY COMPUTER ENTERTAINMENT

TOSHIBA

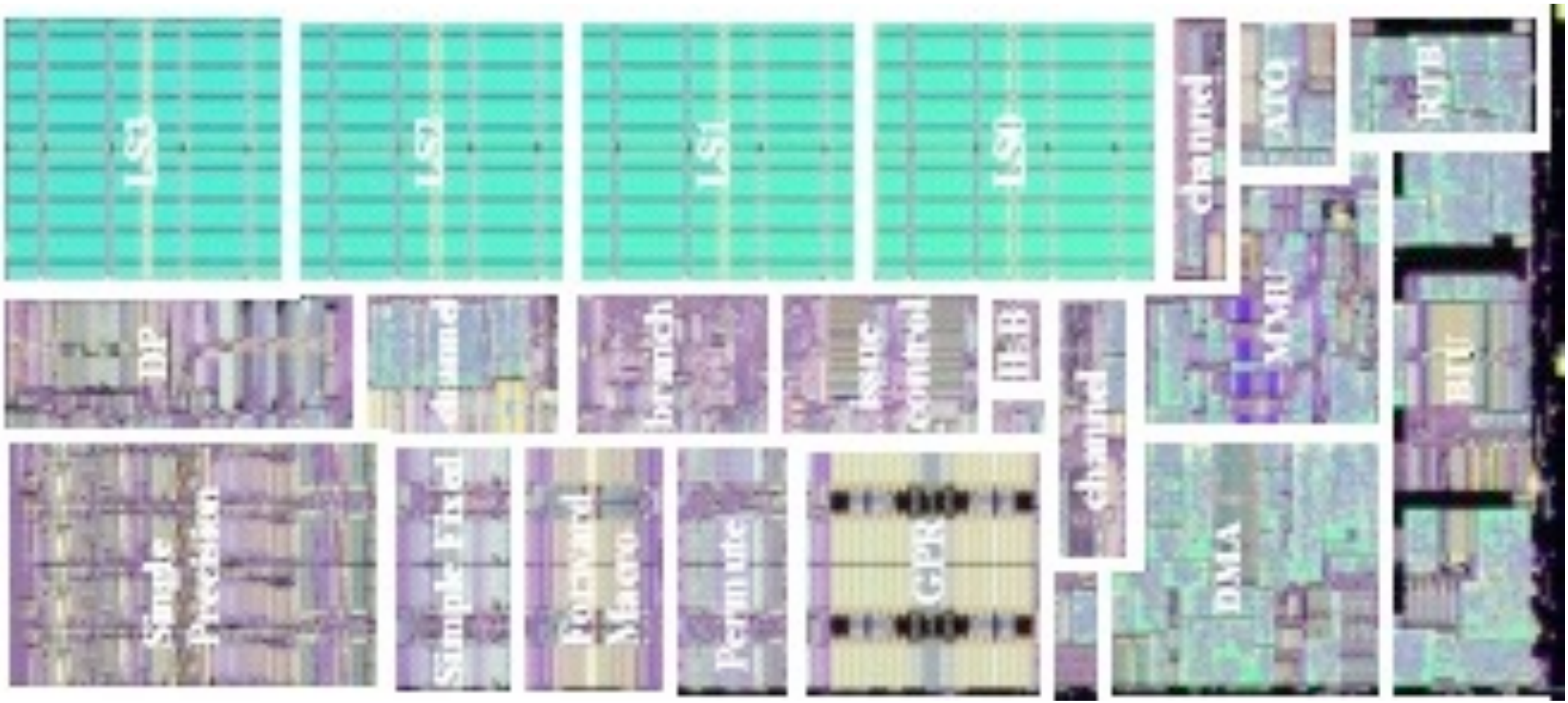


Cell (PS3 Chip): 1 CPU + 8 “SPUs”



UC Regents Fall 2011 © UCB

One Synergistic Processing Unit (SPU)



SPU issues 2 inst/cycle (in order) to 7 execution units

256 KB Local Store, 128 128-bit Registers

SPU fills Local Store using DMA to DRAM and network

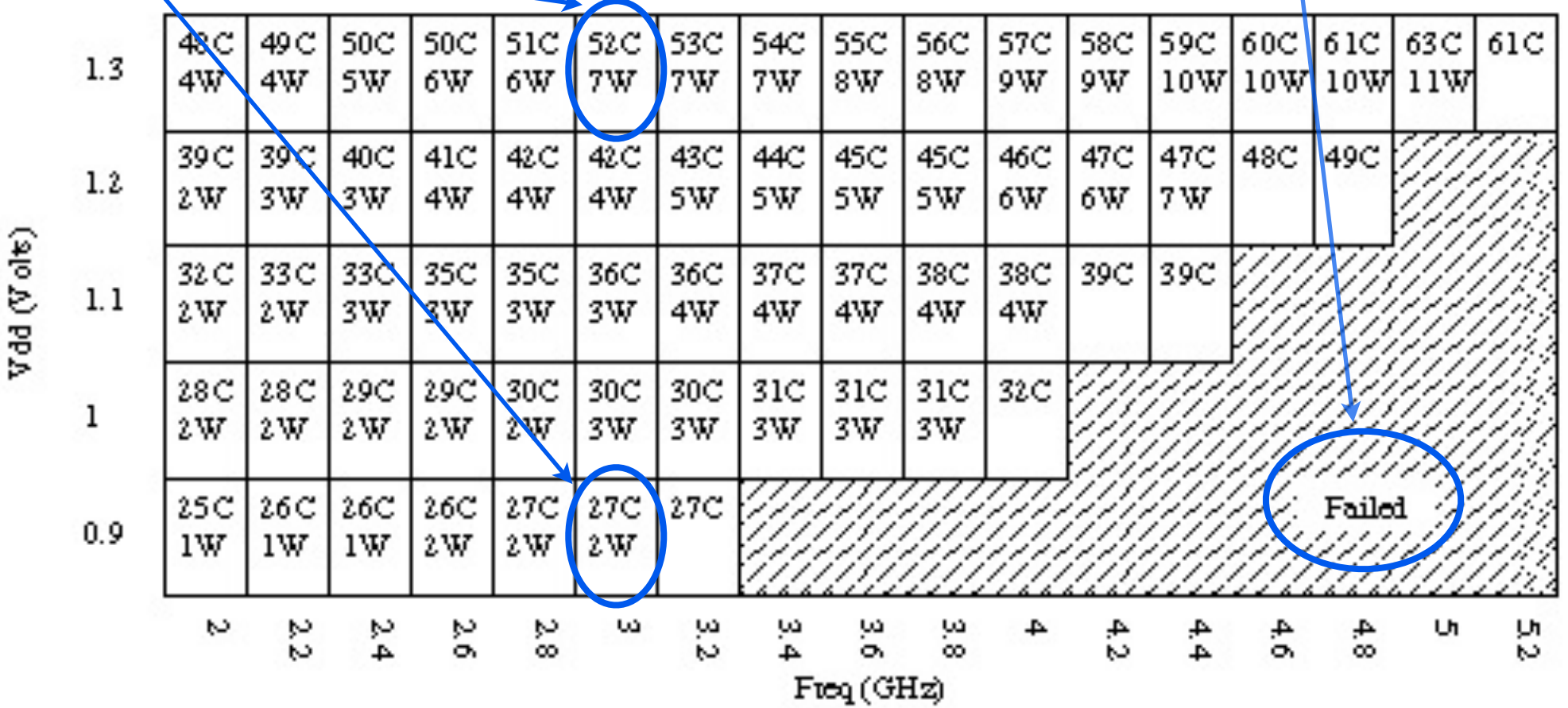
UC Regents Fall 2011 © UCB

A "Schmoo" plot for a Cell SPU ...

The lower Vdd, the less dynamic energy consumption.

$$E_{0->1} = \frac{1}{2} C V_{dd}^2 \qquad E_{1->0} = \frac{1}{2} C V_{dd}^2$$

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

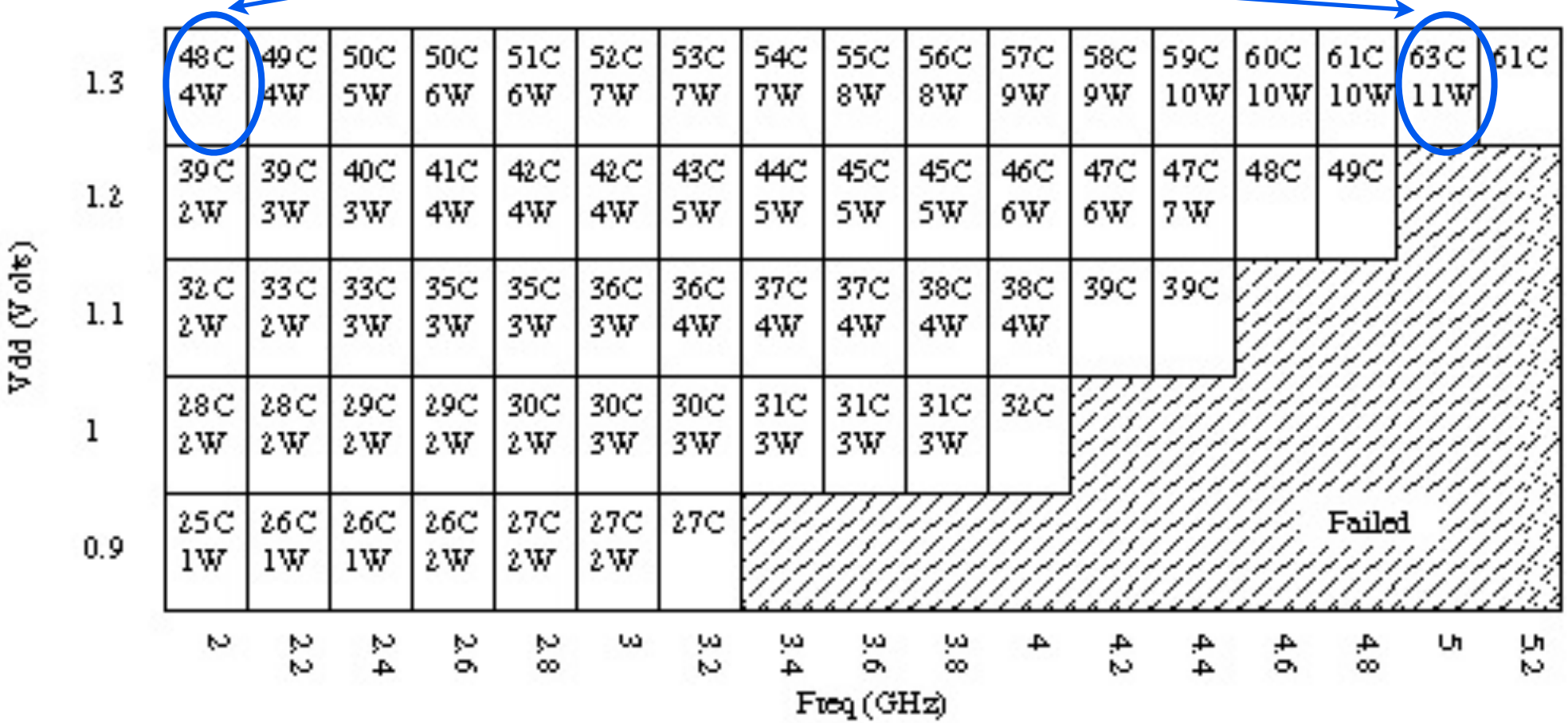


UC Regents Fall 2011 © UCB

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 ...

$$E_{0->1} = \frac{1}{2} C V_{dd}^2 \qquad E_{1->0} = \frac{1}{2} C V_{dd}^2$$



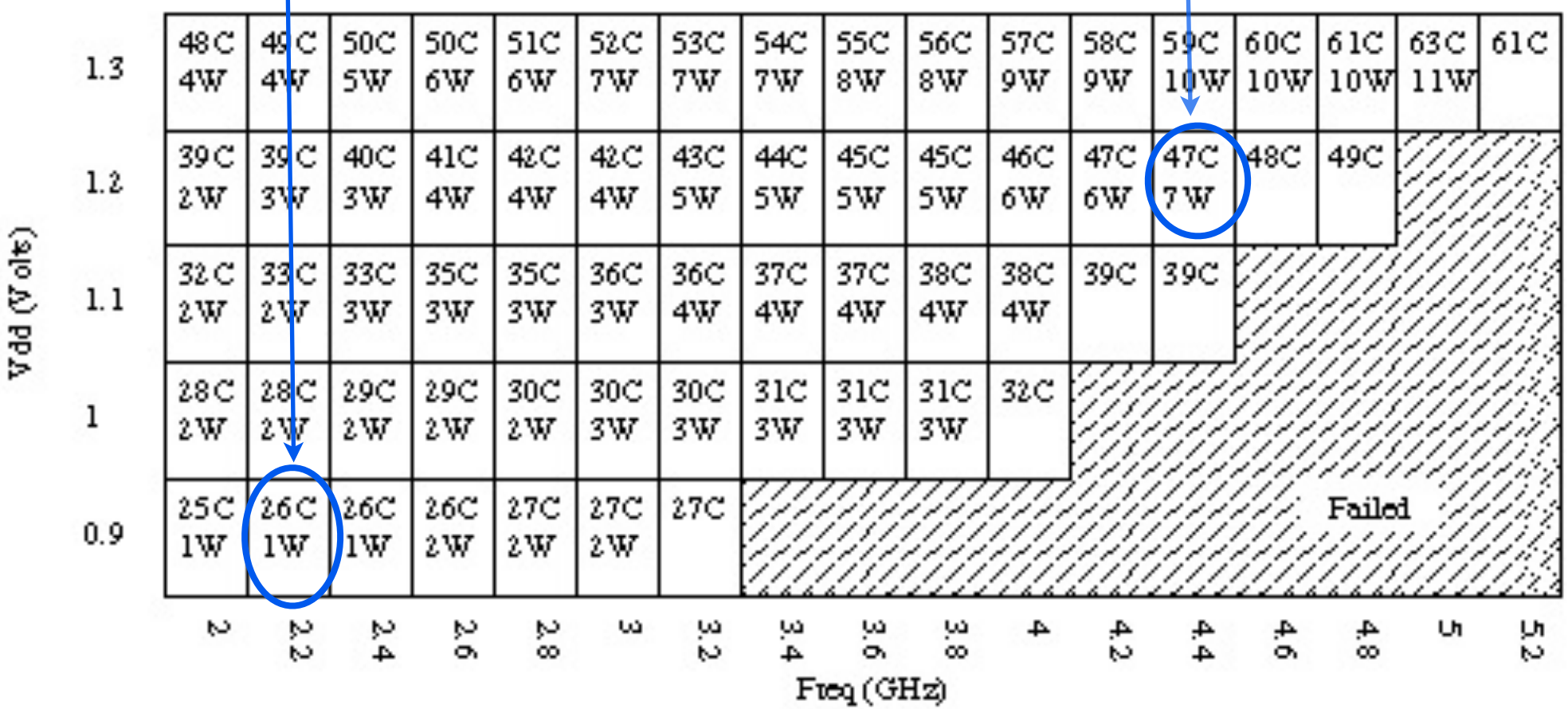
UC Regents Fall 2011 © UCB

Scaling V and f *does* lower energy/op

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

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

If a program that needs a 4.4 Ghz CPU can be recoded to use two 2.2 Ghz CPUs ... big win.

