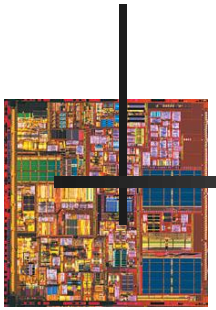
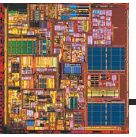


Design for Low Power

A Long Time Ago



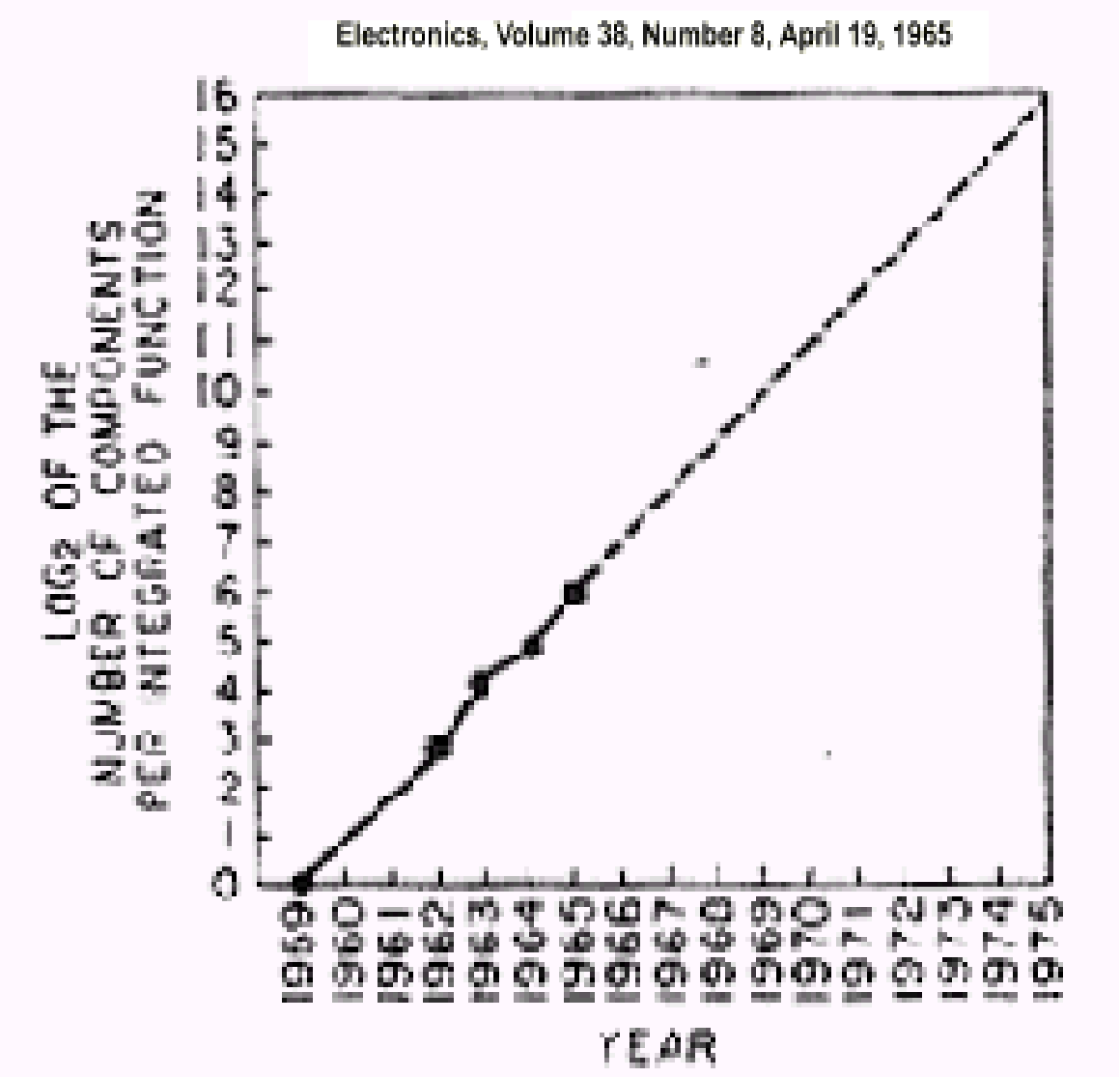
In a building far away

A man made a prediction

On surprisingly little data

That has defined an industry

Moore's Law


Electronics, Volume 38, Number 8, April 19, 1965
LOG2 OF THE NUMBER OF COMPONENTS PER INTEGRATED FUNCTION
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
1959 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
YEAR

Moore's Original Issues

- Design cost
- **Power dissipation**
- What to do with all the functionality possible

ftp://download.intel.com/research/silicon/moorespaper.pdf

Problems of Power Dissipation

- Continuously increasing performance demands

➔ Increasing power dissipation of technical devices

➔ Today: power dissipation is a main problem

- High **Power dissipation** leads to:

☹ Reduced time of operation

☹ Higher weight (batteries)

