CMPEN 411
VLSI Digital Circuits
Spring 2012

Lecture 13: Designing for Low Power

[Adapted from Rabaey's *Digital Integrated Circuits*, Second Edition, ©2003 J. Rabaey, A. Chandrakasan, B. Nikolic]

Review: Designing Fast CMOS Gates

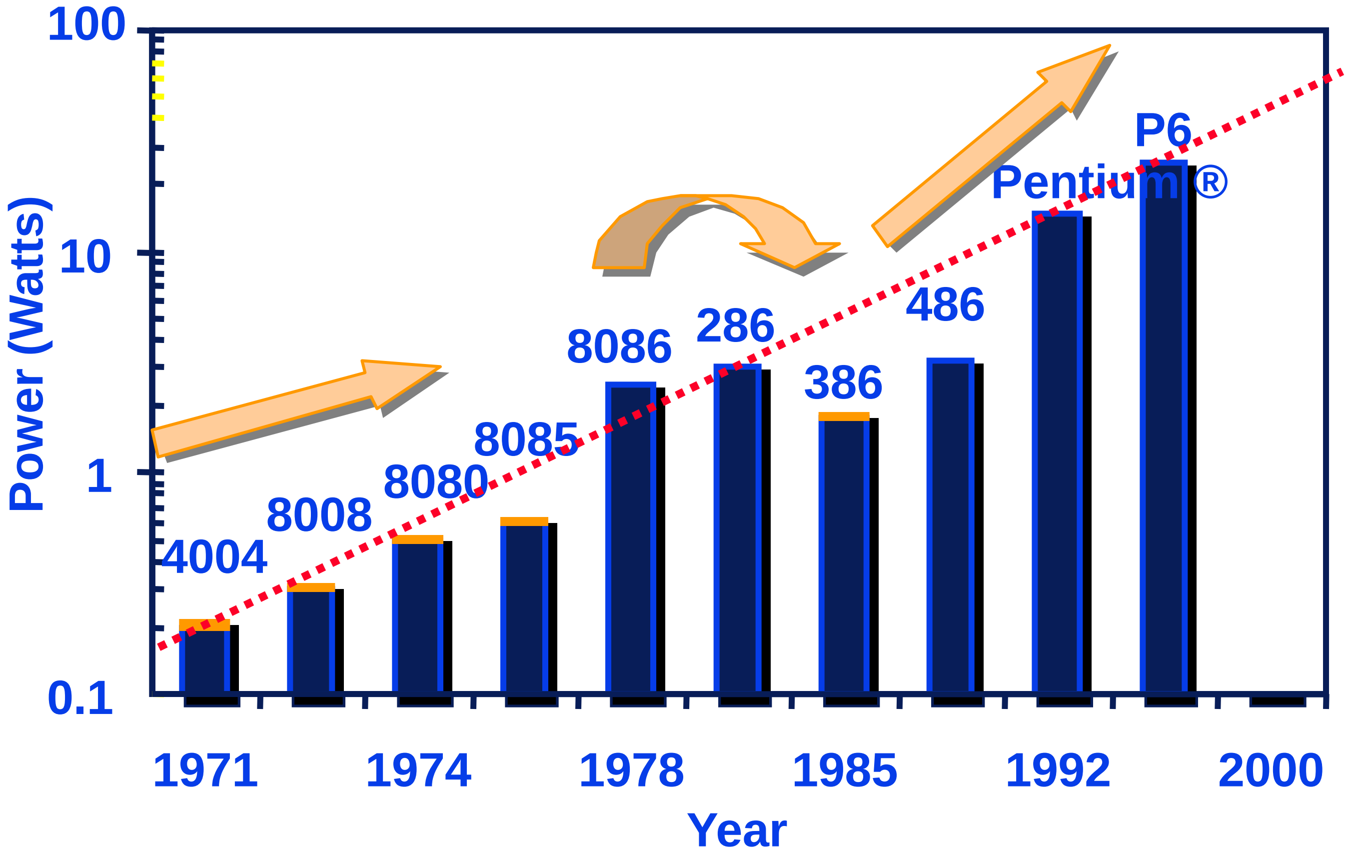
- Transistor sizing
- Progressive transistor sizing
  - fet closest to the output is smallest of series fets
- Transistor ordering
  - put latest arriving signal closest to the output
- Logic structure reordering
  - replace large fan-in gates with smaller fan-in gate network
- Logical effort
- Buffer (inverter) insertion
  - separate large fan-in from large C_L with buffers
  - uses buffers so that the path delay is minimized