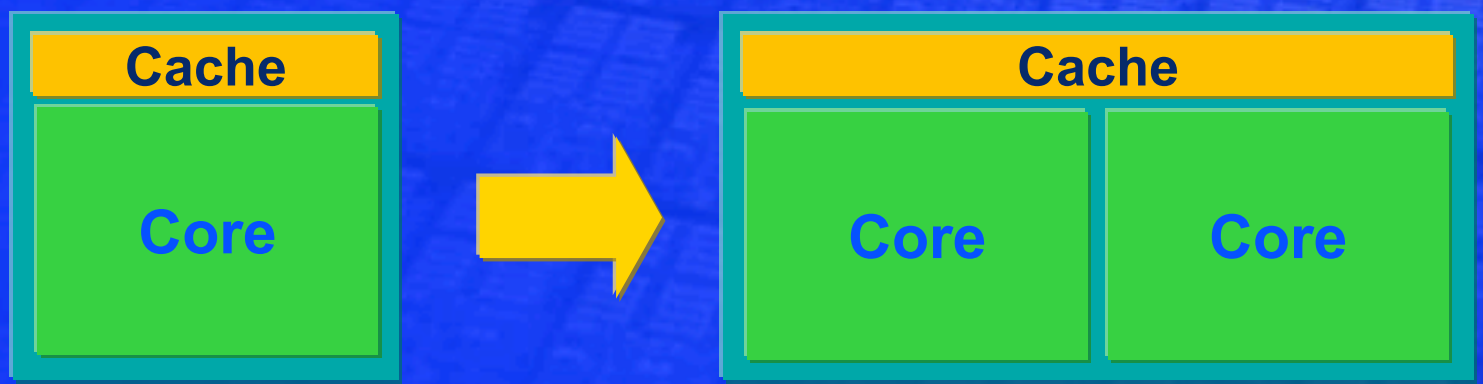


UC Regents Fall 2011 © UCB

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**

From: "Facing the Hot Chips Challenge Again", Bill Holt, Intel, presented at Hot Chips 17, 2005.

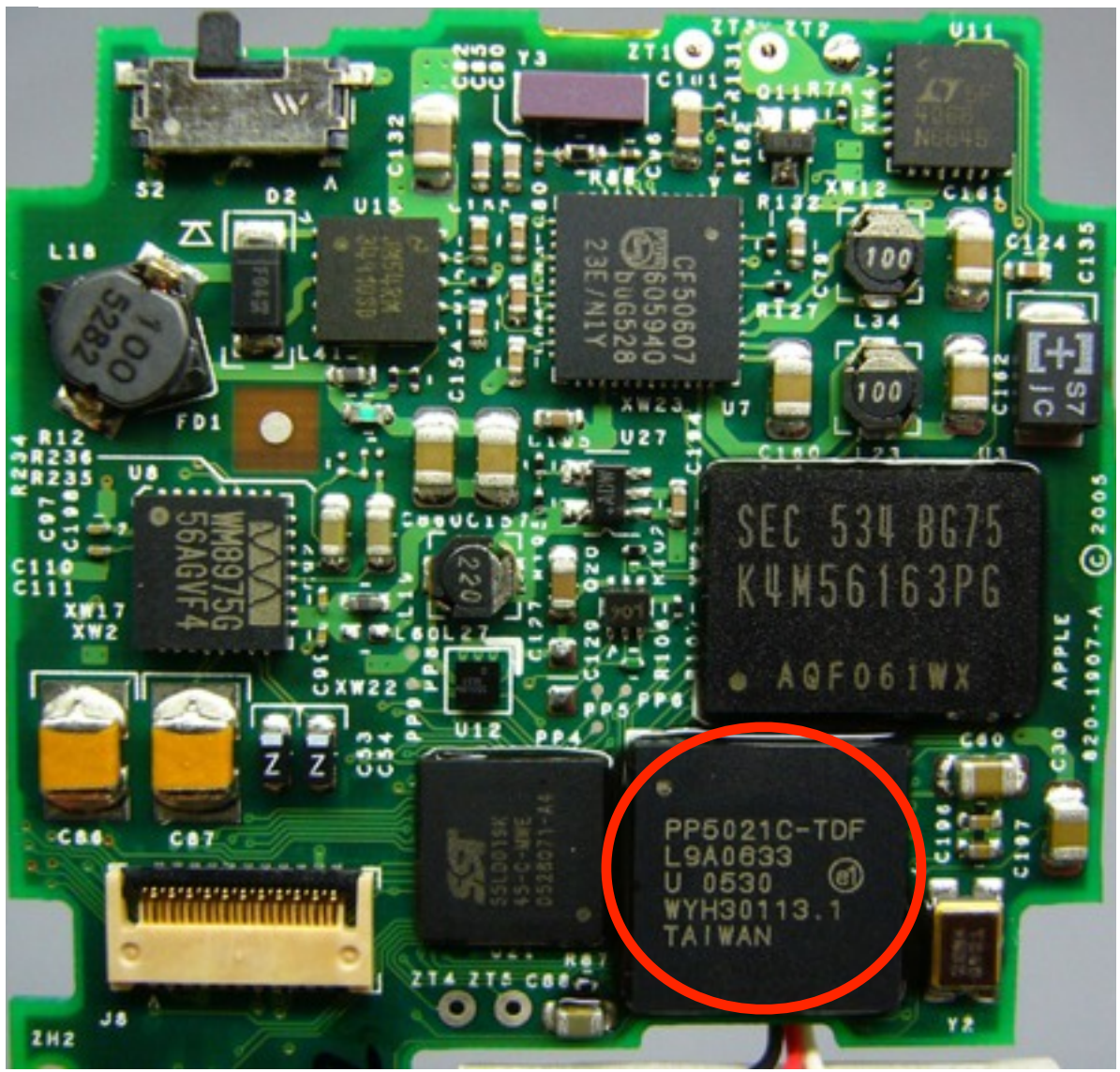
UC Regents Fall 2011 © UCB

How iPod nano 2005 puts its 2 cores to use ...

portalplayer™

PP5020

digital media management system-on-chip



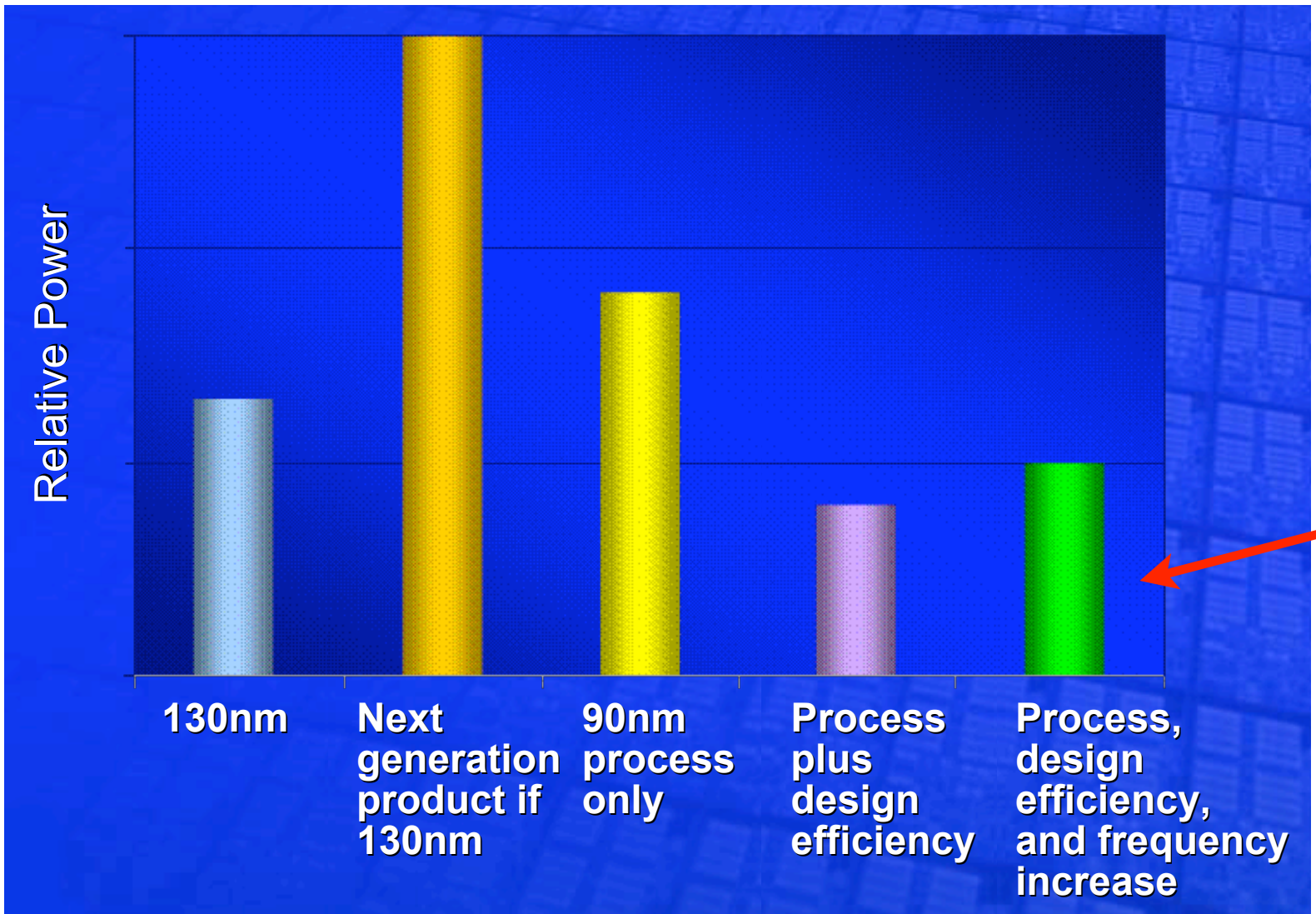
Dual ARM Processors

- Dual 32-bit ARM7TDMI processors
- Up to 80 MHz processor operation per core with independent clock-skipping feature on COP
- Efficient cross-bar implementation providing zero wait state access to internal RAM
- Integrated 96KB of SRAM
- 8KB of unified cache per processor
- Six DMA channels

Two 80 MHz CPUs. Was used in several nano generations, with one CPU doing audio decoding, the other doing photos, etc.

UC Regents Fall 2011 © UCB

Intel: Comparing 2 CPU generations ...



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.

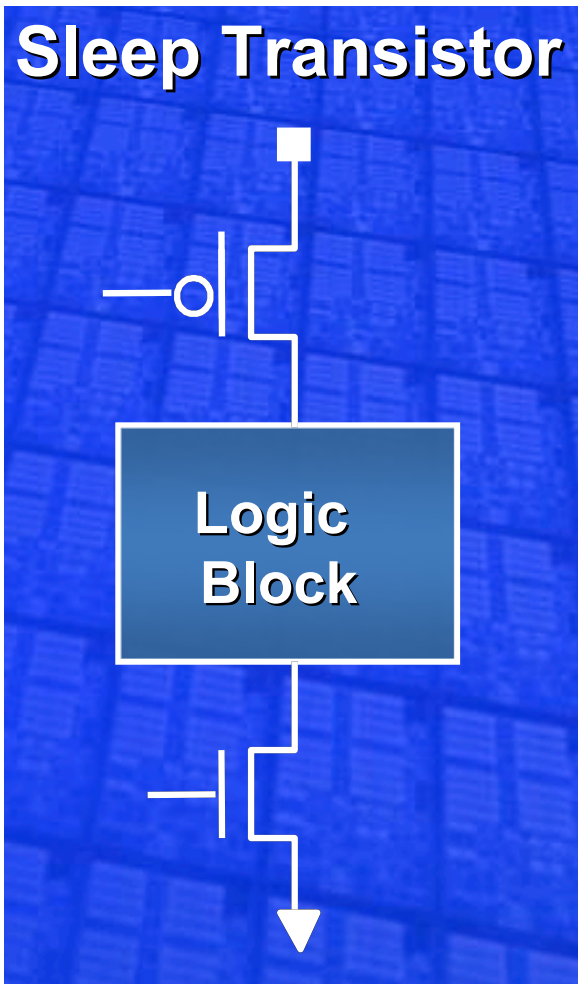
Design tricks: architecture & circuits.

UC Regents Fall 2011 © UCB

Powering down idle circuits

UC Regents Fall 2011 © UCB

Add “sleep” transistors to logic ...



Example: Floating point unit logic.