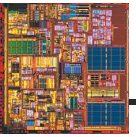
☹ Reduced mobility

☹ High efforts for cooling

☹ Increasing operational costs

☹ Reduced reliability

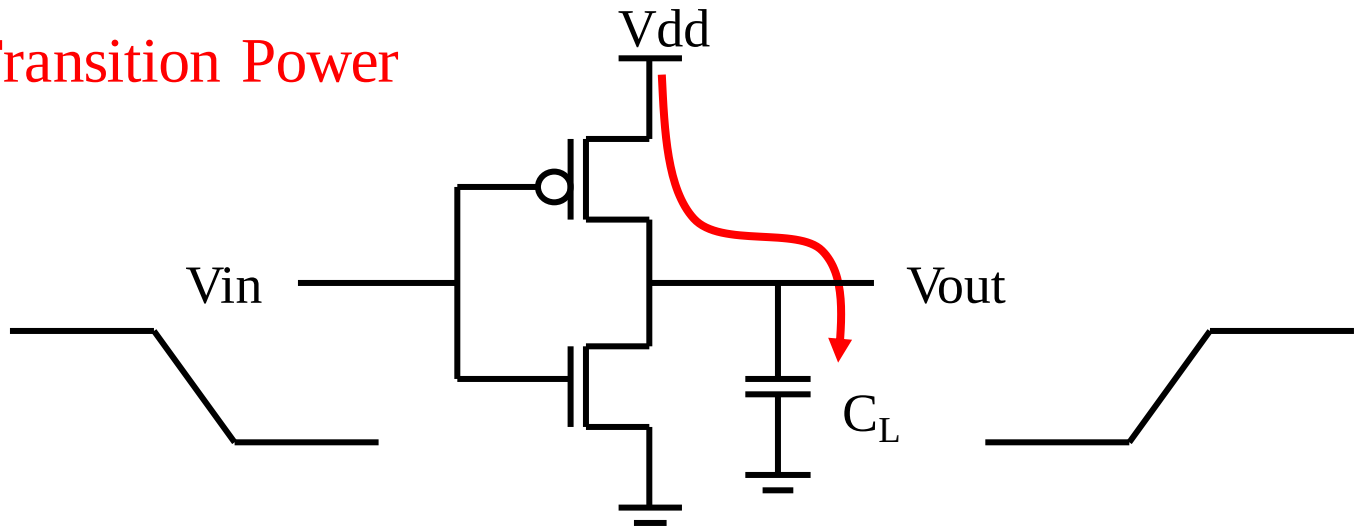
Outline



- What is the problem?
- Power Components
- Historical Context
- Solutions

Dynamic Energy Consumption

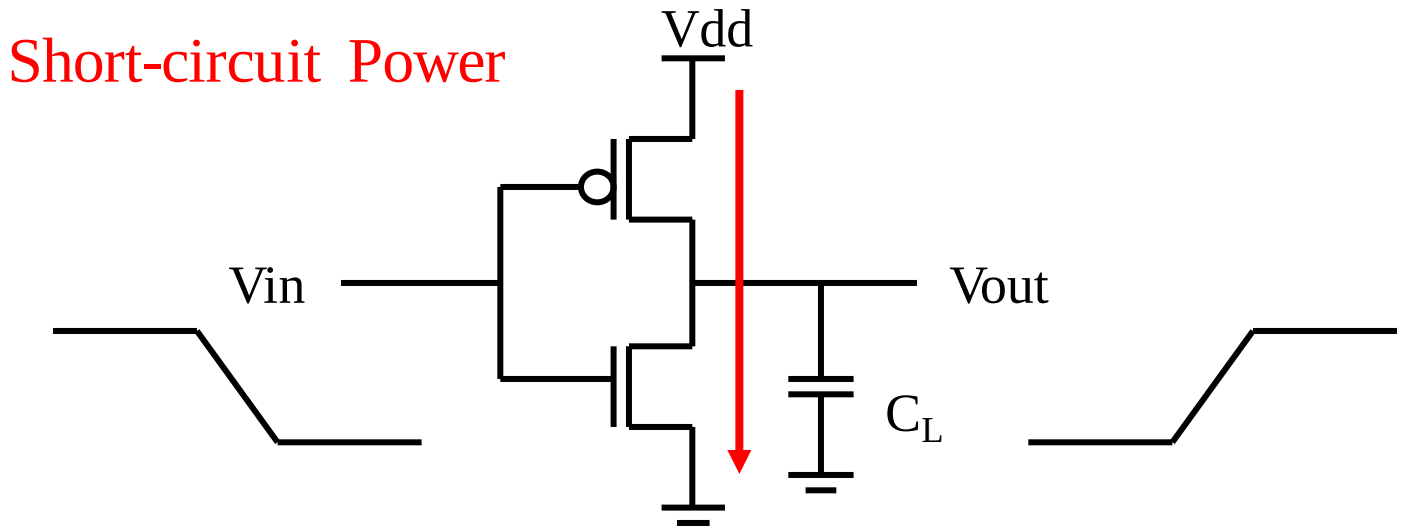
Transition Power



$$\text{Energy/transition} = C_L * V_{DD}^2 * P_{0/1 \rightarrow 1/0}$$

$$\text{Power} = C_L * V_{DD}^2 * f$$

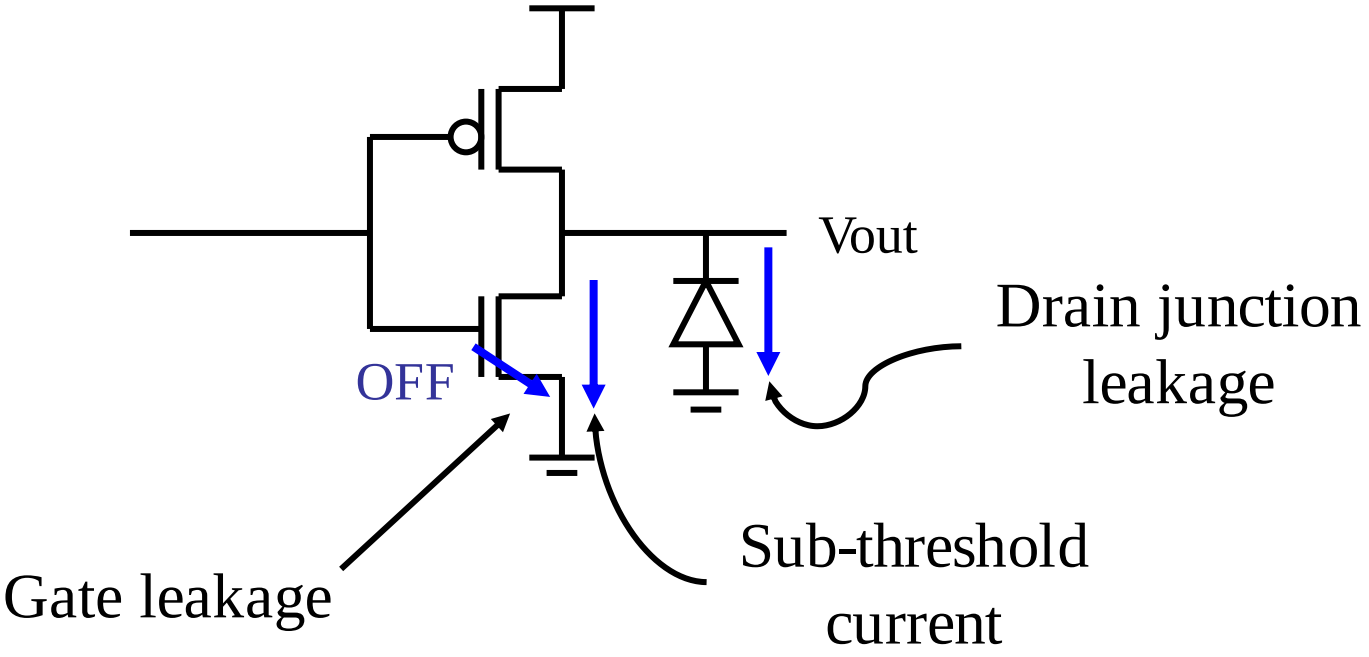
Dynamic Energy Consumption



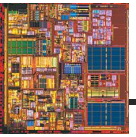
$$\text{Energy/transition} = t_{sc} * V_{DD} * I_{peak} * P_{0/1 \rightarrow 1/0}$$

$$\text{Power} = t_{sc} * V_{DD} * I_{peak} * f$$

Leakage Energy



Independent of switching



Power Equations in CMOS

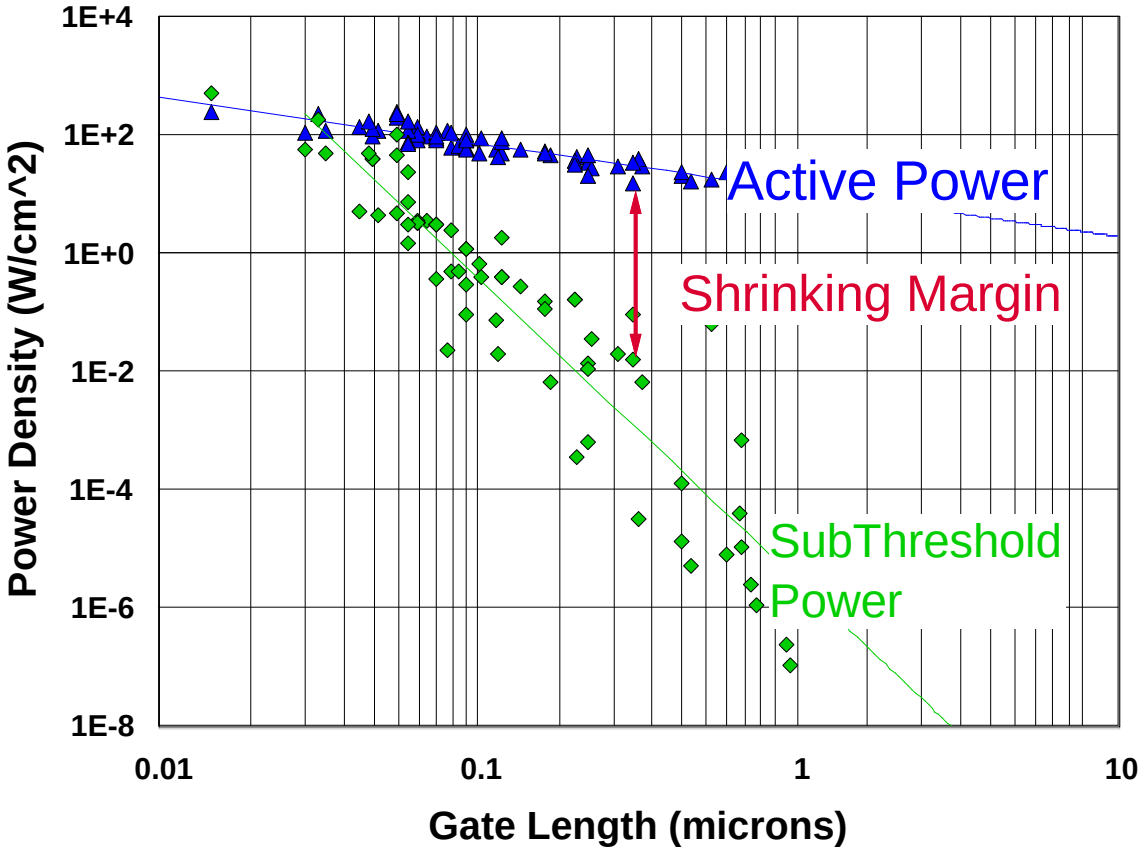
$$P = \alpha f C_L V_{DD}^2 + V_{DD} I_{peak} (P_{0\rightarrow1} + P_{1\rightarrow0}) + V_{DD} I_{leak}$$

Dynamic power (≈ 40 - 70% today and decreasing relatively)

Short-circuit power (≈ 10 % today and decreasing absolutely)

Leakage power (≈ 20 – 50 % today and increasing)