Why Power Matters

- Packaging costs
- Power supply rail design
- Chip and system cooling costs
- Noise immunity and system reliability
- Battery life (in portable systems)
- Environmental concerns
  - Office equipment accounted for 5% of total US commercial energy usage in 1993
  - *Energy Star* compliant systems

Why worry about power? – Power Dissipation

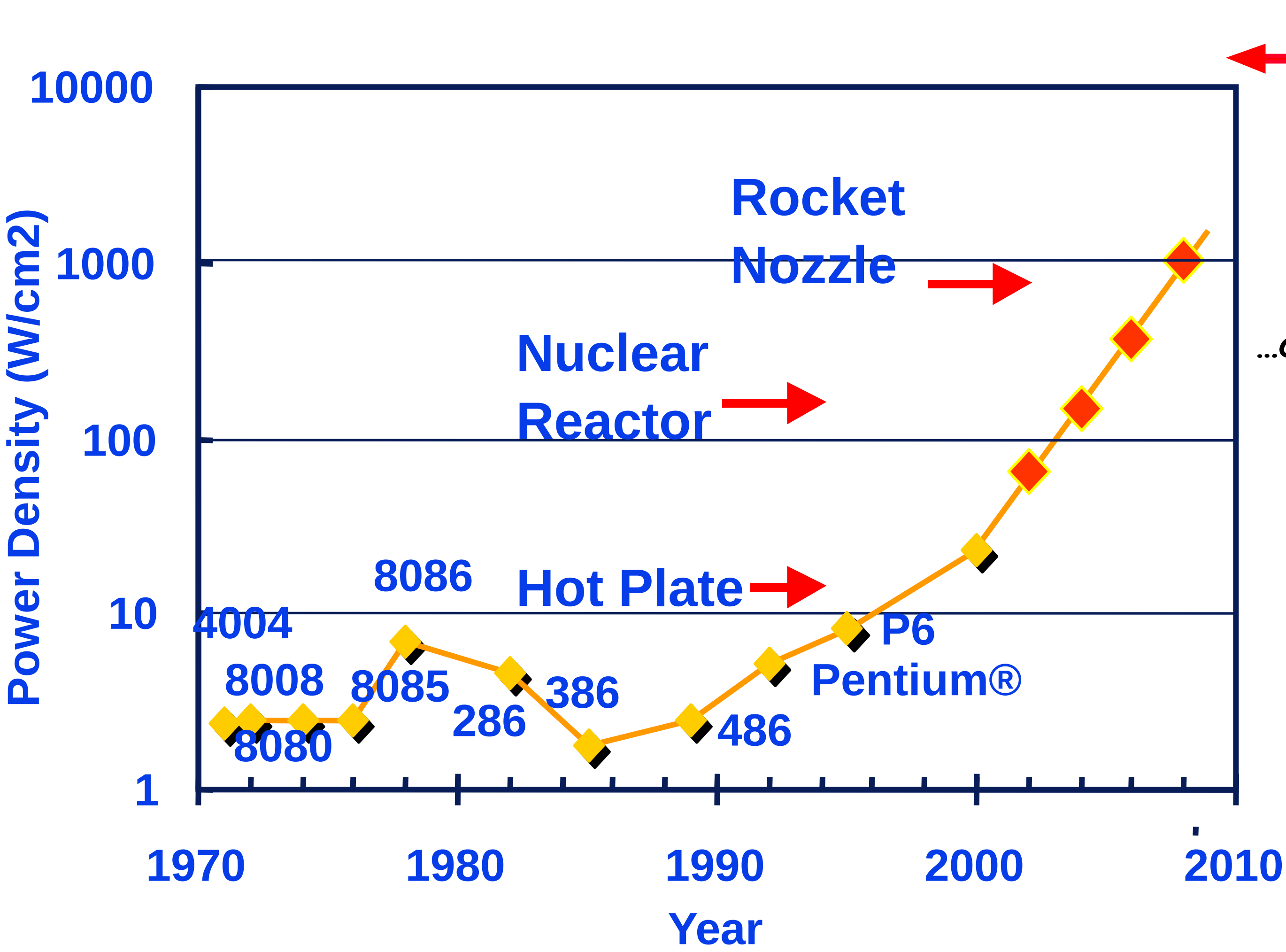
Lead microprocessors power continues to increase