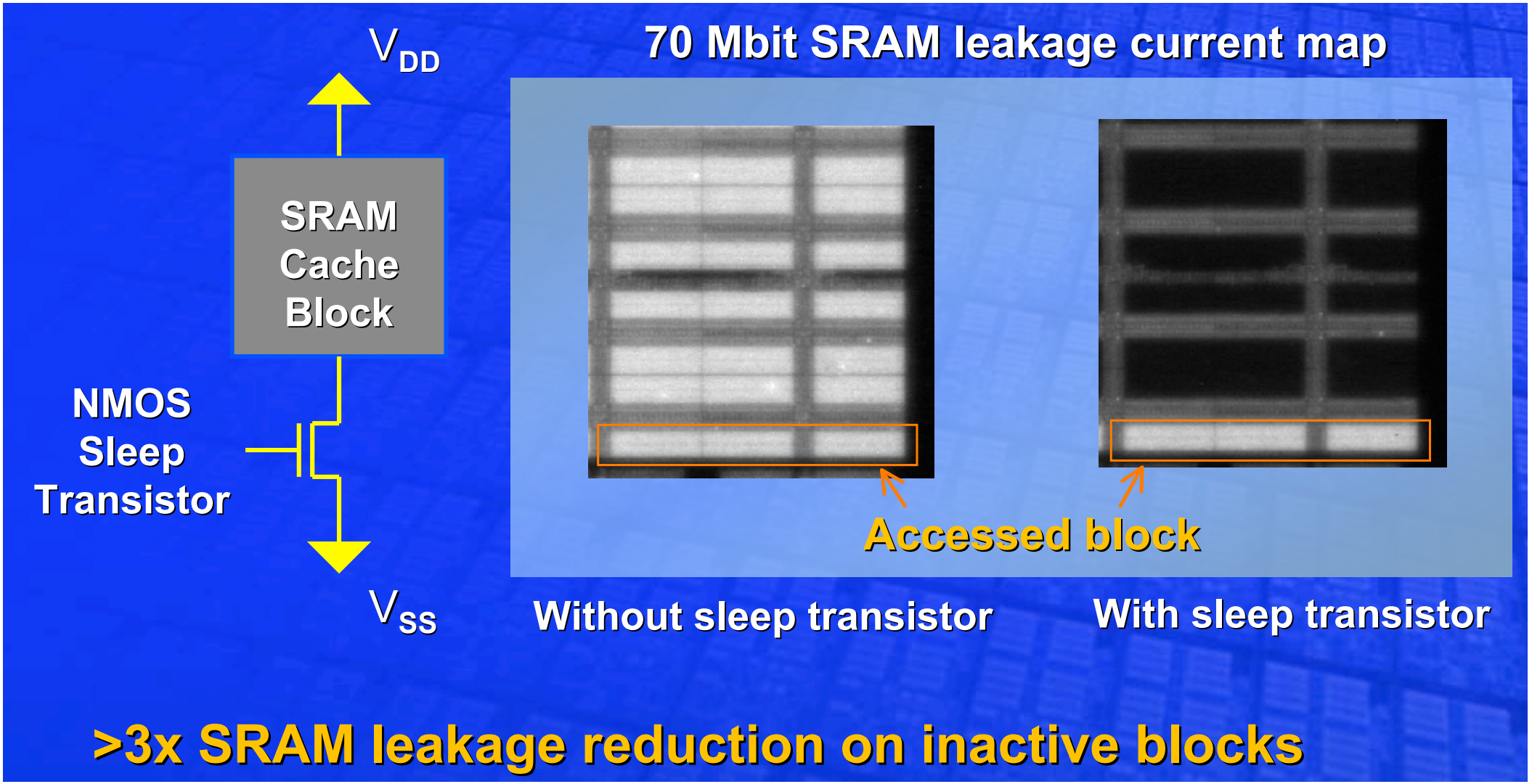
When running fixed-point instructions, put logic “to sleep”.

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

--- Presence of sleep transistors slows down the clock rate when the logic block is in use.

UC Regents Fall 2011 © UCB

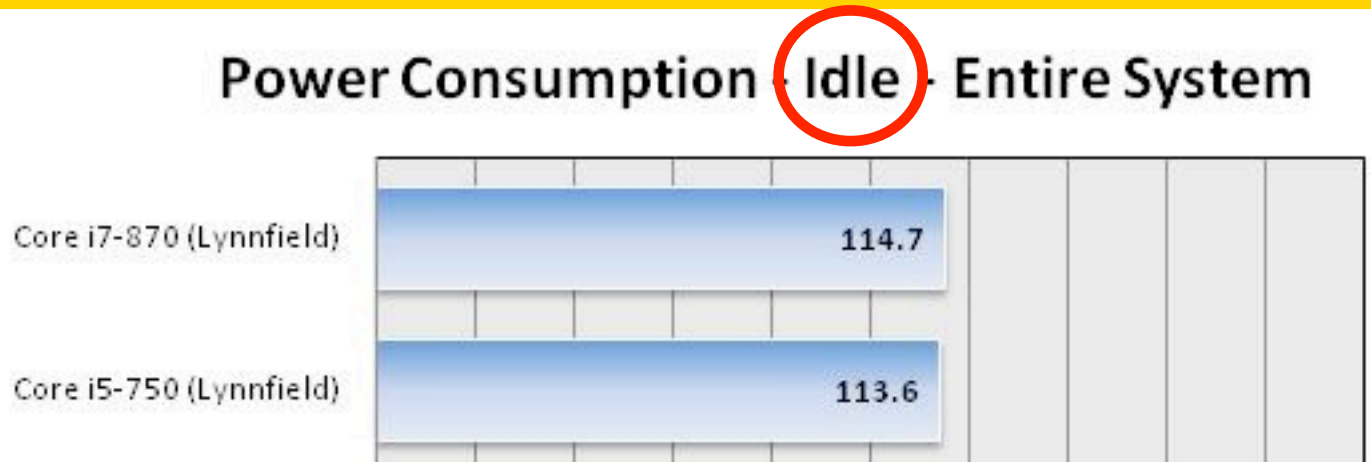
Intel example: Sleeping cache blocks



From: "Facing the Hot Chips Challenge Again", Bill Holt, Intel, presented at Hot Chips 17, 2005.

UC Regents Fall 2011 © UCB

Fall 2009: Intel Mainstream Desktop



From: PC Perspective website

UC Regents Fall 2011 © UCB

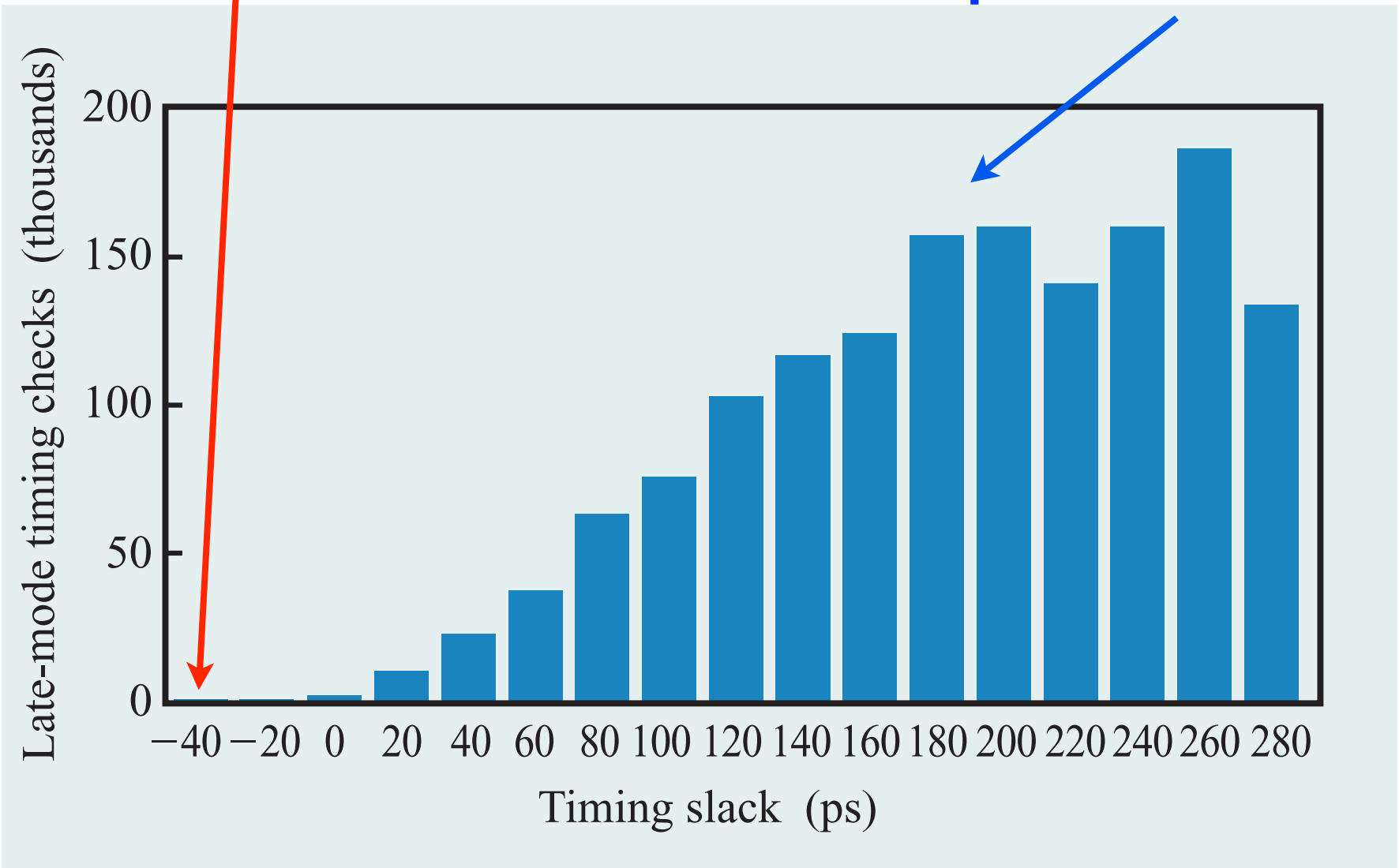
Slow down “slack paths”

UC Regents Fall 2011 © UCB

Fact: Most logic on a chip is "too fast"

The critical path

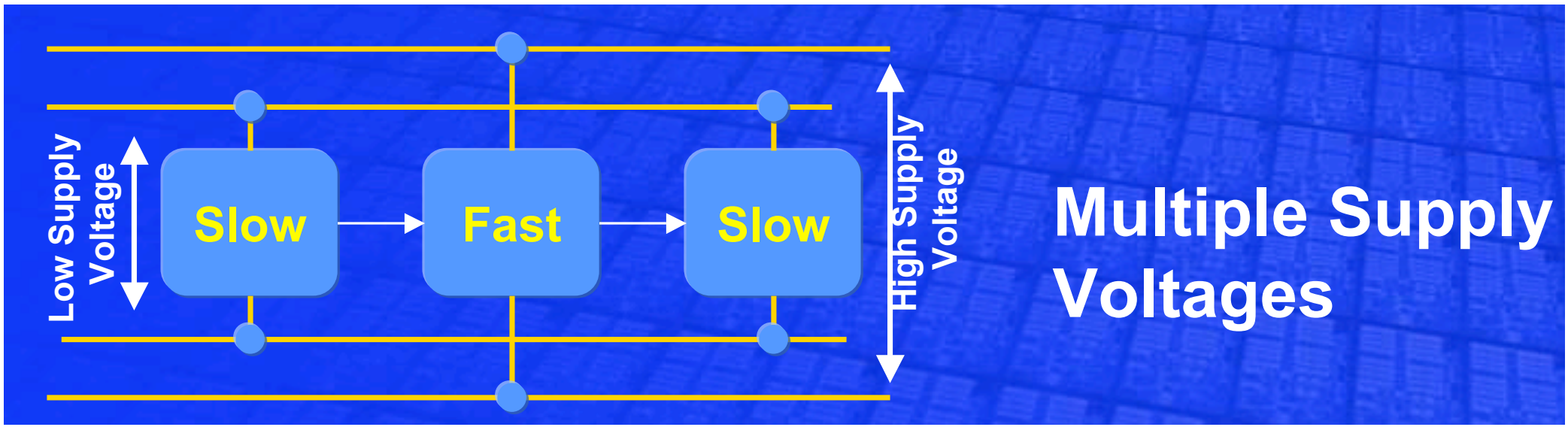
Most wires have hundreds of picoseconds to spare.



From "The circuit and physical design of the POWER4 microprocessor", IBM J Res and Dev, 46:1, Jan 2002, J.D. Warnock et al.

UC Regents Fall 2011 © UCB

Use several supply voltages on a chip …



Why use multi-Vdd? We can reduce dynamic power by using low-power Vdd for logic off the critical path.

What if we can't do a multi-Vdd design?
In a multi-Vt process, we can reduce leakage power on the slow logic by using high-Vth transistors.

From: "Facing the Hot Chips Challenge Again", Bill Holt, Intel, presented at Hot Chips 17, 2005.

UC Regents Fall 2011 © UCB

LOW POWER ARM 1136JF-S™ DESIGN

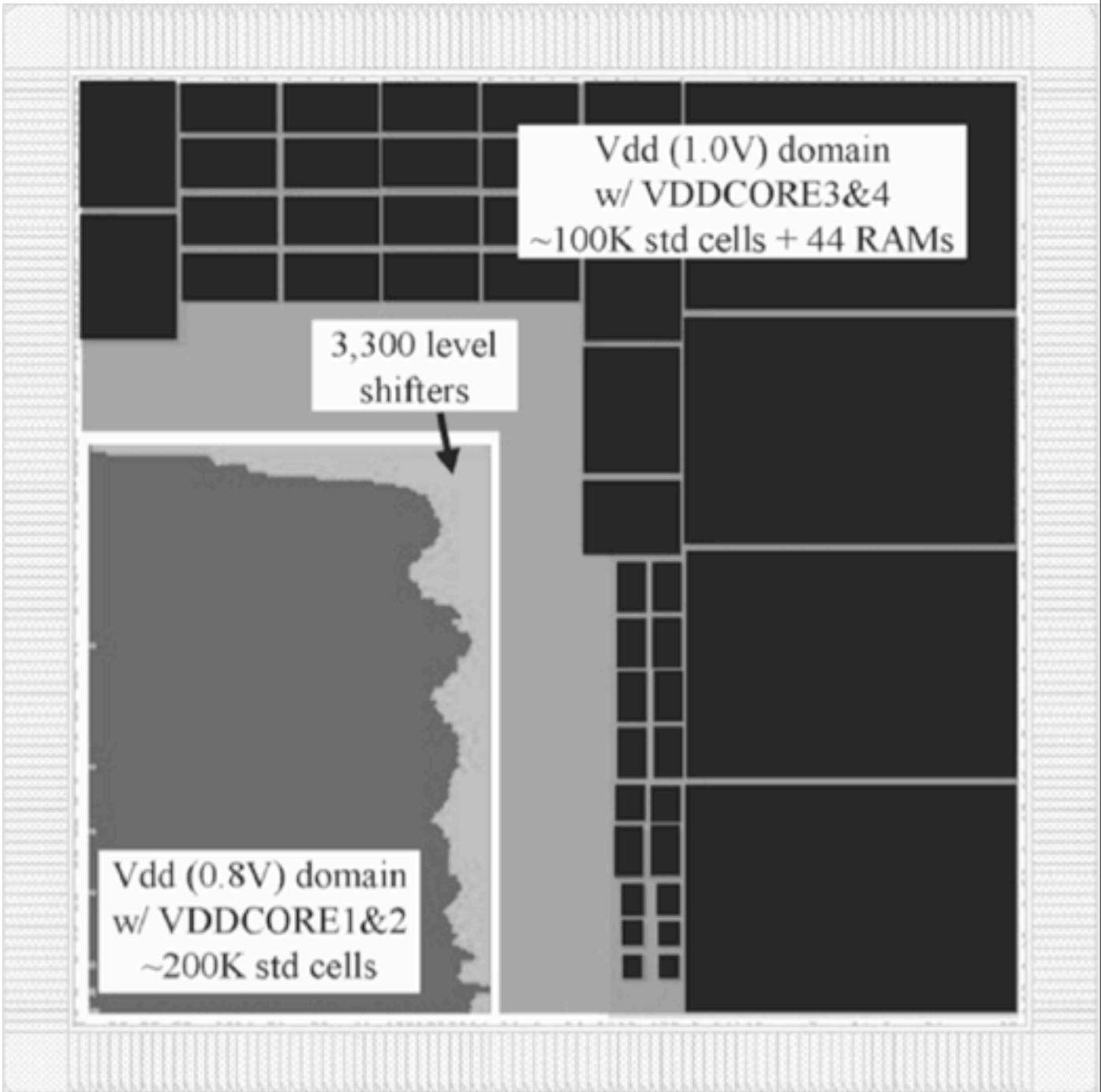
George Kuo, Anand Iyer
Cadence Design Systems, Inc.
San Jose, CA 95134, USA

Logical partition into 0.8V and 1.0V nets done manually to meet 350 MHz spec (90nm).