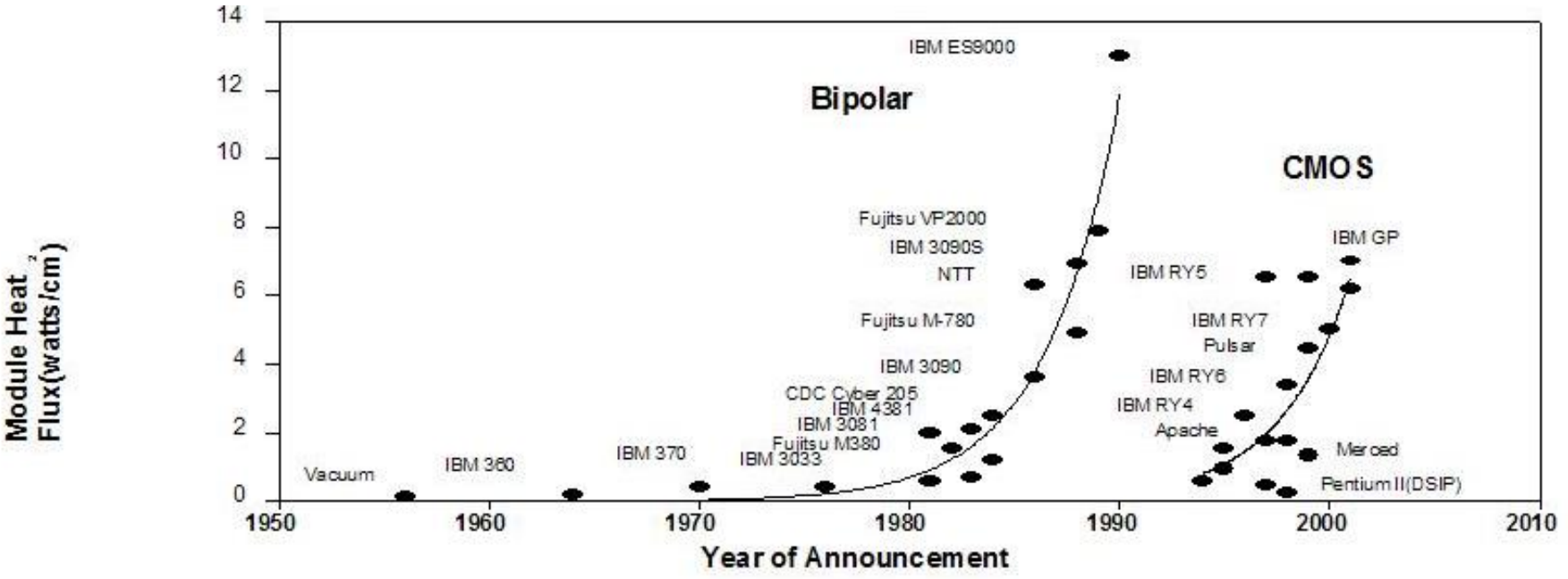
Dynamic vs Static Power



Source: Leon Stok, DAC 42©

The 80's Power Problem

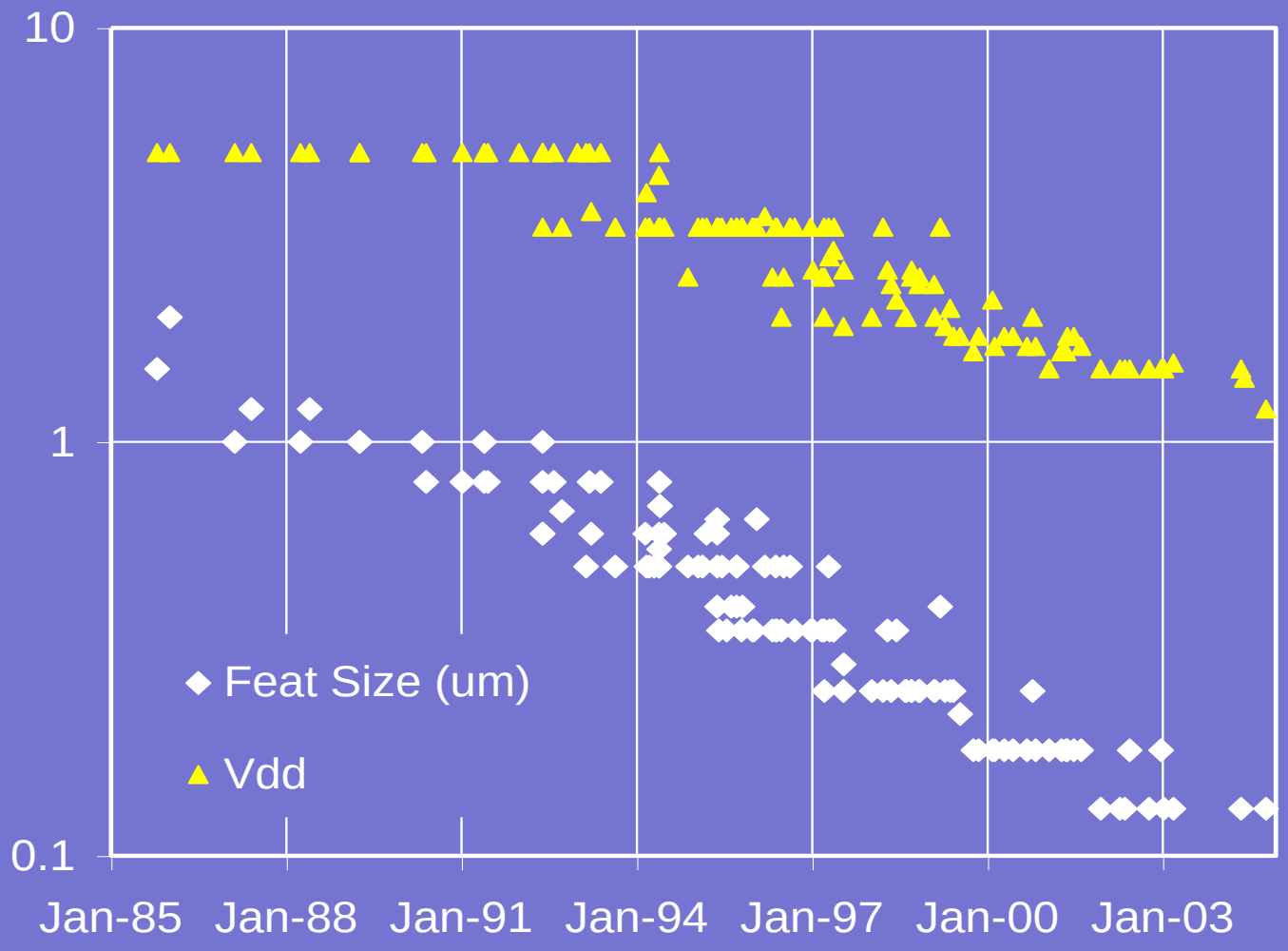
- Until mid 80s technology was mixed
  - nMOS, bipolar, some CMOS
- Supply voltage was not scaling / power was rising
  - nMOS, bipolar gates dissipate static power