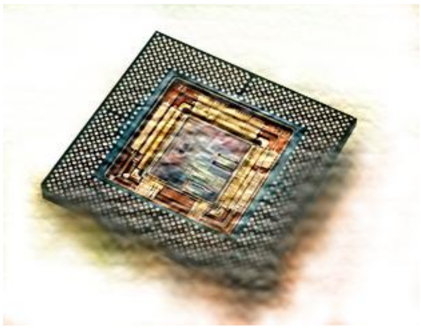
Power delivery and dissipation will be prohibitive

Source: Borkar, De Intel®

Why worry about power? – Chip Power Density



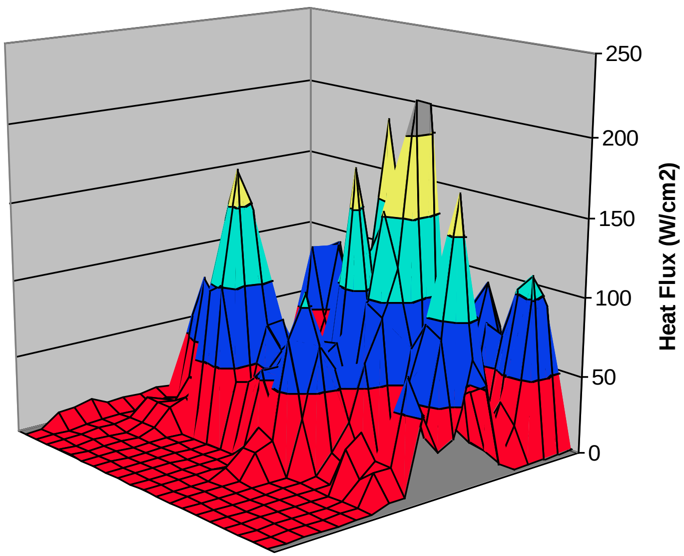
…chips might become hot…