Level-shifter insertion and placement done automatically.

Dynamic power in 0.8V section cut 50% below baseline.

Leakage power in 1.0V section cut 70% below baseline.



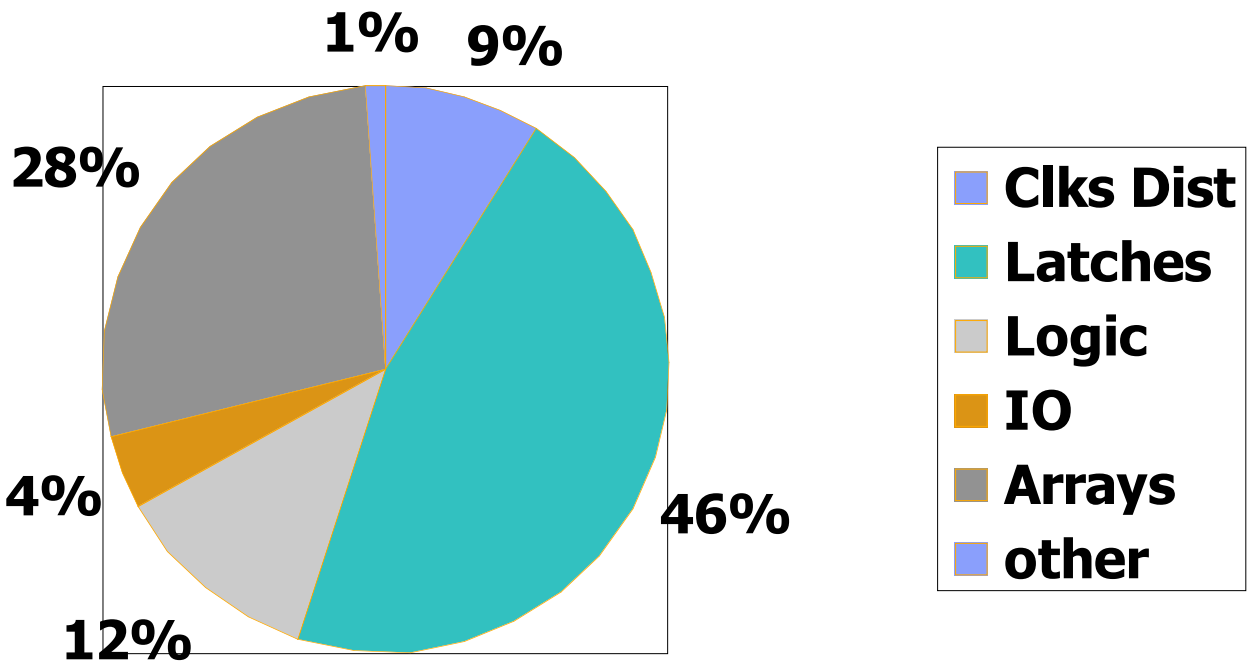
From a chapter from new book on ASIC design by Chinnery and Keutzer (UCB).

UC Regents Fall 2011 © UCB

Gating clocks to save power

UC Regents Fall 2011 © UCB

On a CPU, where does the power go?



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

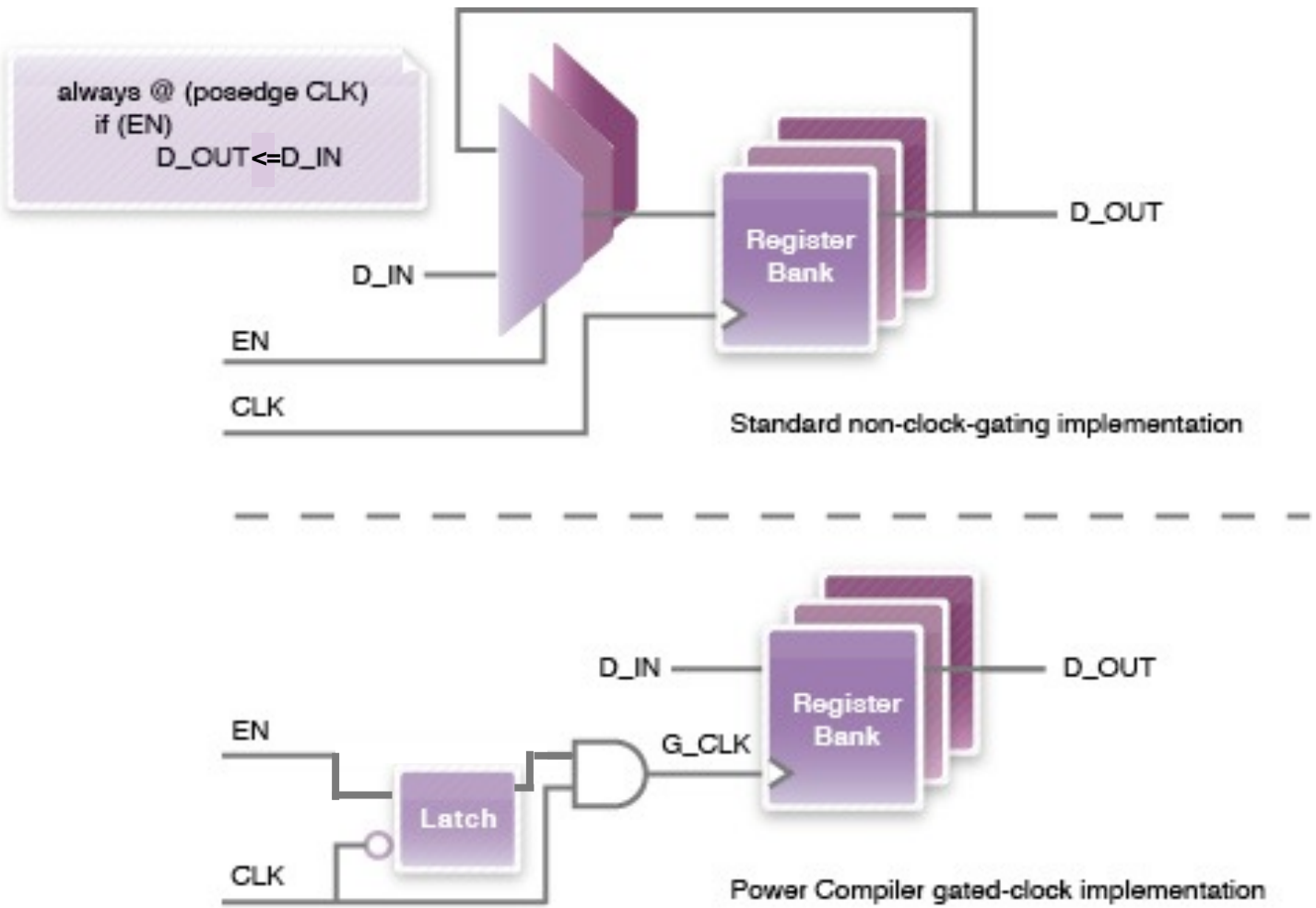
Most of the time, the latches don't change state.

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

From: Bose, Martonosi, Brooks: Sigmetrics-2001 Tuto

UC Regents Fall 2011 © UCB

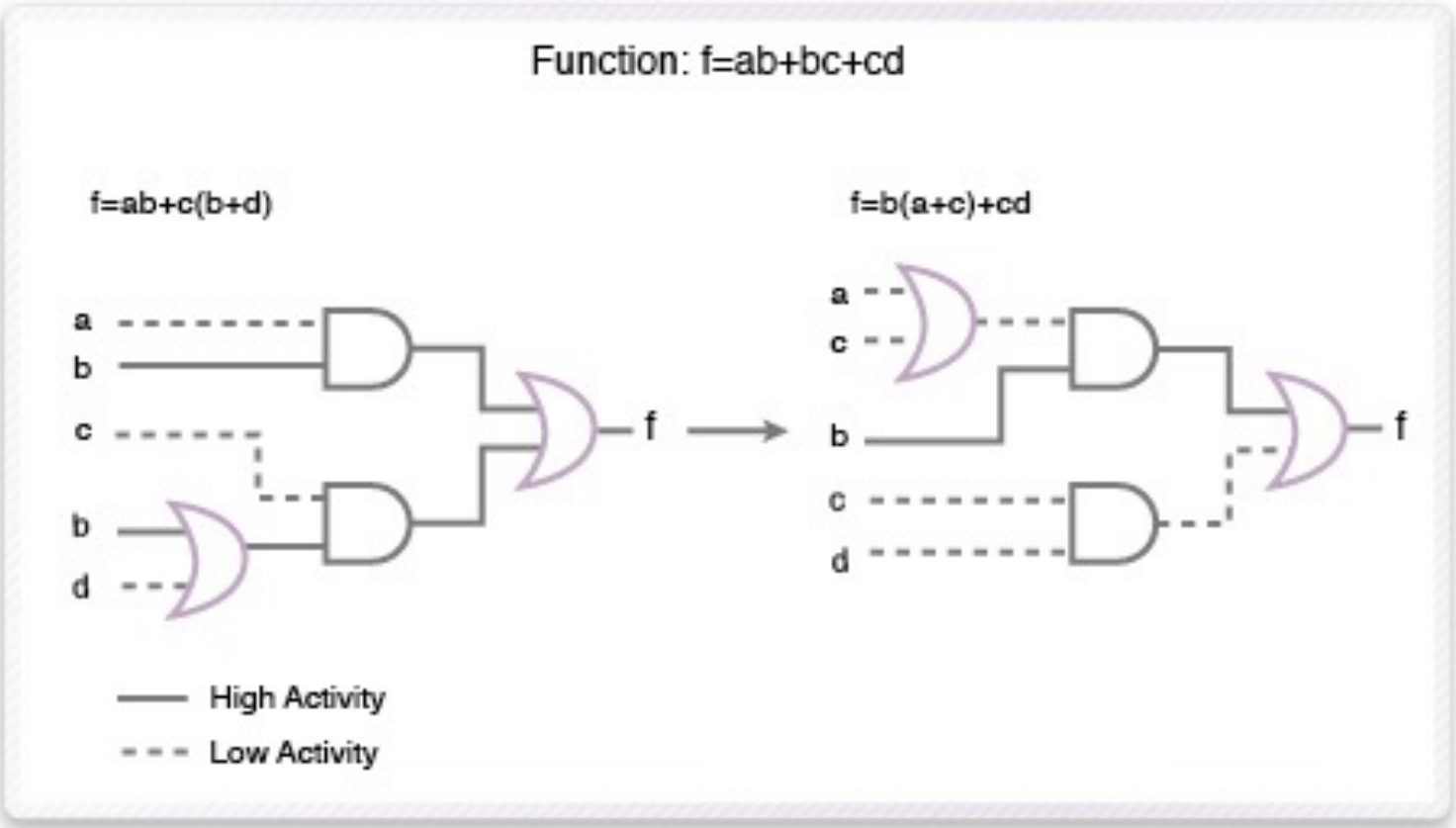
Synopsis Power Compiler can do this ...



"Up to 70% power savings at the block level, for applicable circuits" Synopsis Data Sheet

UC Regents Fall 2011 © UCB

Power Compiler also can do this ...



10-20% push-button power savings, using techniques like this one.

UC Regents Fall 2011 © UCB

Thermal Management

UC Regents Fall 2011 © UCB

Keep chip cool to minimize leakage power

A recipe for thermal runaway

Junction Temperature (T_J °C)	Normalized Static Power or I_{CCINTQ} Typical
25	1.00
50	1.46
85	2.50
100	3.14