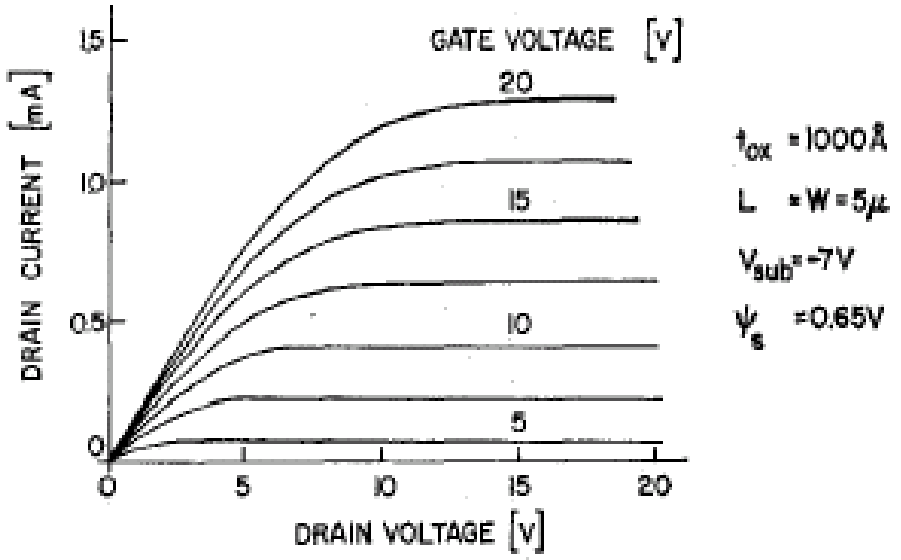
From Roger Schmidt, IBM Corp

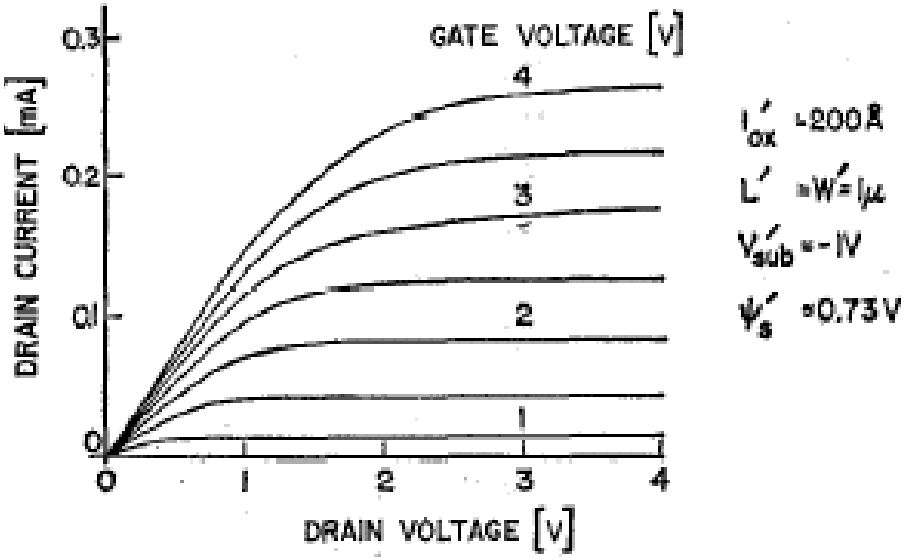
Solution: Move to CMOS

- And then scale Vdd



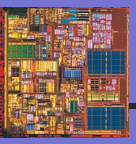
Scaling MOS Devices





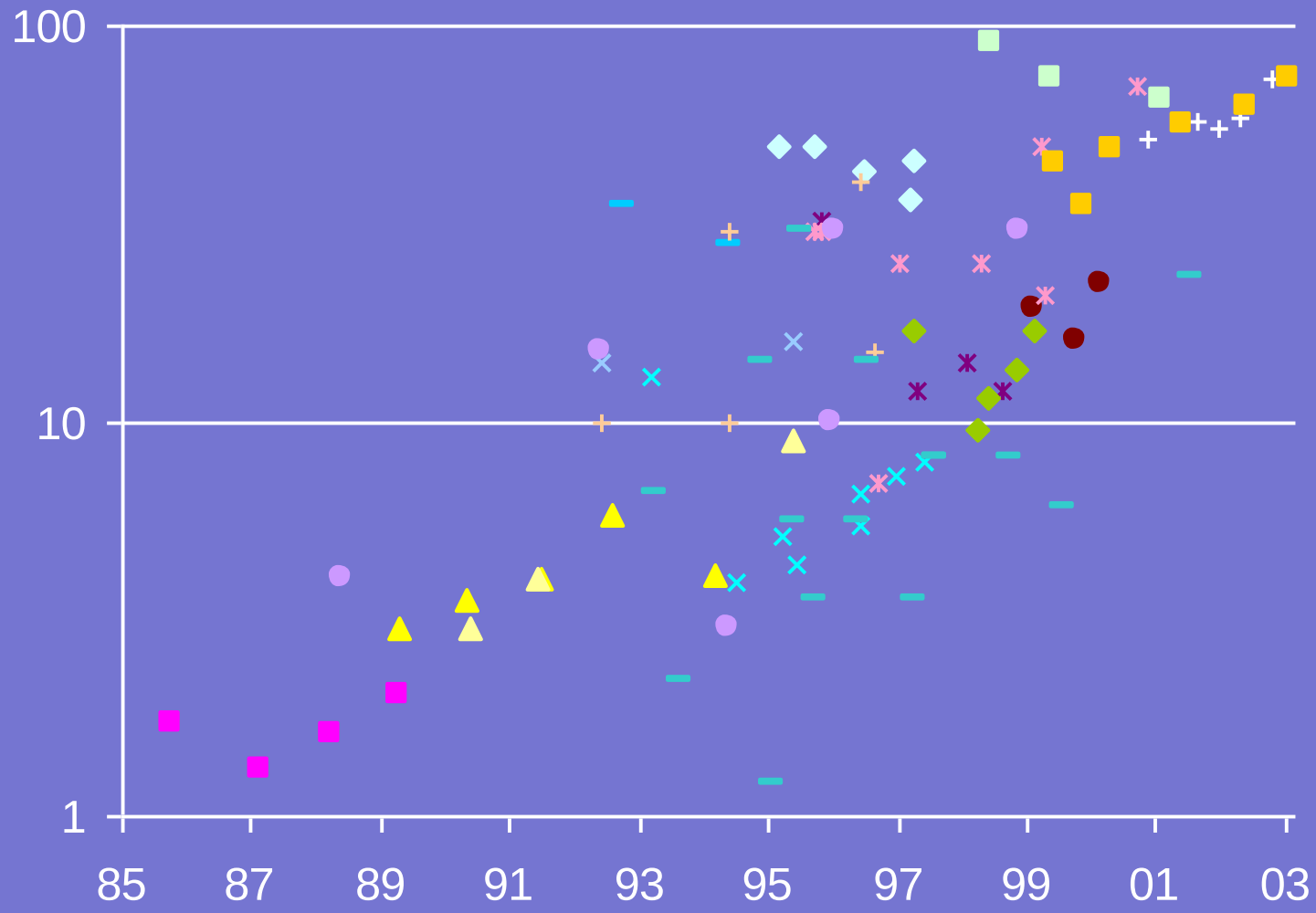
JSSC Oct 74, pg 256

- In this ideal scaling
  - V scales to αV, L scales to αL
  - So C scales to αC, i scales to αi (i/μ is stable)
  - Delay = CV/I scales as α
  - Energy = CV^2 scales as α^3

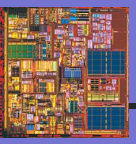


Processor Power

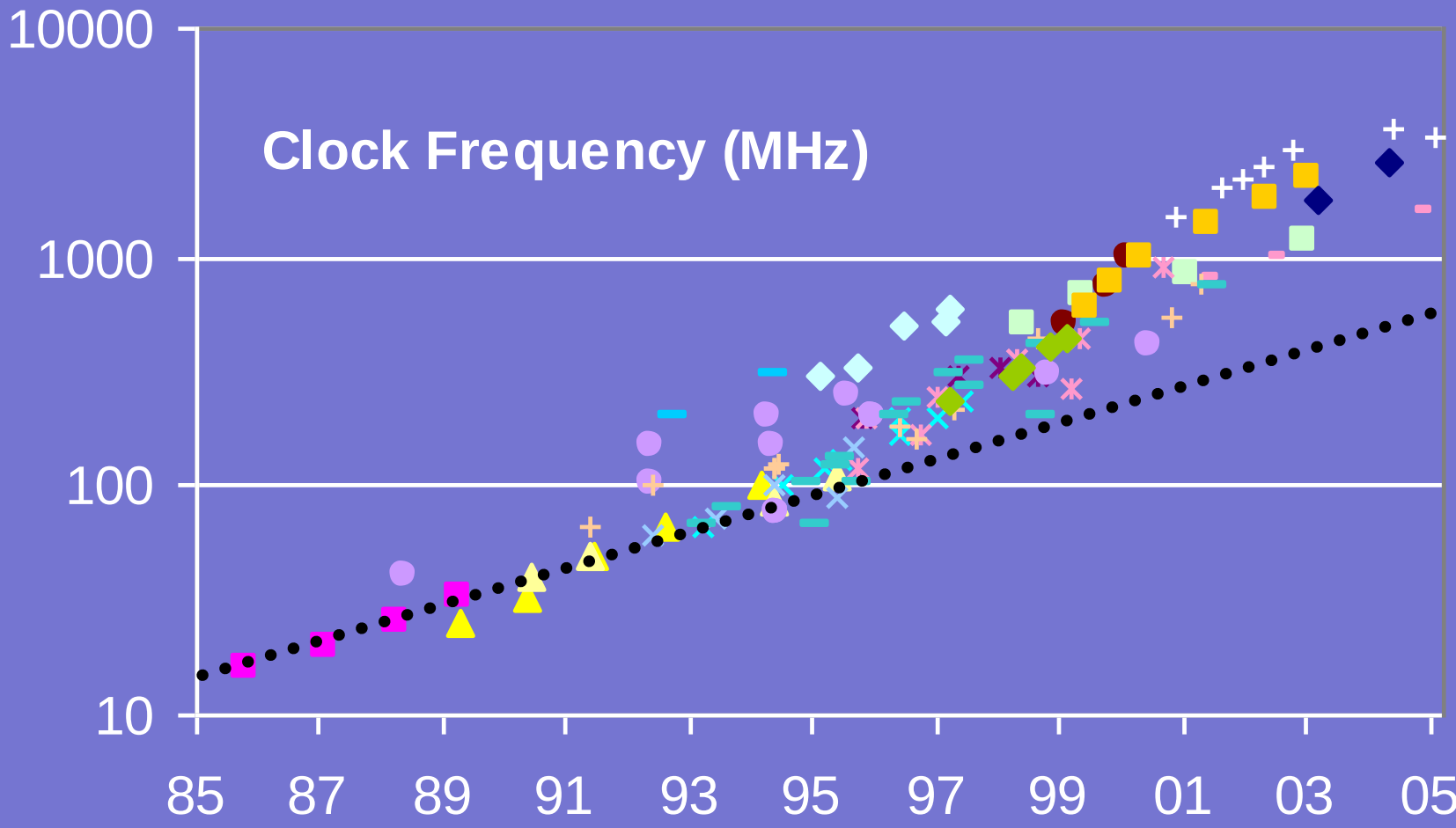
- Continued to grow, even when Vdd was scaled