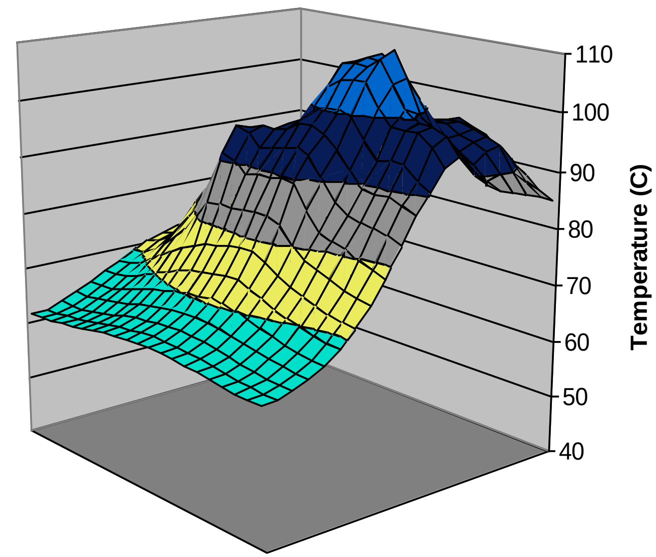
Source: Borkar, De Intel®

Chip Power Density Distribution => Heat

Power Map



On-Die Temperature



- Heat dissipation => temperature,

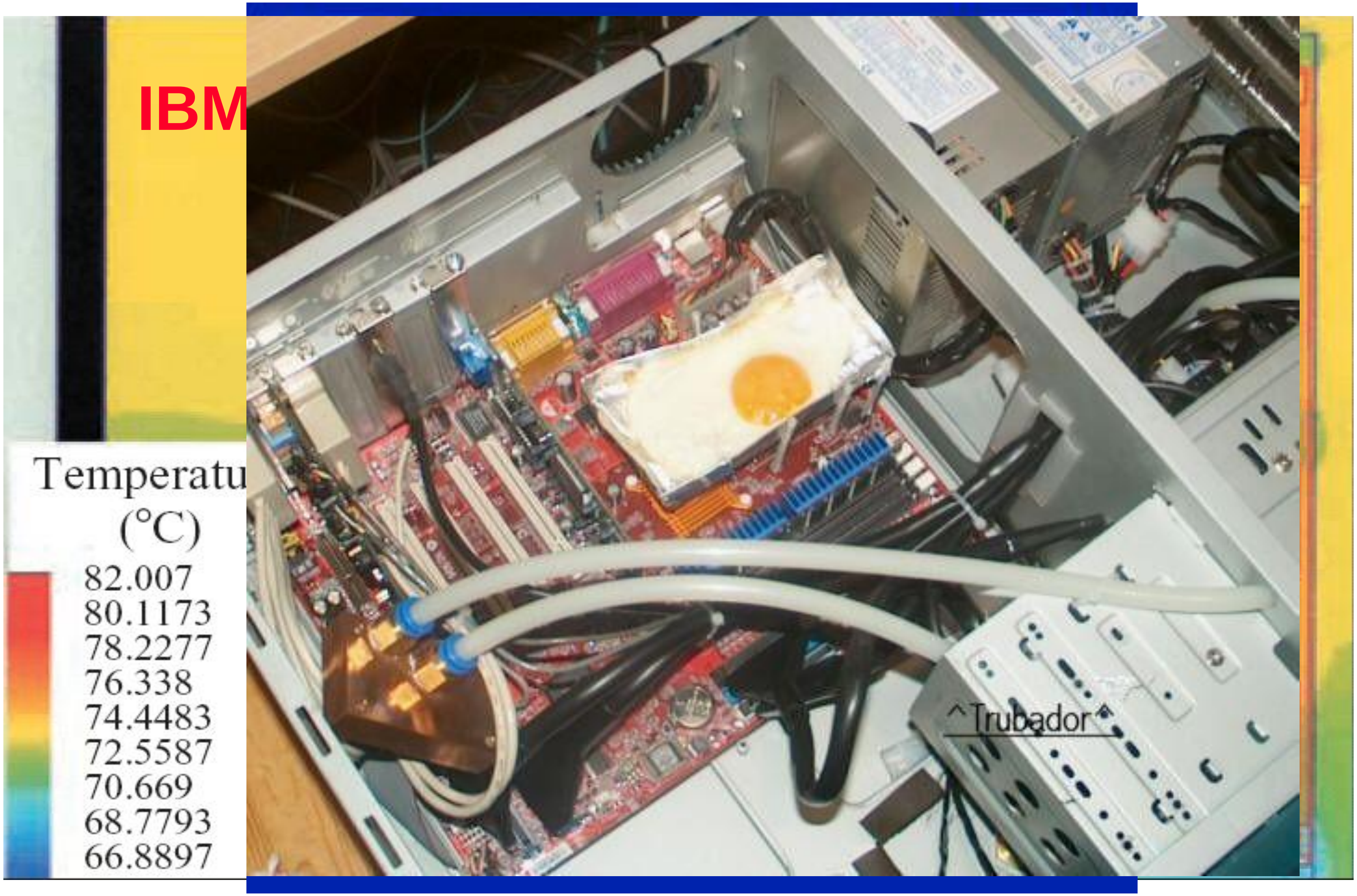
  higher power density => higher temperature