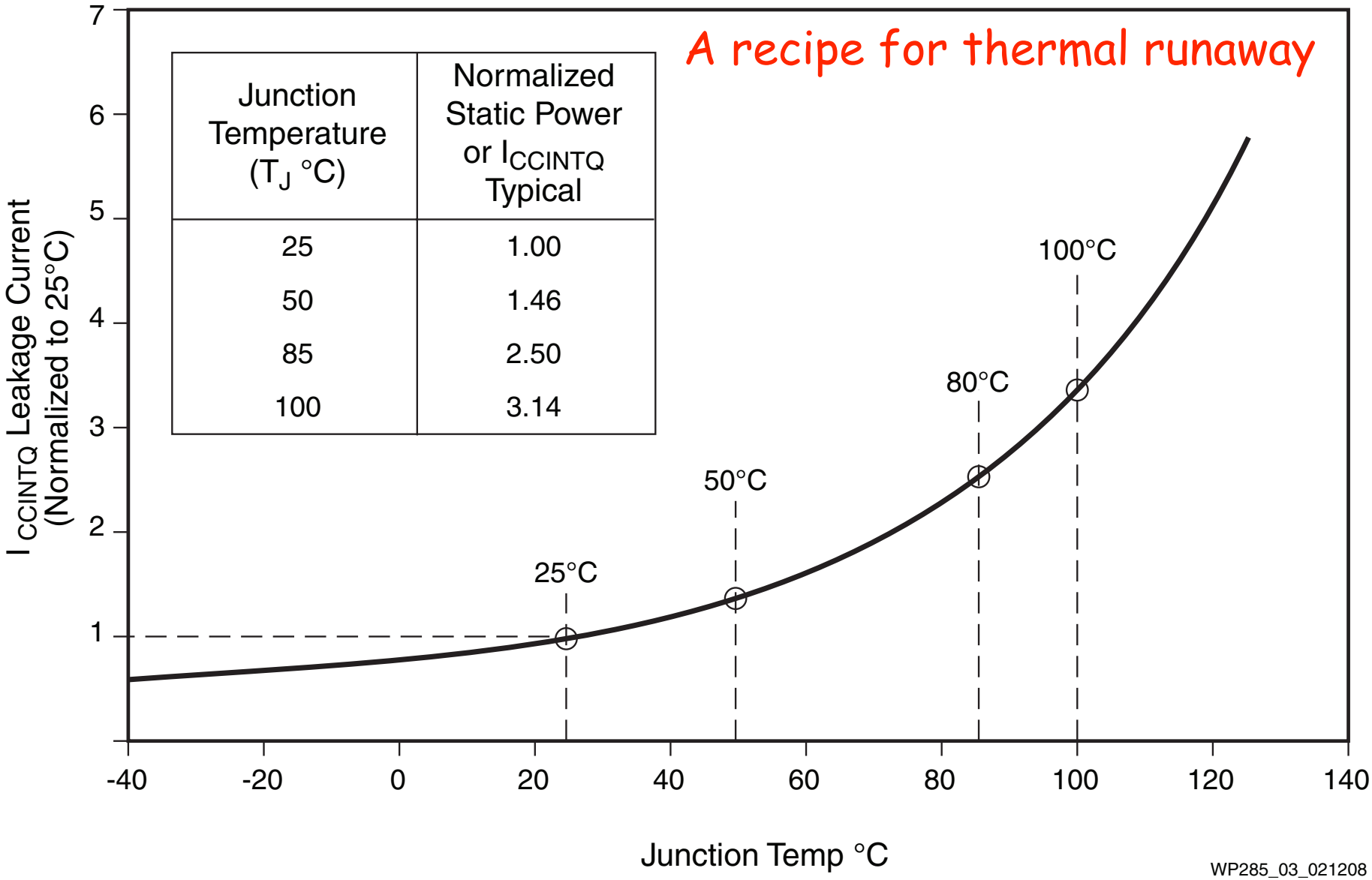
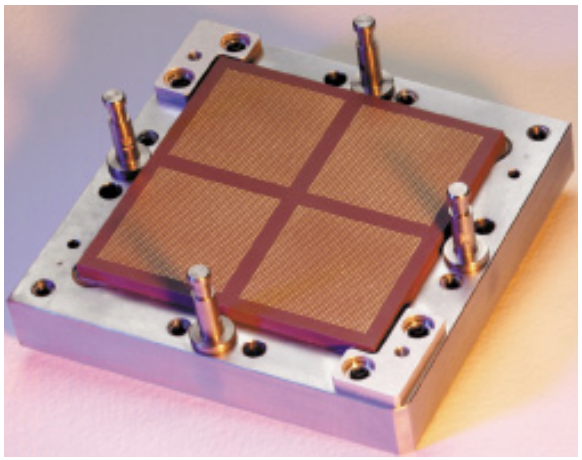


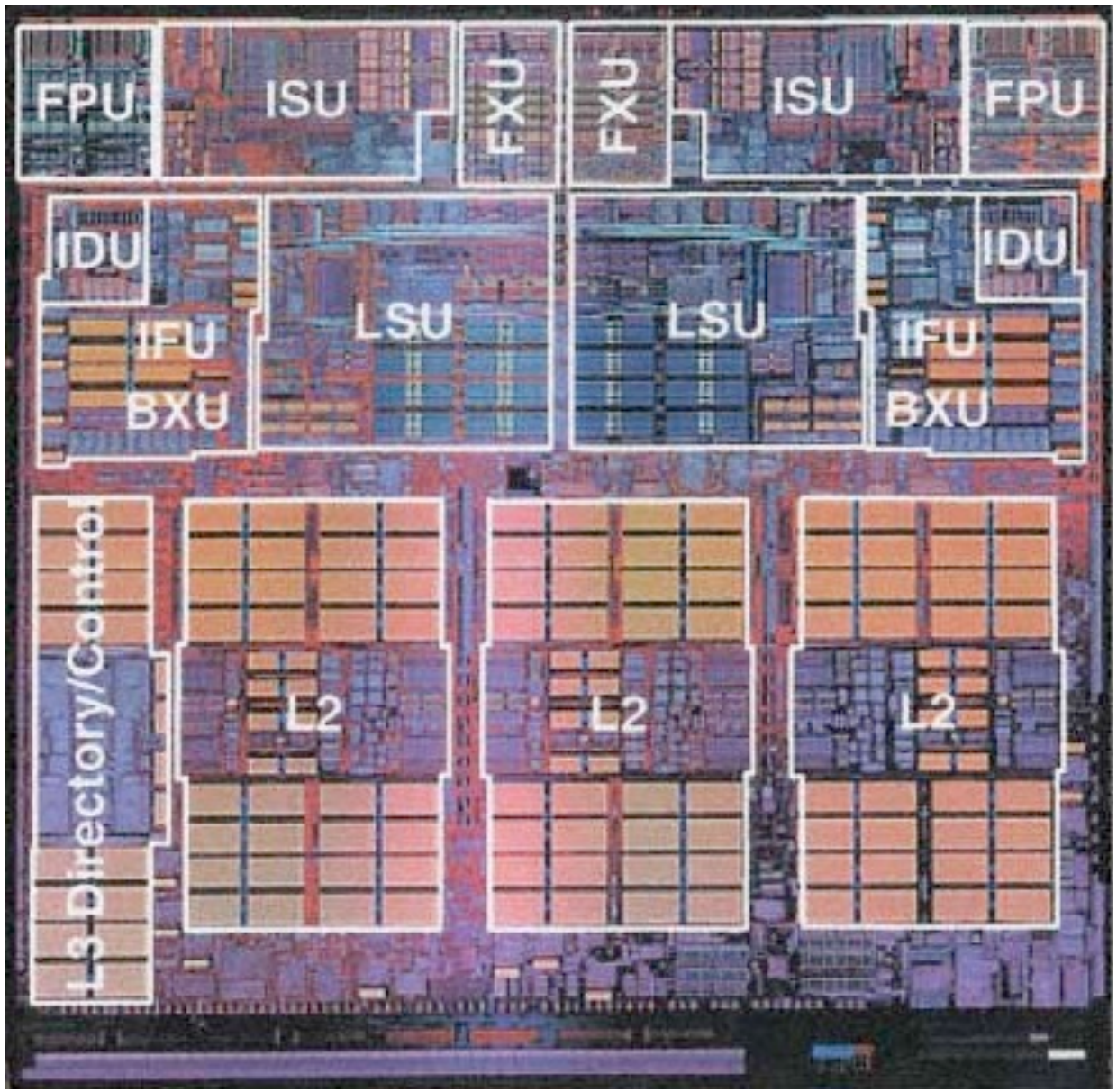
Figure 3: **I_{CCINTQ} vs. Junction Temperature with Increase Relative to 25°C**

IBM Power 4: How does die heat up?



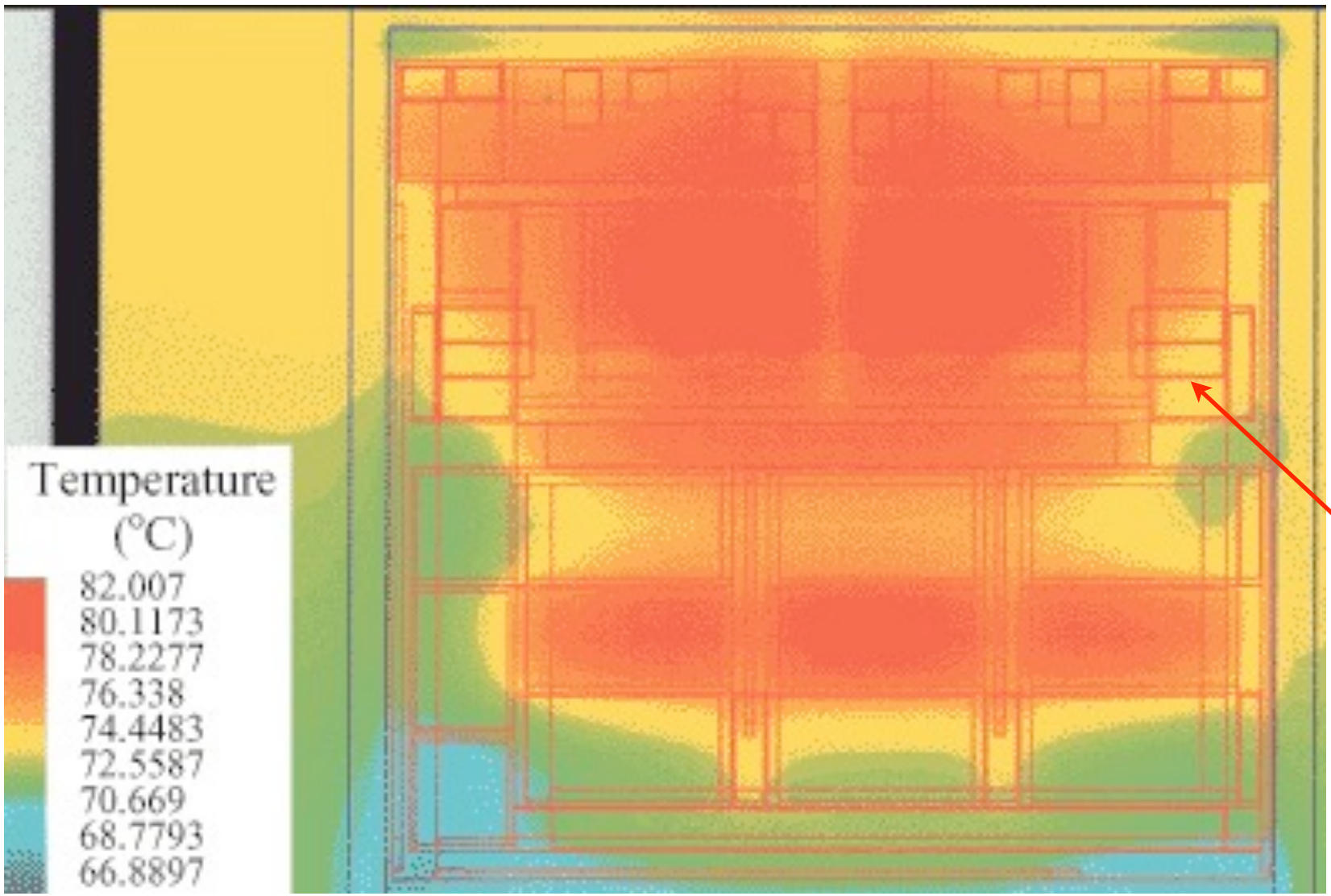
4 dies on a multi-chip module

2 CPUs per die



UC Regents Fall 2011 © UCB

115 Watts: Concentrated in "hot spots"



Hot spots

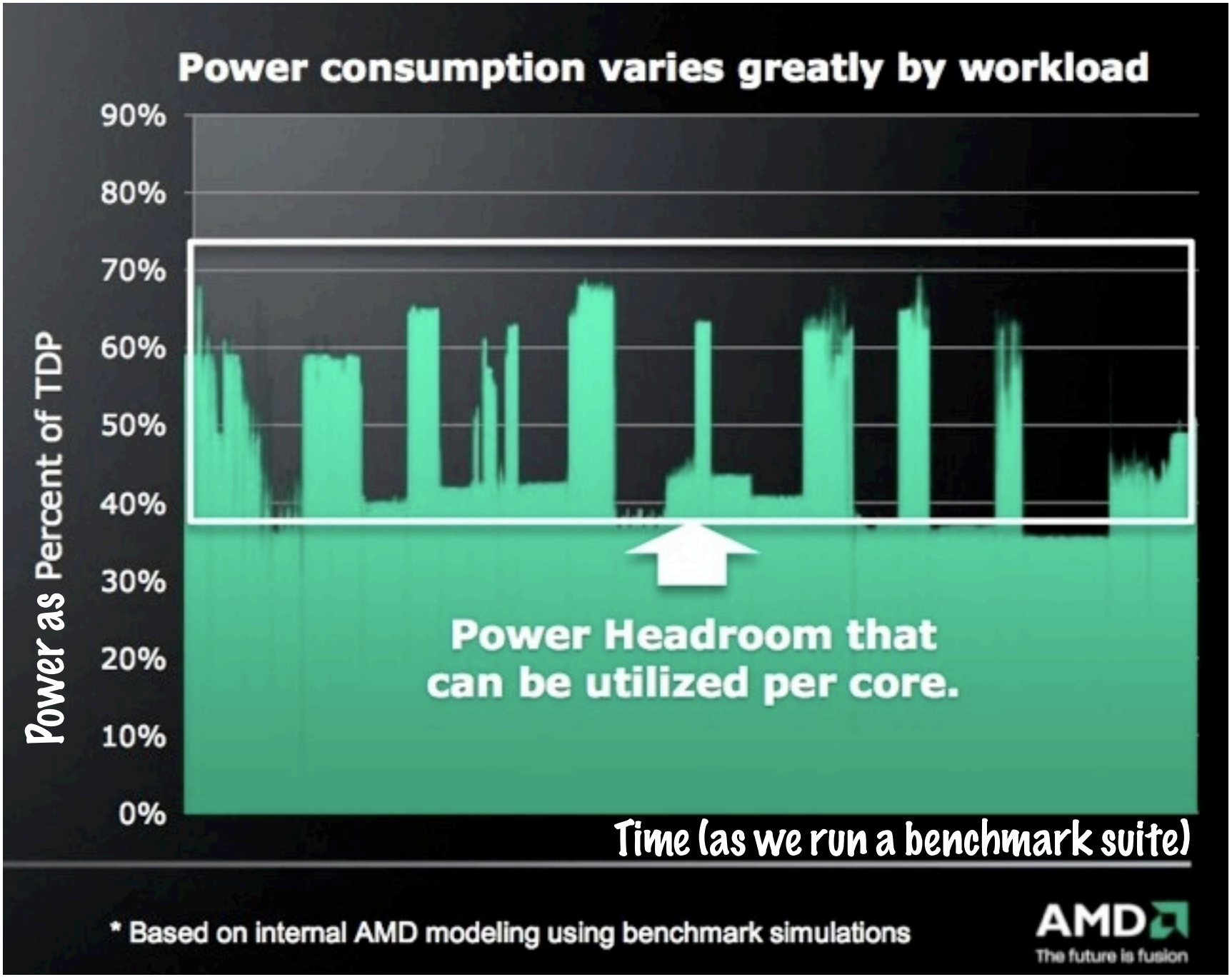
Fixed point units

Cache logic

66.8 C == 152 F 82 C == 179.6

UC Regents Fall 2011 © UCB

Idea: Monitor temperature, servo clock speed



TDP = Thermal Design Point

UC Regents Fall 2011 © UCB

Intel implementation: “Turbo Boost”

Processor	Intel Core i7-870 2.93 GHz				Intel Core i7-860 2.80 GHz				Intel Core i5-750 2.66 GHz			
Processor Cores	4				4				4			
Active Cores	1C	2C	3C	4C	1C	2C	3C	4C	1C	2C	3C	4C
Maximum Intel® Turbo Boost Technology Bin Upside	5	4	2	2	5	4	1	1	4	4	1	1
Maximum Intel® Turbo Boost Technology Frequency	3.6	3.46	3.2	3.2	3.46	3.33	2.93	2.93	3.2	3.2	2.8	2.8