Why Power Increased

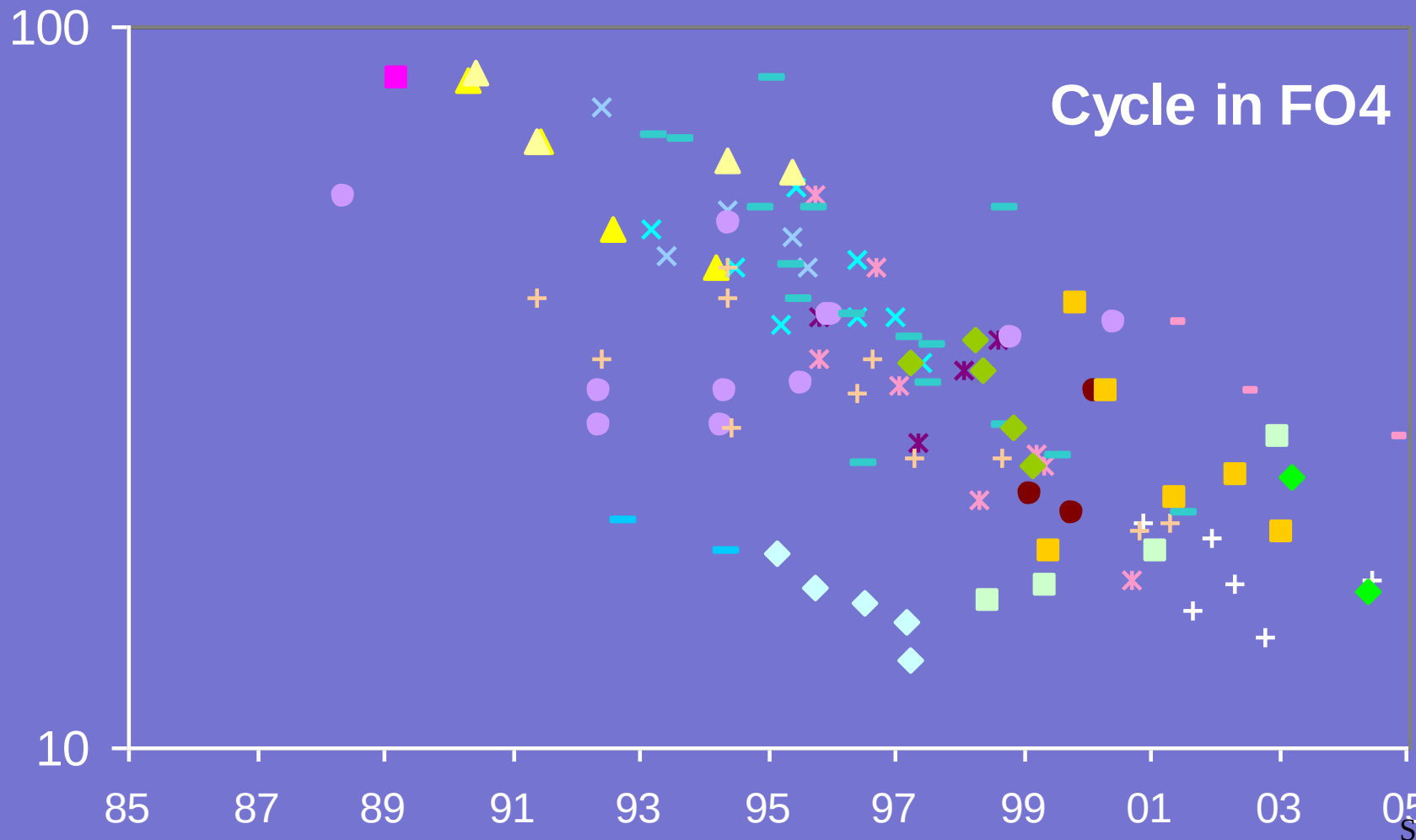


- Growing die size, fast frequency scaling



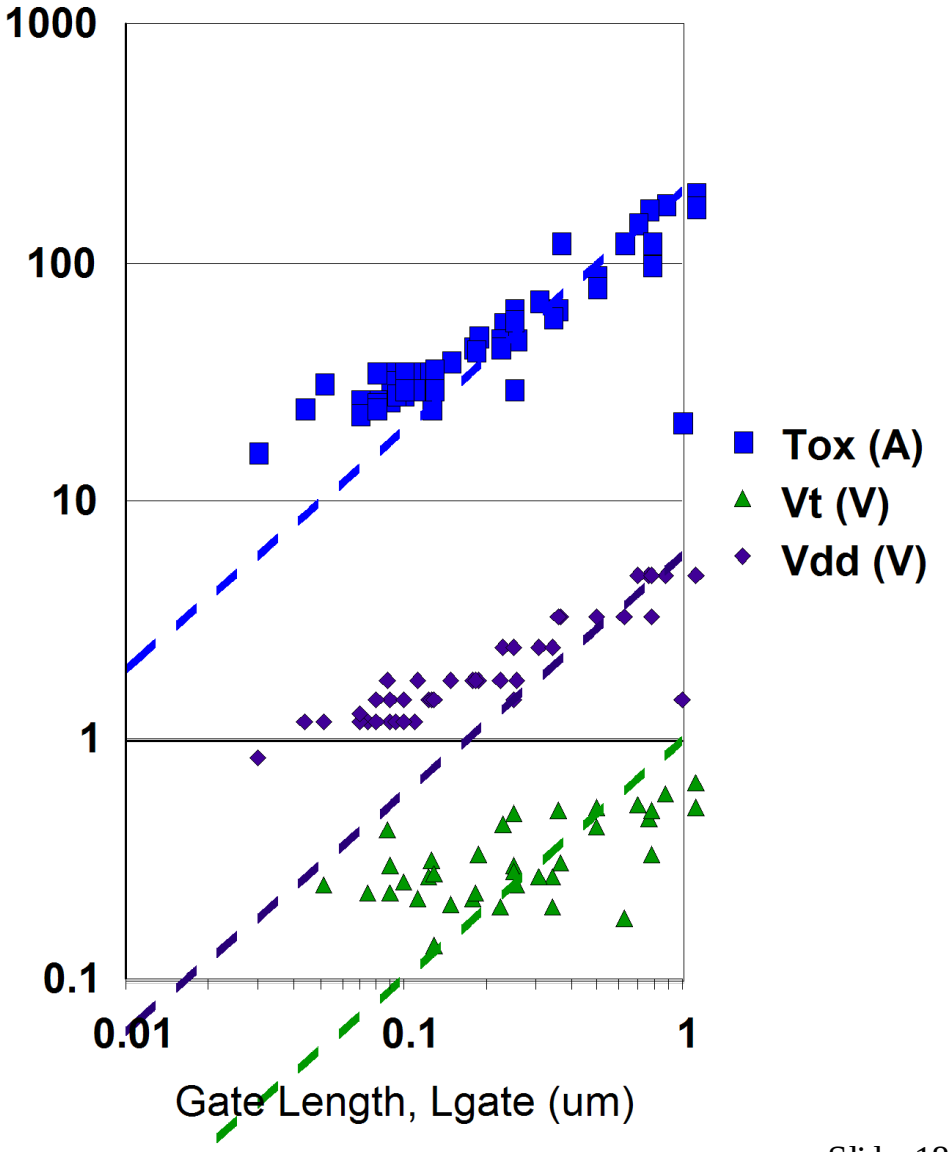
Good News

- Die growth & super frequency scaling have stopped



Bad News

- Voltage scaling has stopped as well
  - kT/q does not scale
  - Vth scaling has power consequences
- If Vdd does not scale
  - Energy scales slowly

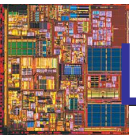


ISSCC, Feb. 2001, Keynote

Patrick P. Gelsinger
Senior Vice President
General Manager
Digital Enterprise Group
INTEL CORP.

"Ten years from now, microprocessors will run at 10GHz to 30GHz and be capable of processing 1 trillion operations per second – about the same number of calculations that the world's fastest supercomputer can perform now.

"Unfortunately, if nothing changes these chips will produce as much heat, for their proportional size, as a nuclear reactor. . . ."

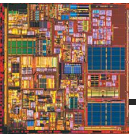


Low Power Design Techniques

Three main classes of methods to reduce energy:

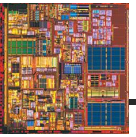
- Cheating
  - Reducing the performance of the design
- Reducing waste
  - Stop using energy for stuff that does not produce results
  - Stop waiting for stuff that you don't need (parallelism)
- Problem reformulation
  - Reduce work (less energy and less delay)

Cheating



- Many low-power papers talk only about energy
  - Don't consider performance
- Reducing performance can always reduce energy
  - But there are many ways to reduce performance
- Good technique must lower the optimal curve
  - "Sensitivity" of technique
    - Must be better than current curve
    - This depends on location on the curve

Reducing Energy Waste



- Clock gating
  - If a section is idle, remove clock
  - Removes clock power
  - Prevents any internal node from transitioning
- Create system power states
  - Turn on subsystems only when they are needed
  - Can have different "off" states
    - Power vs. wakeup time
    - Disk (do you stop it from spinning?)