- Power density is not uniformly distributed across the chip

Intel Pentium 4: (0.18 um) 64 W @ 217 mm^2

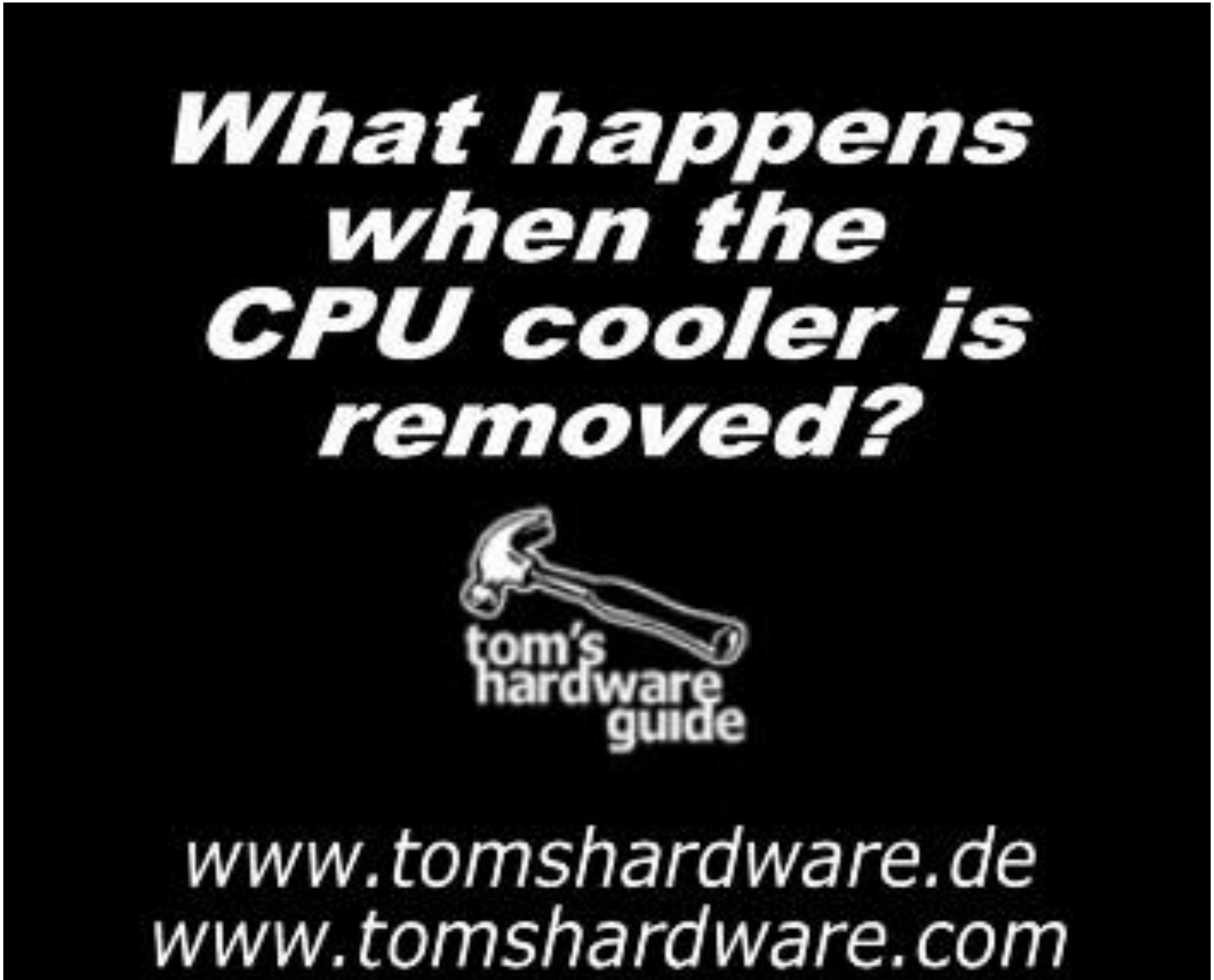
Intel Pentium 4: (90 nm) 103W @ 112 mm^2