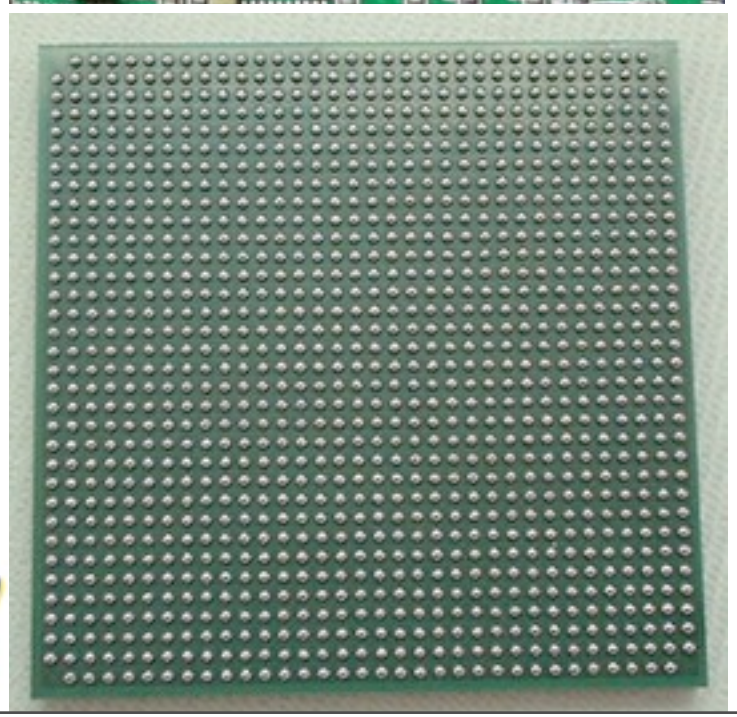
Bin Size = 133 MHz

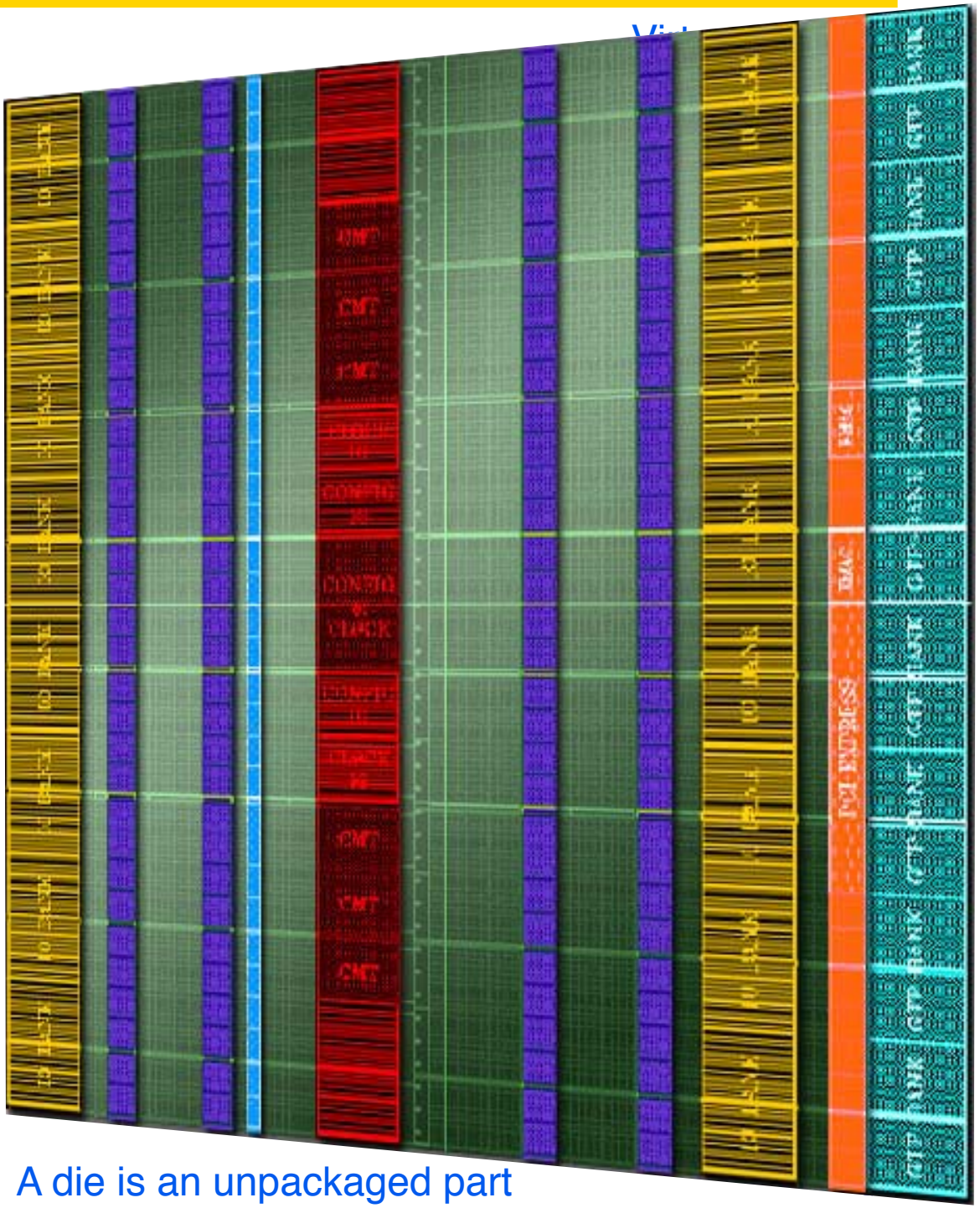
UC Regents Fall 2011 © UCB

Power and Energy in FPGAs

UC Regents Fall 2011 © UCB

FPGA: Xilinx Virtex-5 XC5VLX110T





A die is an unpackaged part

Colors represent different types of resources:

Logic

Block RAM

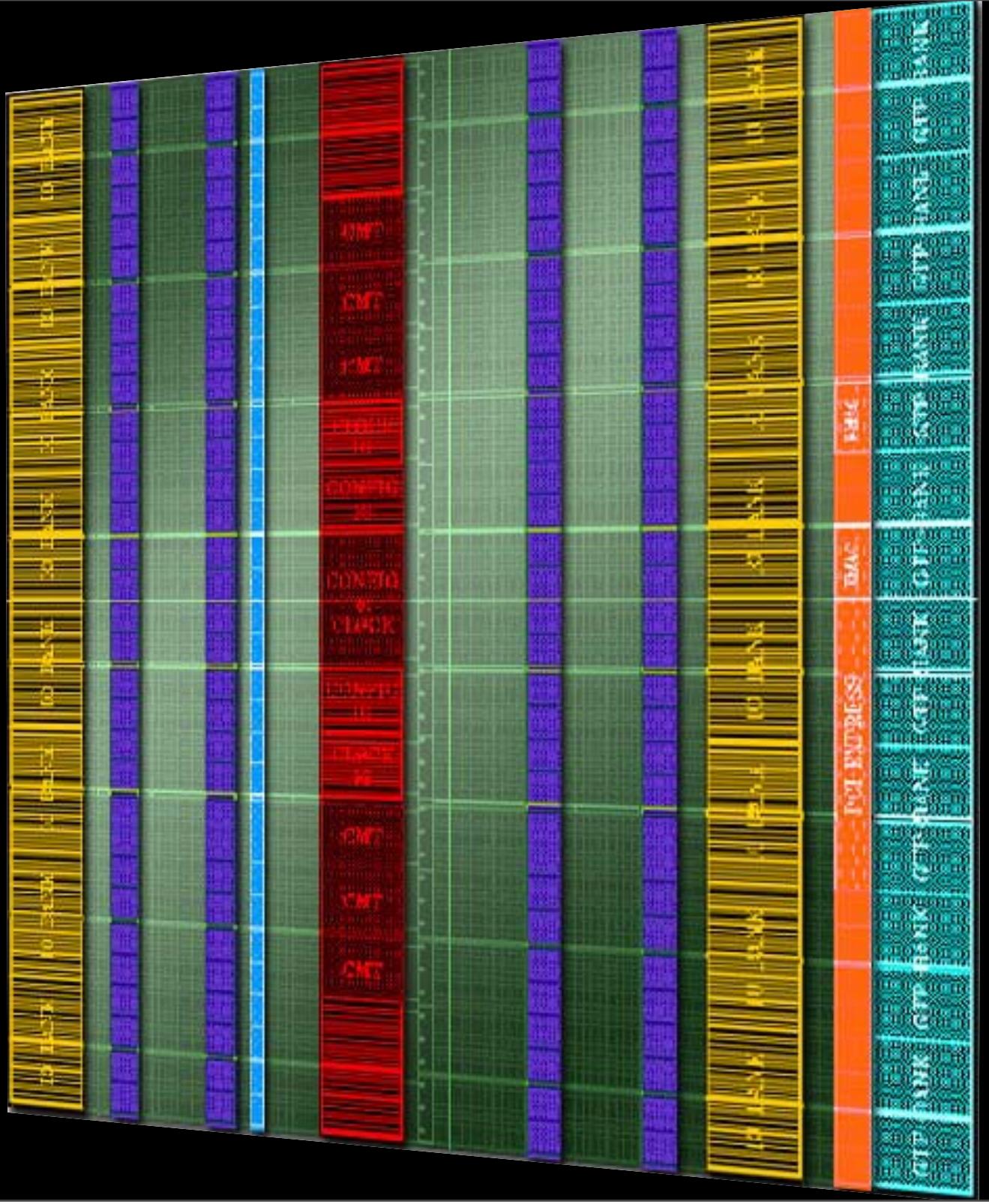
DSP (ALUs)

Clocking

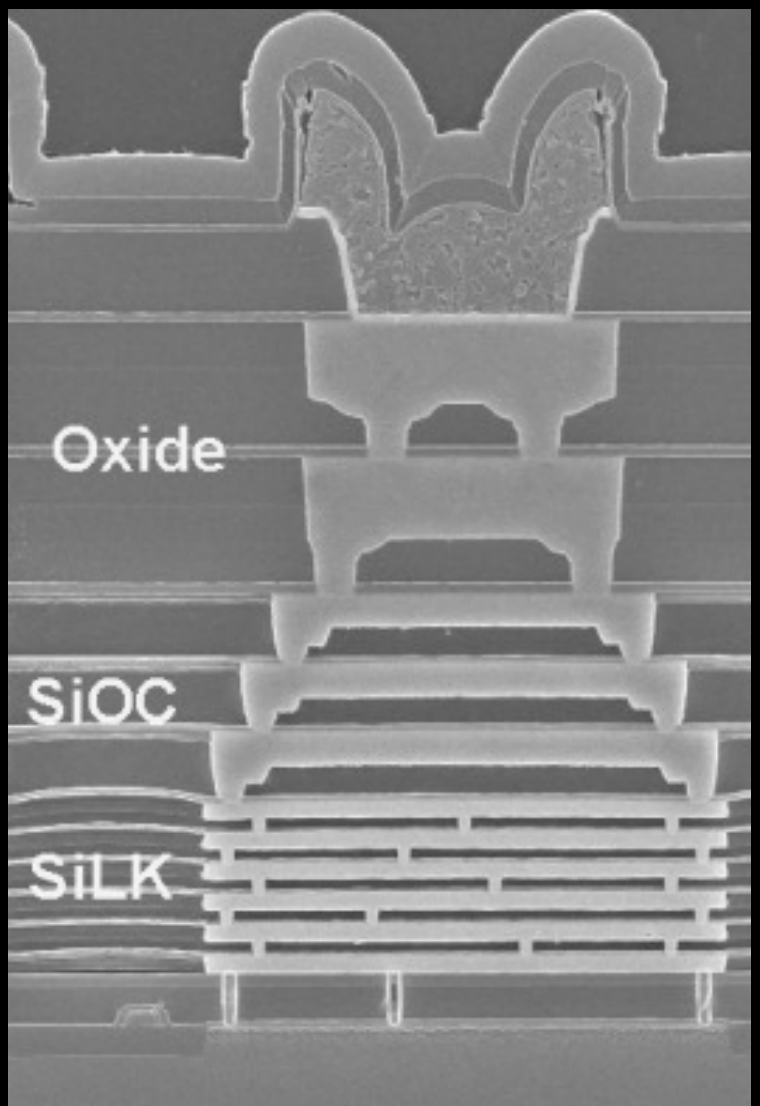
I/O

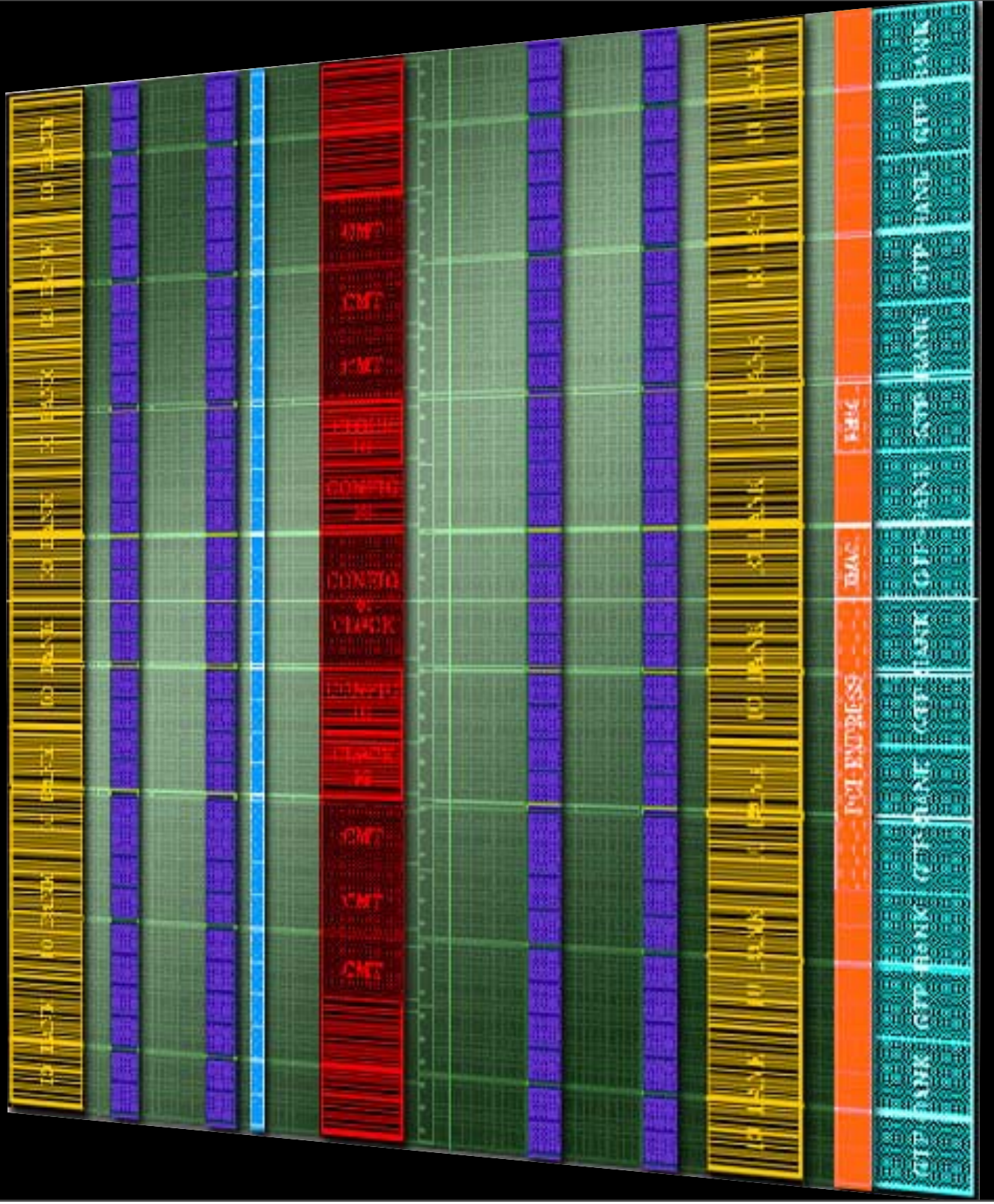
Serial I/O + PCI

A routing fabric runs throughout the chip to wire everything together.