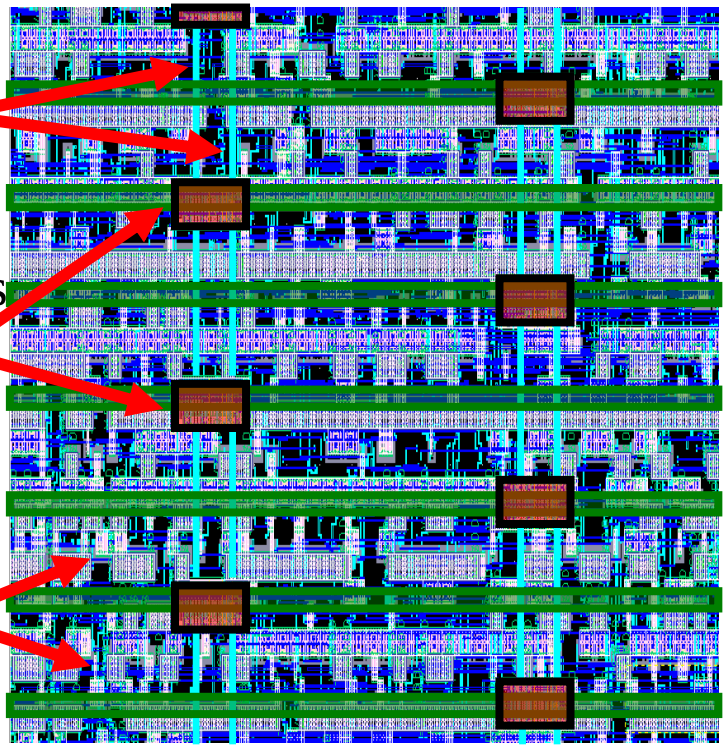
Embedded Power Gating

- Since transistors still leak when power is off

Power Switch Control Signals

Embedded Power Switches

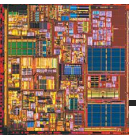
Rows of Standard Cells



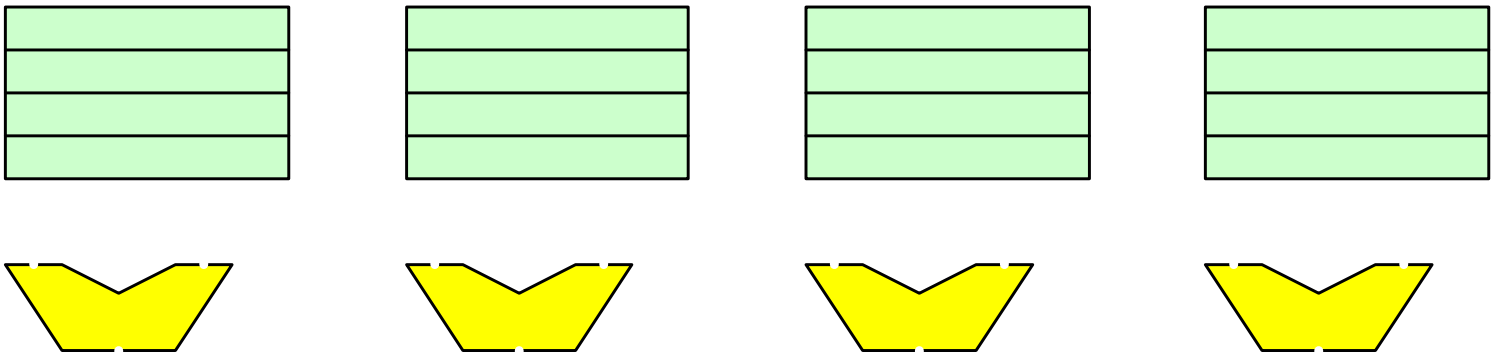
- Can reduce leakage
  - 250x reported
- But costs
  - Performance
  - Drop in Vdd, Gnd

Royannez, et al, 90nm Low Leakage SoC Design Techniques for Wireless Applications, ISSCC 2005

Parallelism

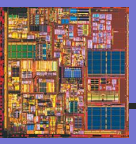


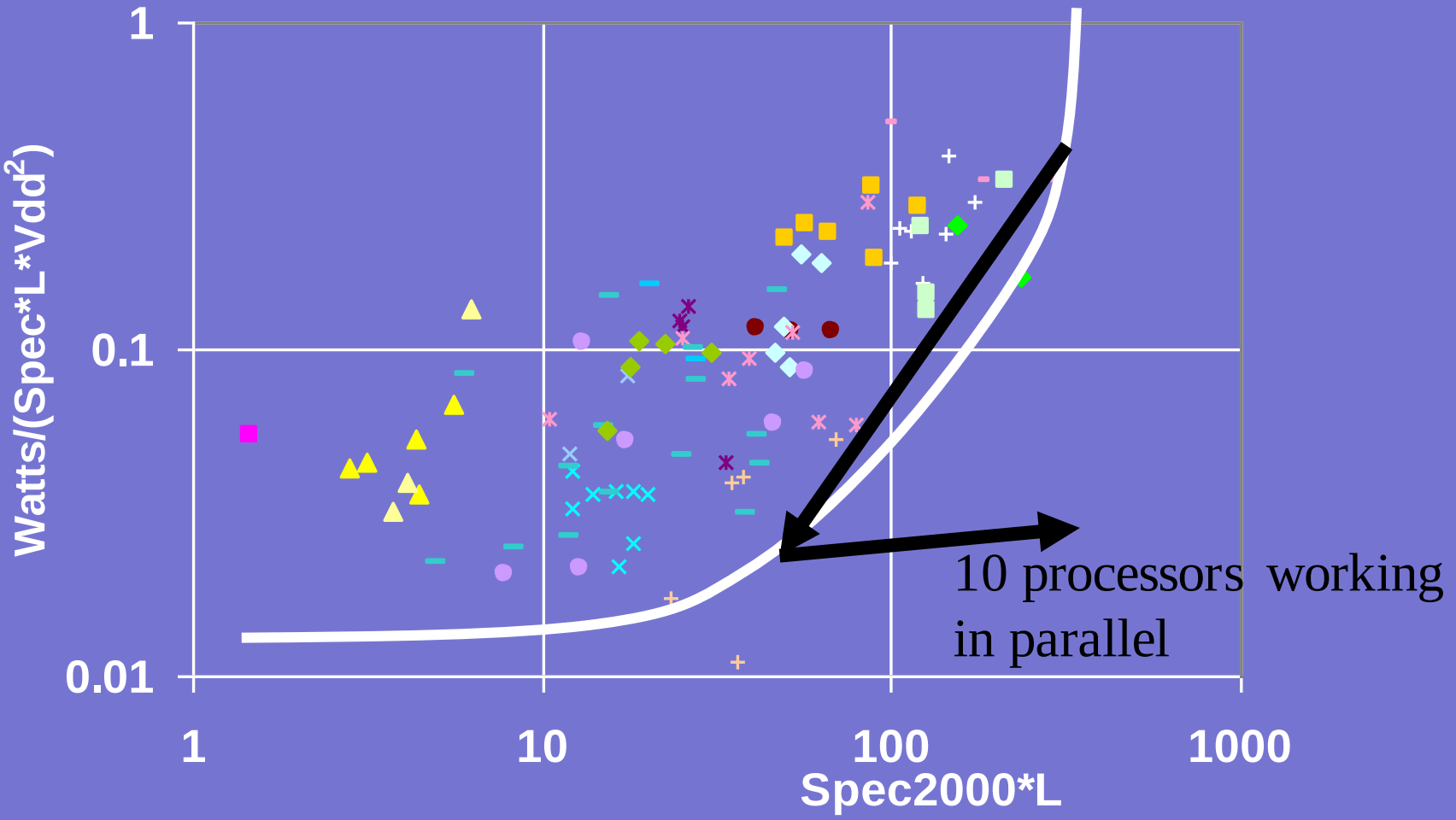
- If the application has data parallelism



- Parallelism is a way to improve performance
  - With low additional energy cost