Examples


IBM
Temperatu
(°C)
82.007
80.1173
78.2277
76.338
74.4483
72.5587
70.669
68.7793
66.8897
^Trubador^

Apple Power G5

Problem Illustration



Power and temperature are BAD

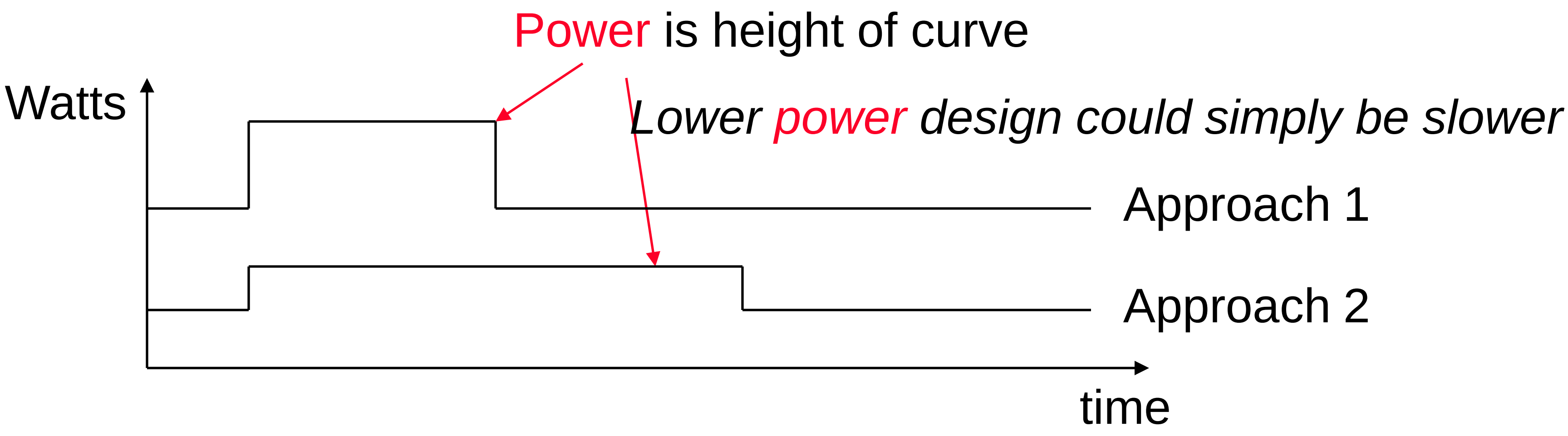
- and can be EVIL

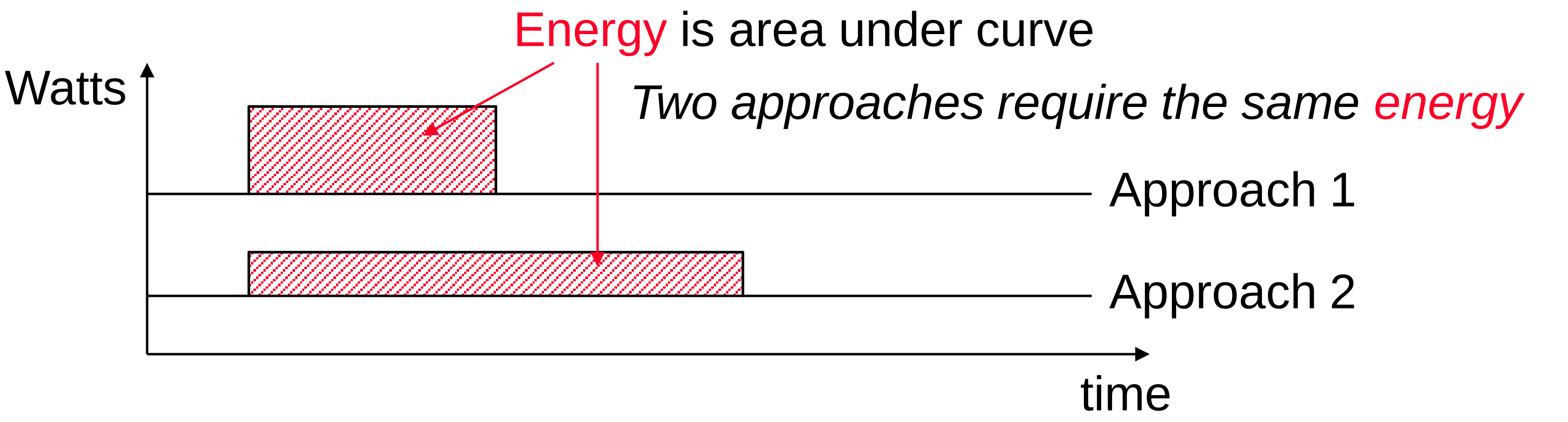
Source: Tom's Hardware Guide
http://www6.tomshardware.com/cpu/01q3/010917/heatvideo-01.html