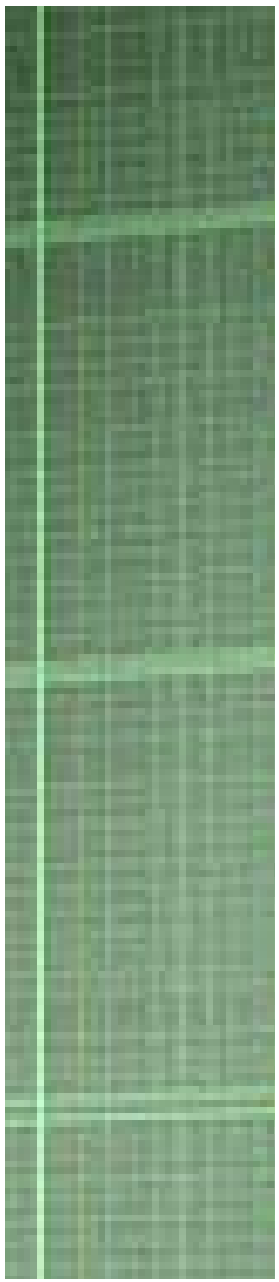


Routing fabric requires many interconnect layers.

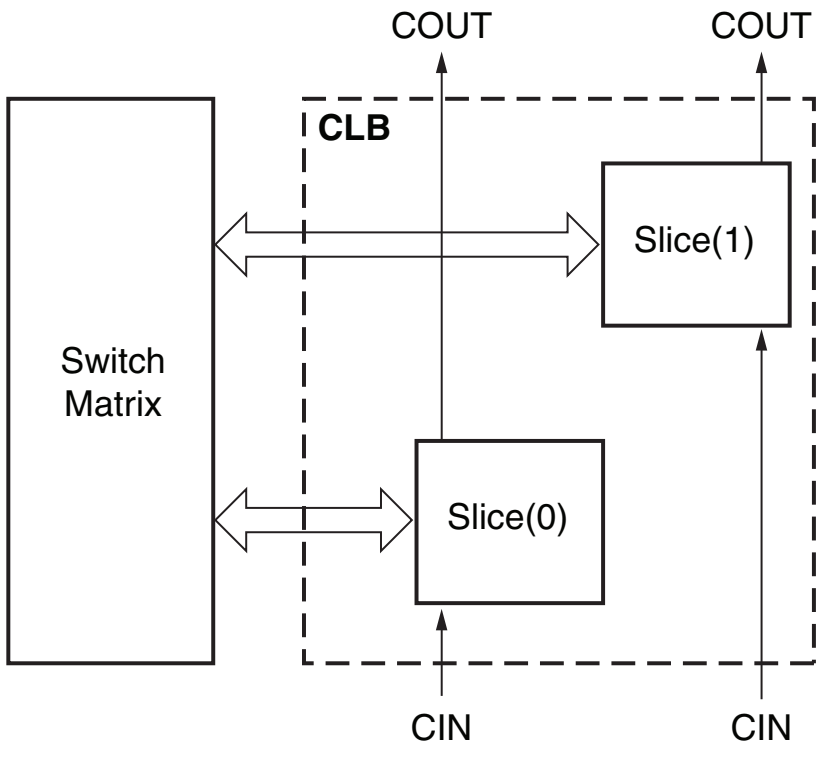




Power Issue #1: Switching Fabric 'C'



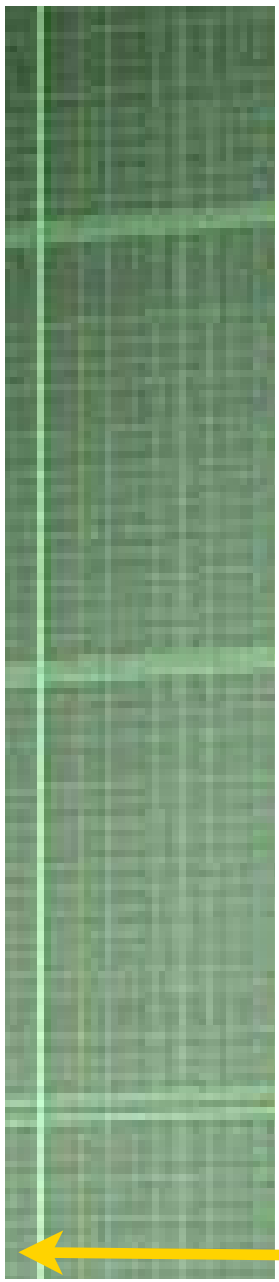
Slices define regular connections to the switching fabric, and to slices in CLBs above and below it on the die.



The LX110T has 17,280 slices.

UC Regents Fall 2011 © UCB

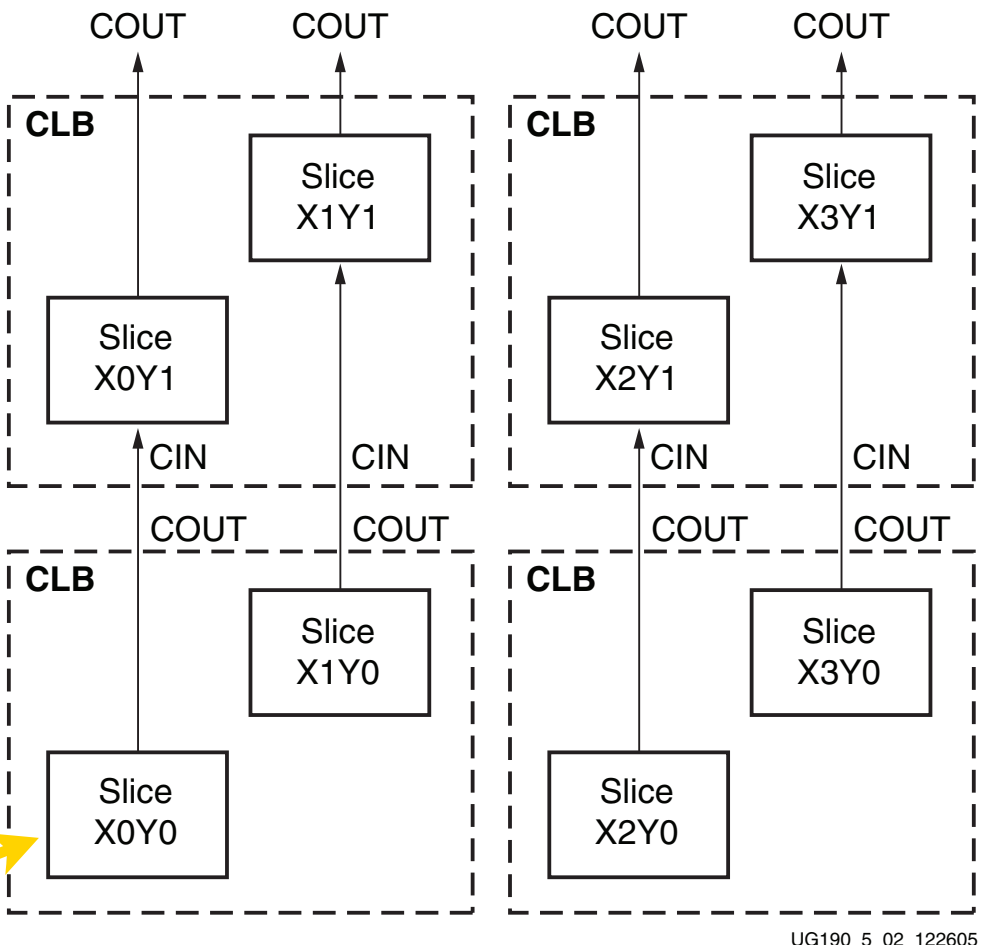
How slices appear on the die ...



X0, X2, ... are lower CLB slices.

X1, X3, ... are upper CLB slices.

Y0, Y1, ... are CLB column positions.

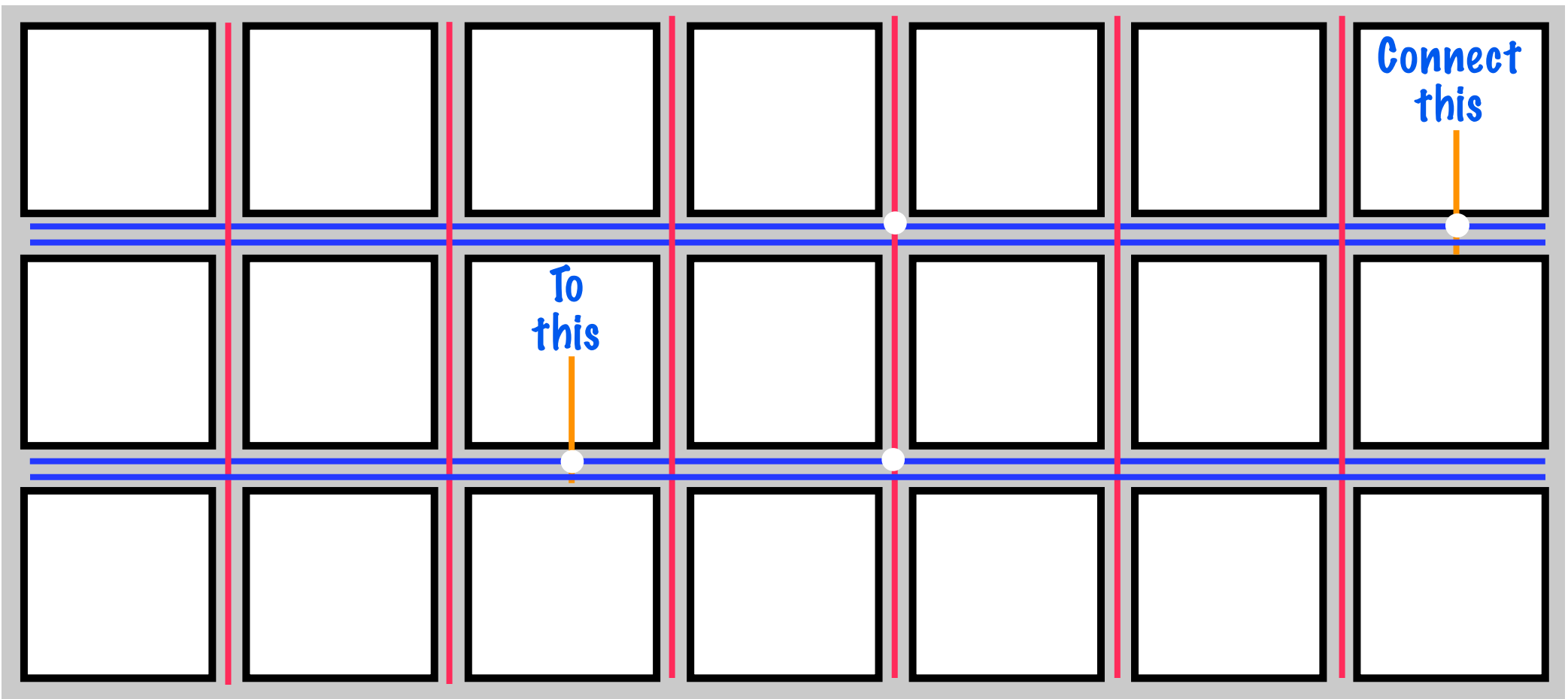


Lower-left corner of the die.

UC Regents Fall 2011 © UCB

Simple model of FPGA interconnect ...

Why 'C' is so big: (1) each white dot is a transistor switch
(2) path may not be shortest length (3) all wires are too long!