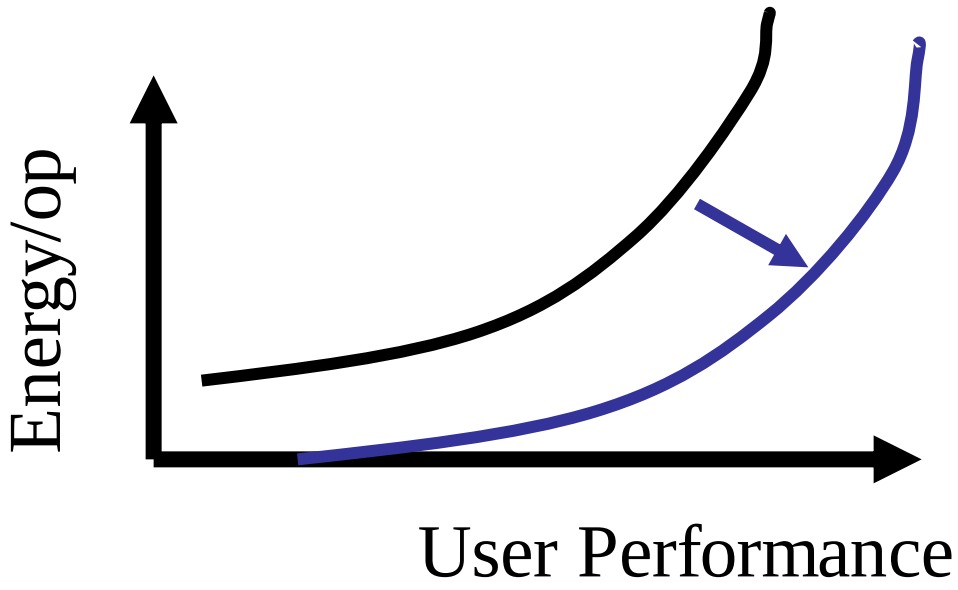
Existing Processors



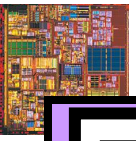


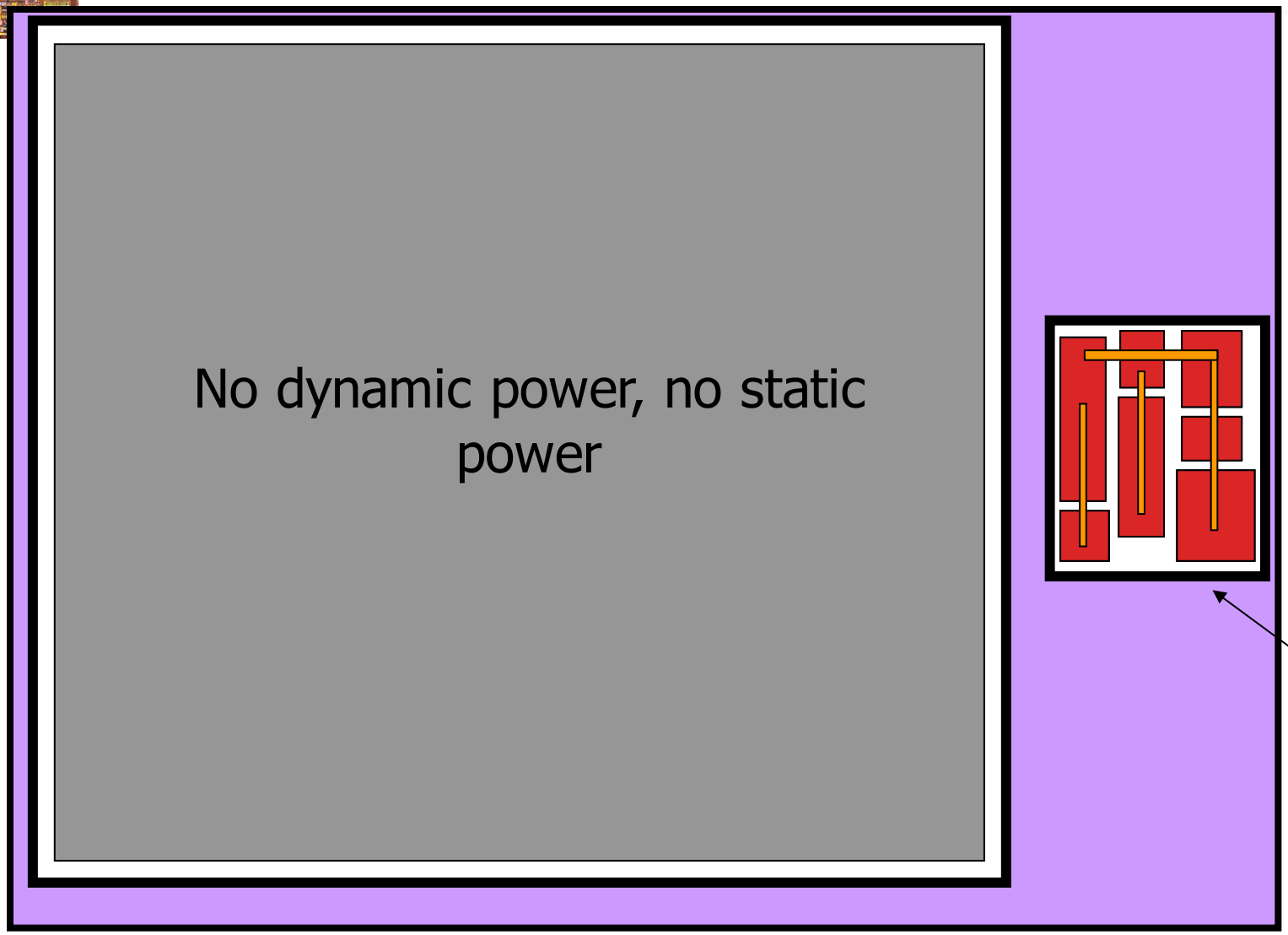
Problem Reformulation

- Best way to save energy is to do less work
  - Energy directly reduced by the reduction in work
  - But required time for the function decreases as well
    - Convert this into extra power gains
  - Shifts the optimal curve down and to the right



Saving power using multi-cores

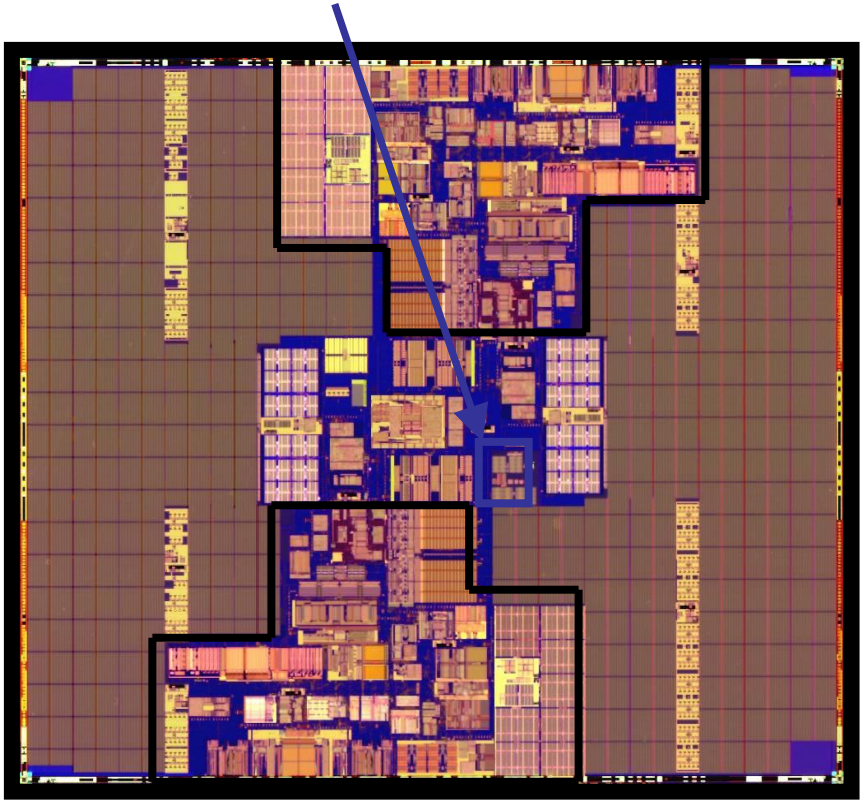




[Kumar et al]

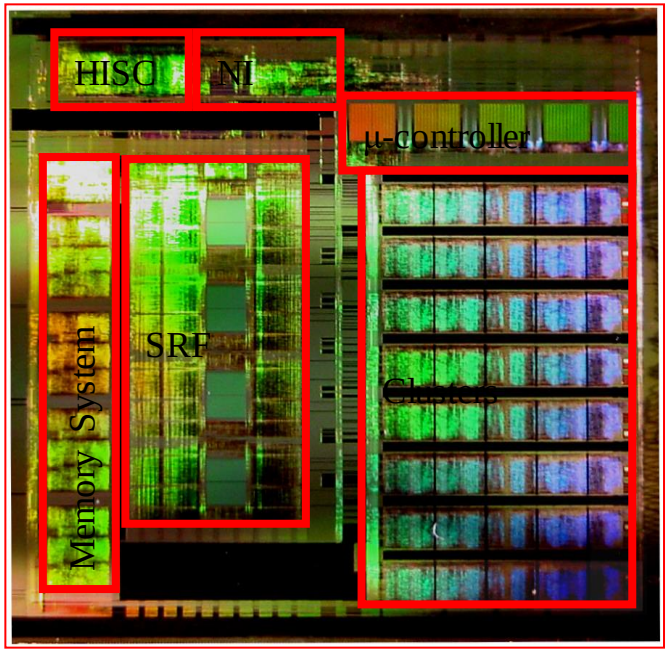
Constant Power Scaling

- Foxton controller on Itanium II
  - Raises Vdd/boosts F when most units idle
  - Lowers Vdd for parallel code to stay in budget



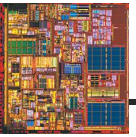
Exploit Specialization

- Optimize execution units for different applications
  - Reformulate the hardware to reduce needed work
  - Can improve energy efficiency for a class of applications
- Stream / Vector processing is a current example
  - Exploit locality, reuse
  - High compute density



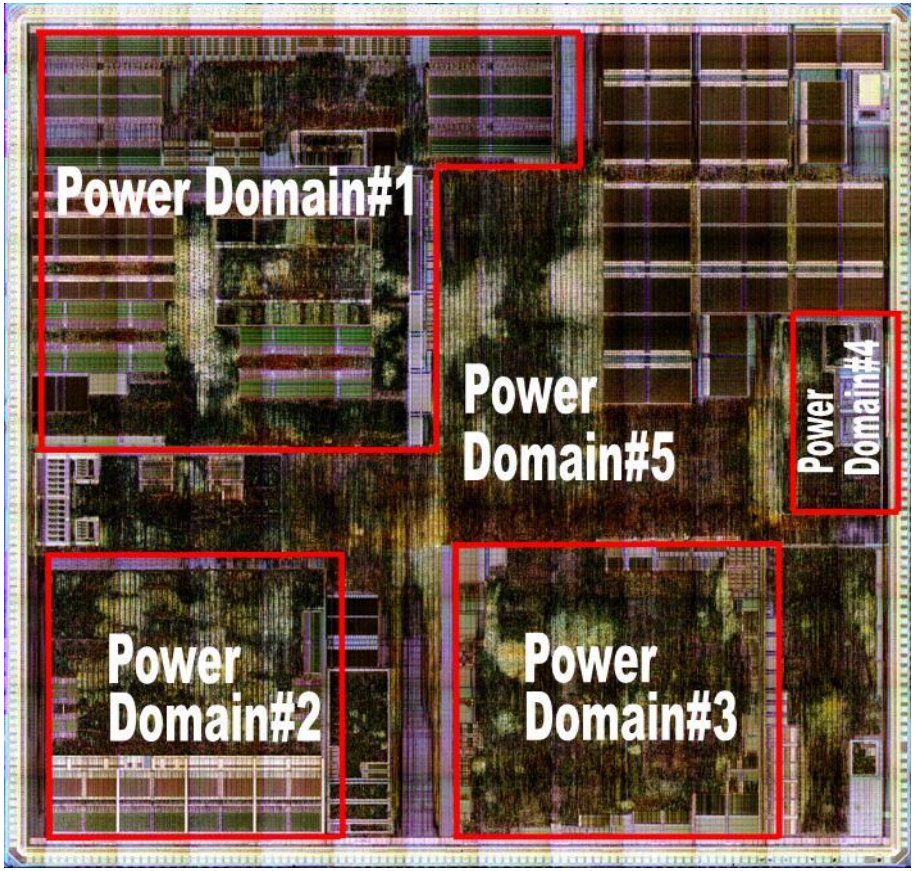
Bill Dally et al, Stanford Imagine

Exploit Integration



- Moving units onto one chip
  - Reduces the number of I/Os on system
  - I/O can take significant power today
  - Allows even larger integration