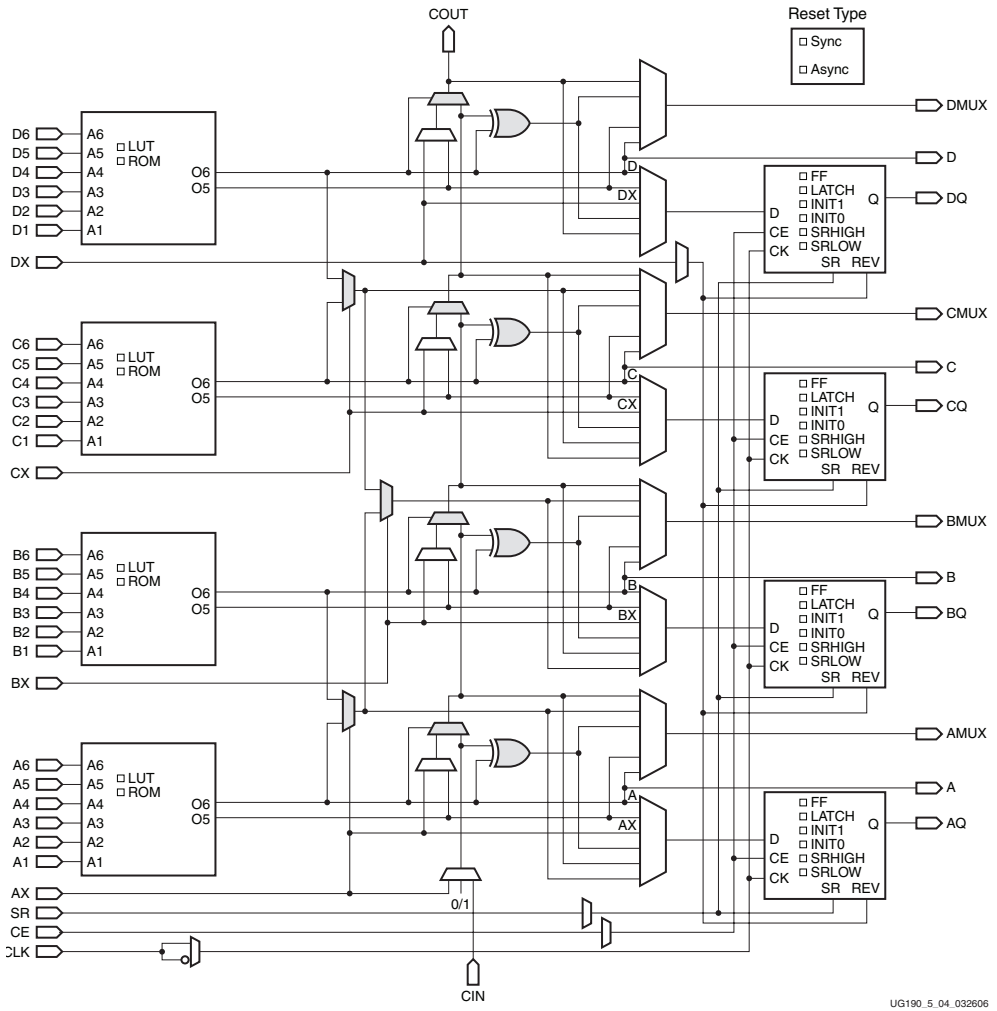


In a non-FPGA (ASIC) chip, this wire may have 10 times less capacitance, and thus use 10 times less power on each flip !

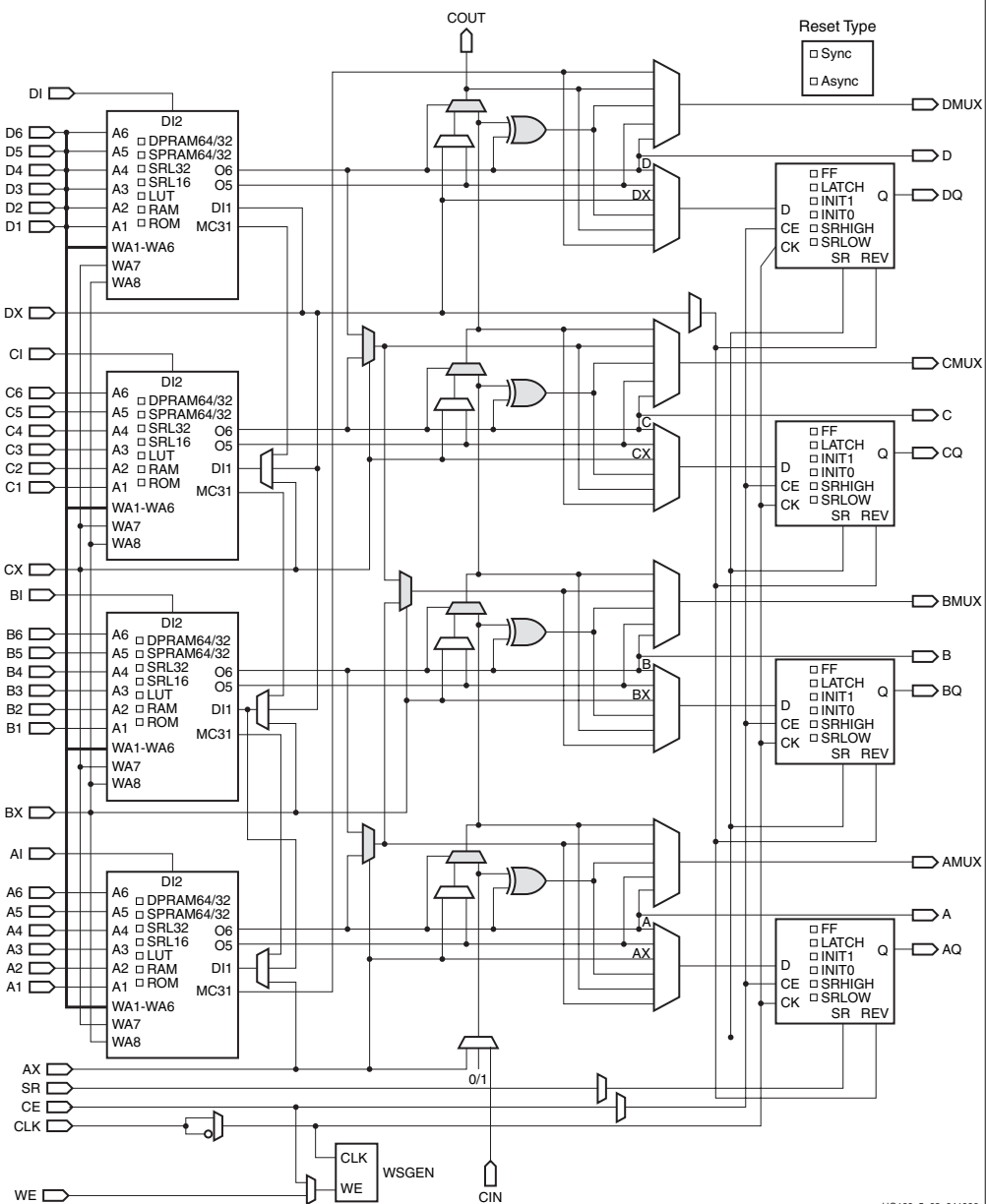
UC Regents Fall 2011 © UCB

Power Issue #2: Unused slice logic.

SLICEL



SLICEM



Unused logic that cannot be turned off wastes static power.

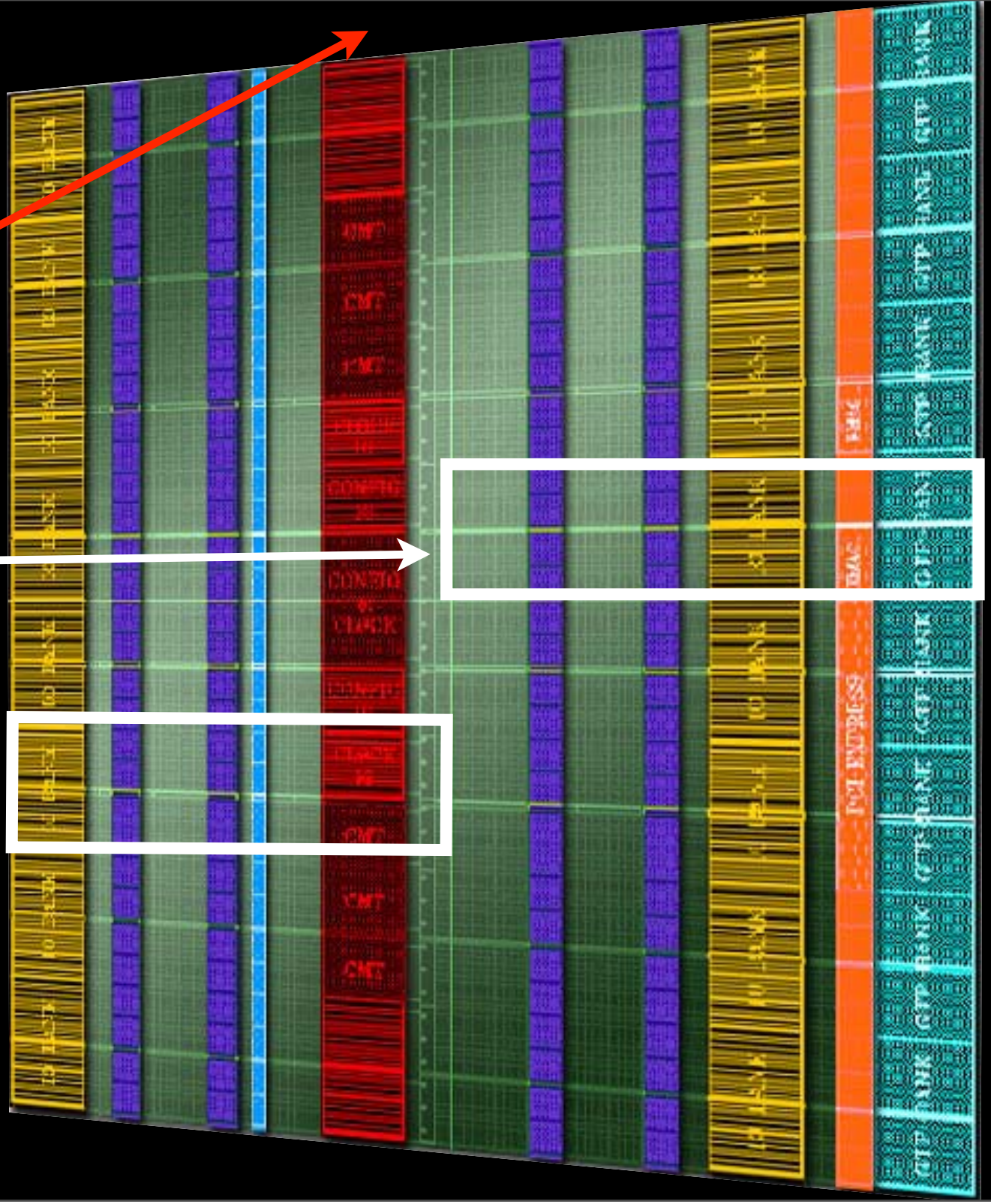
UC Regents Fall 2011 © UCB

Power issue #3:
Clock power.

32 global clock wires go down the red column.

Any 10 may be sent to a clock region.

Much of the clock wire is a clock wire to nowhere.
Big power waste.

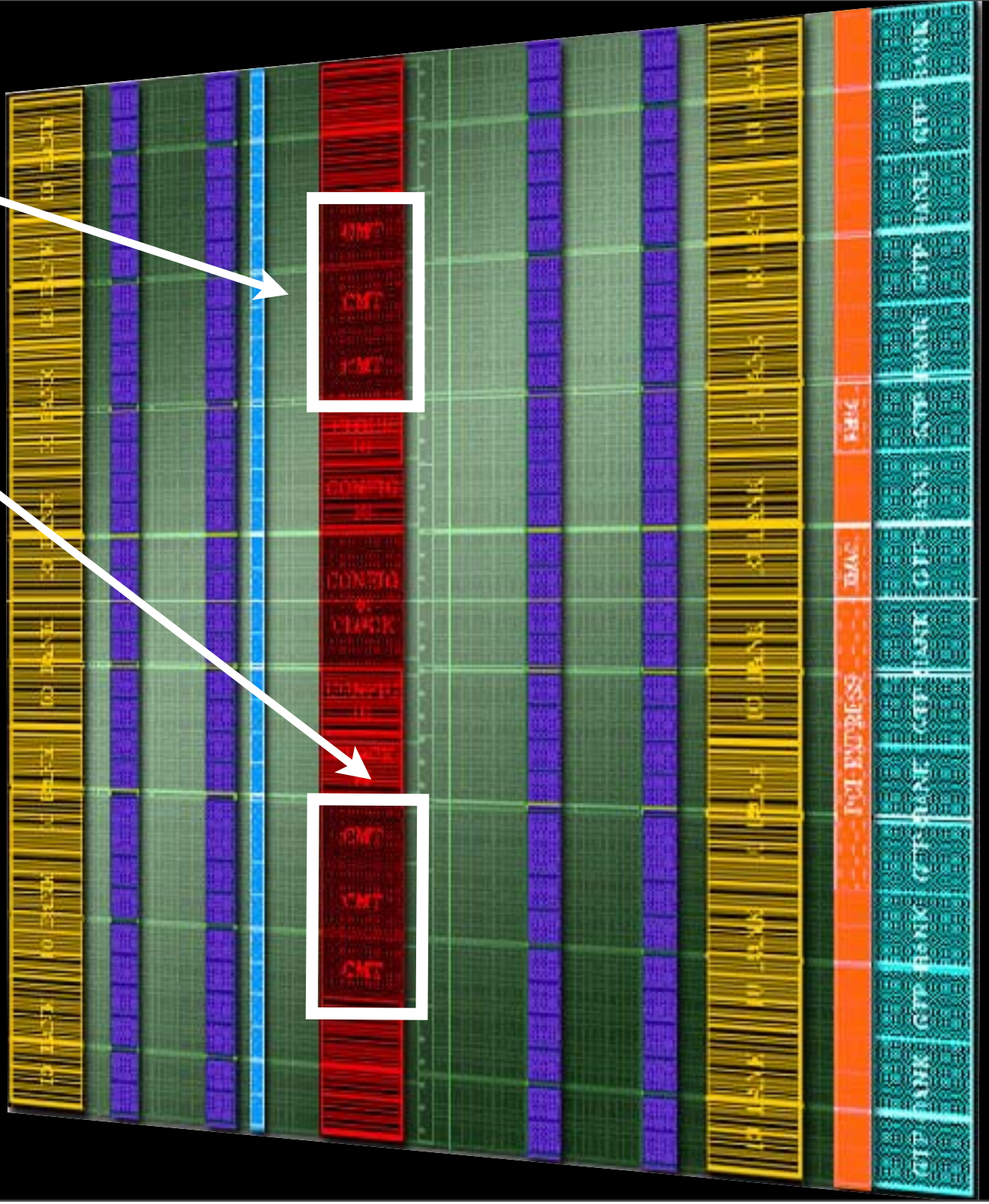


LX110T:

12
Digital Clock
Managers (DCM)

6 Phase Locked
Loops (PLL)

A non-FPGA
(ASIC) chip
can address the
problems DCMs
and PLLs solve
using less power.



Power Issue #4: CAD speed tradeoffs ...

An example design from a Xilinx "white paper" ...

CAD set for high speed.

1 x 32.5k x 36-bit	Area	Fmax MHz	Total Dynamic Power (mW)	Dynamic RAM Power (mW)
Speed Implementation	70 Slices 73 RAMB16	139	135.6	99.6
XST 9.2i-power	751 Slices 65 RAMB16	123	48	2.4

CAD set for low power.

A small drop in clock speed, but a big power win. But, design size explodes, and the tools run slow.

UC Regents Fall 2011 © UCB

Upcoming: The Project Lecture

Wed Sep 21	KA	Lecture 7: Project details.
Thu Sep 22	BZ	Discussion 4: Debugging, ALUs

UC Regents Fall 2011 © UCB