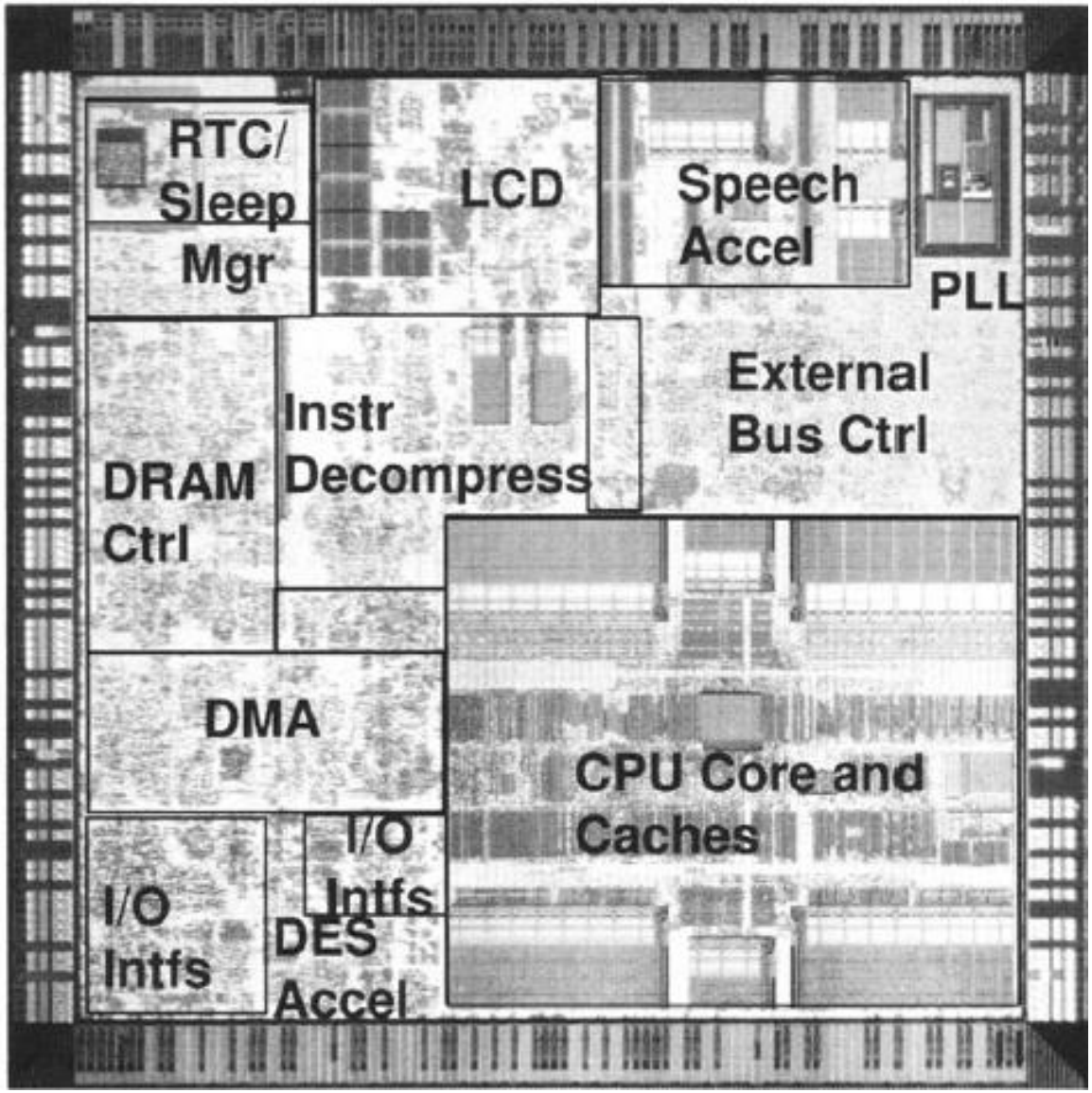
TI - OMAP2420



- Specialization
- And power domains
  - Most units are off
- OMAP 2420
  - 5 Power Domains
  - #1: MCU Core
  - #2: DSP Core
  - #3: Graphic Accelerator
  - #4: Core + Periph.
  - #5: Always On logic

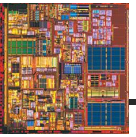
Royannez, et al, 90nm Low Leakage SoC Design Techniques for Wireless Applications, ISSCC 2005

Low-Power PowerPC



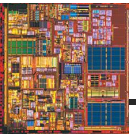
Nowka et al., Low-power PowerPC, ISSCC

Low Power Design



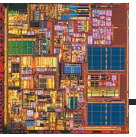
- Reduce dynamic power
  - α:
  - C:
  - V_{DD}:
  - f:
- Reduce static power

Low Power Design



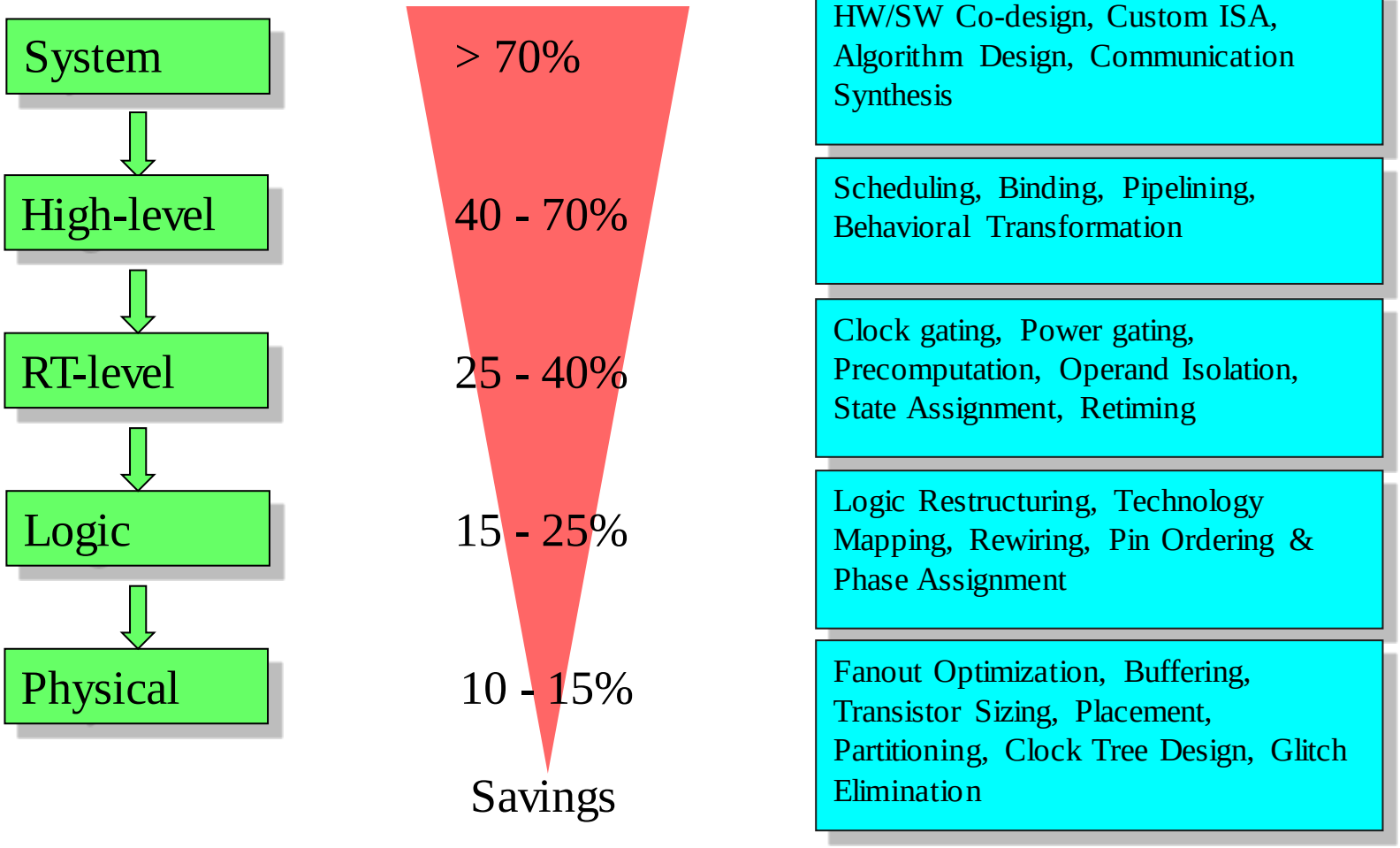
- Reduce dynamic power
  - α: clock gating, sleep mode
  - C: small transistors (esp. on clock), short wires
  - V_{DD}: lowest suitable voltage
  - f: lowest suitable frequency
- Reduce static power

Low Power Design



- Reduce dynamic power
  - α: clock gating, sleep mode
  - C: small transistors (esp. on clock), short wires
  - V_{DD}: lowest suitable voltage
  - f: lowest suitable frequency
- Reduce static power
  - Selectively use low V_t devices
  - Leakage reduction:
    stacked devices, body bias, low temperature

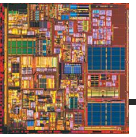
Power Saving Opportunities



A power-conscious design methodology addresses power at every level of the design hierarchy

Source: Pedram, 1999

Conclusions



- Power is the ultimate limiter for technology scaling.
  - Besides the physical limits
- Dynamic power has been dominating until about five years ago
  - Leakage power catches up due to Vt and Tox scaling
- Power optimization has changed the semiconductor industry
- Need to optimize power in each design stage
- Early and accurate power analysis is essential