Power and Energy Figures of Merit

- Power consumption in Watts
  - determines battery life in hours
- Peak power
  - determines power ground wiring designs
  - sets packaging limits
  - impacts signal noise margin and reliability analysis
- Energy efficiency in Joules
  - rate at which power is consumed over time
- Energy = power * delay
  - Joules = Watts * seconds
  - lower energy number means less power to perform a computation at the same frequency

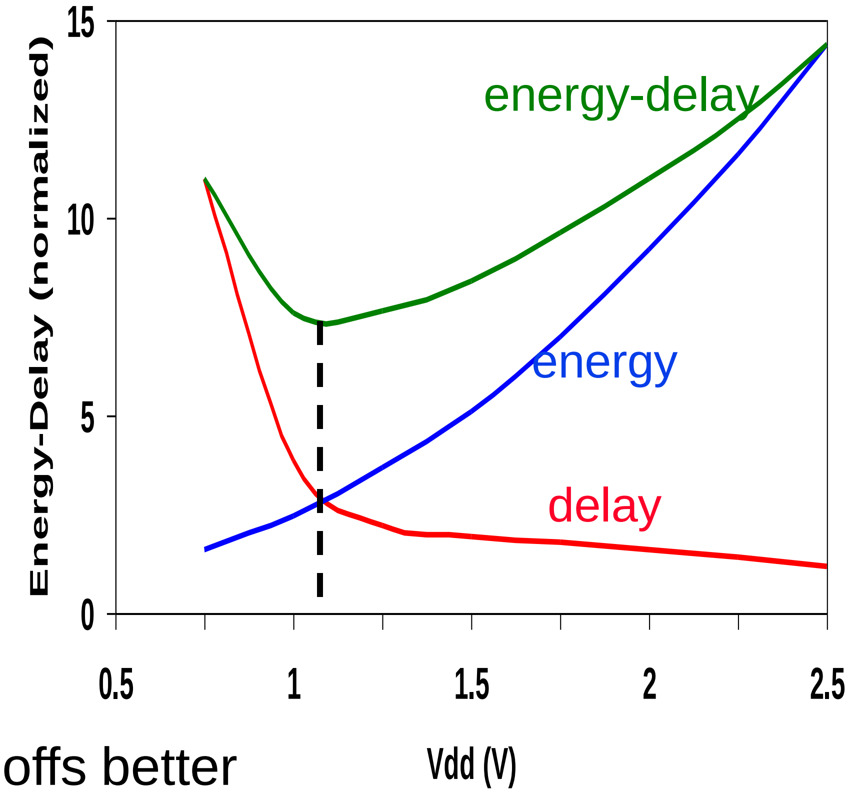
Power versus Energy





PDP and EDP

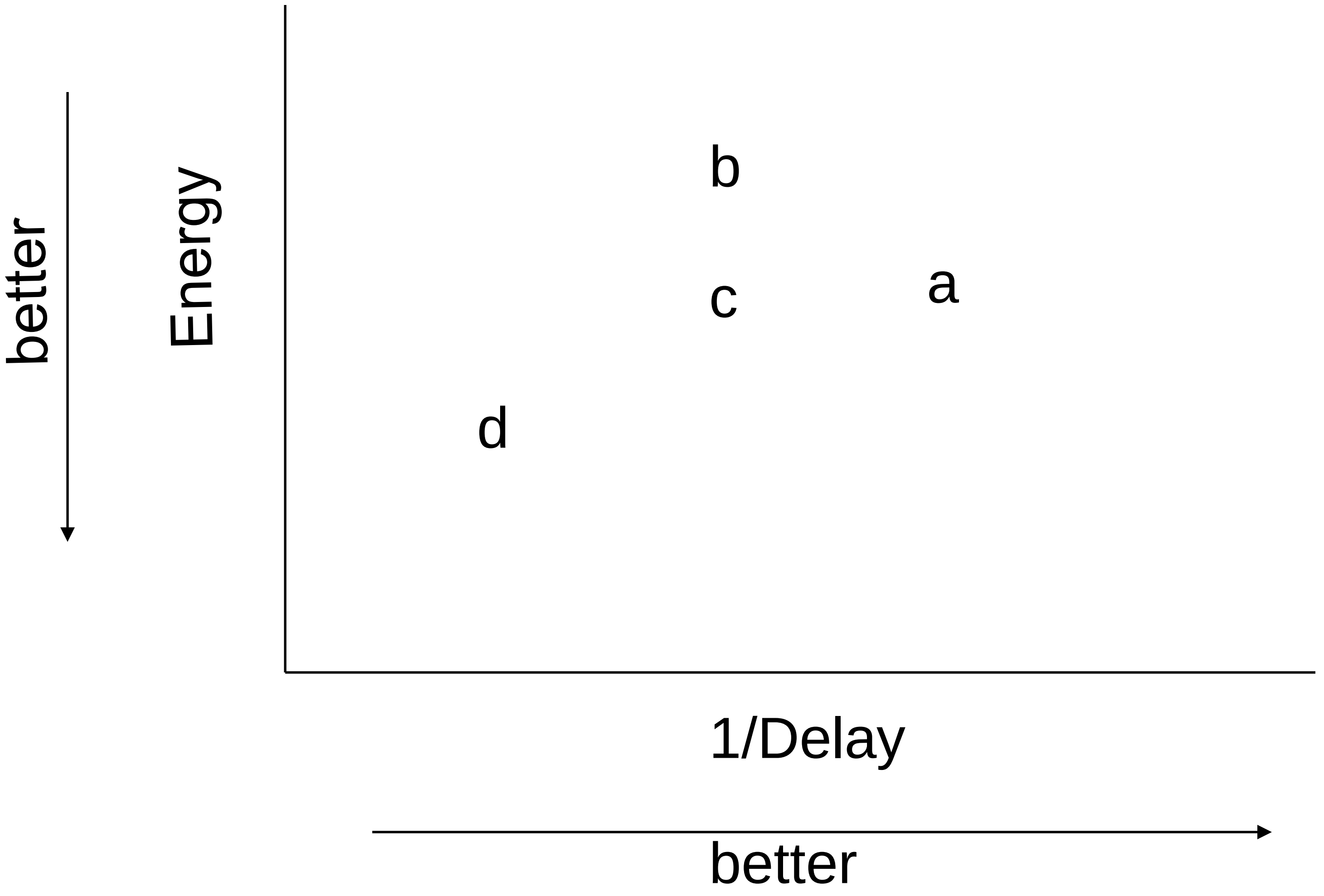
- Power-delay product (PDP) = $P_{av} * t_p = (C_L V_{DD}^2)/2$
  - PDP is the average energy consumed per switching event (Watts * sec = Joule)
  - lower power design could simply be a slower design
- Energy-delay product (EDP) = PDP * $t_p = P_{av} * t_p^2$
  - EDP is the average energy consumed multiplied by the computation time required
  - takes into account that one can trade increased delay for lower energy/operation (e.g., via supply voltage scaling that increases delay, but decreases energy consumption)



  - allows one to understand tradeoffs better

Understanding Tradeoffs

- Which design is the "best" (fastest, coolest, both) ?



CMOS Power Equations

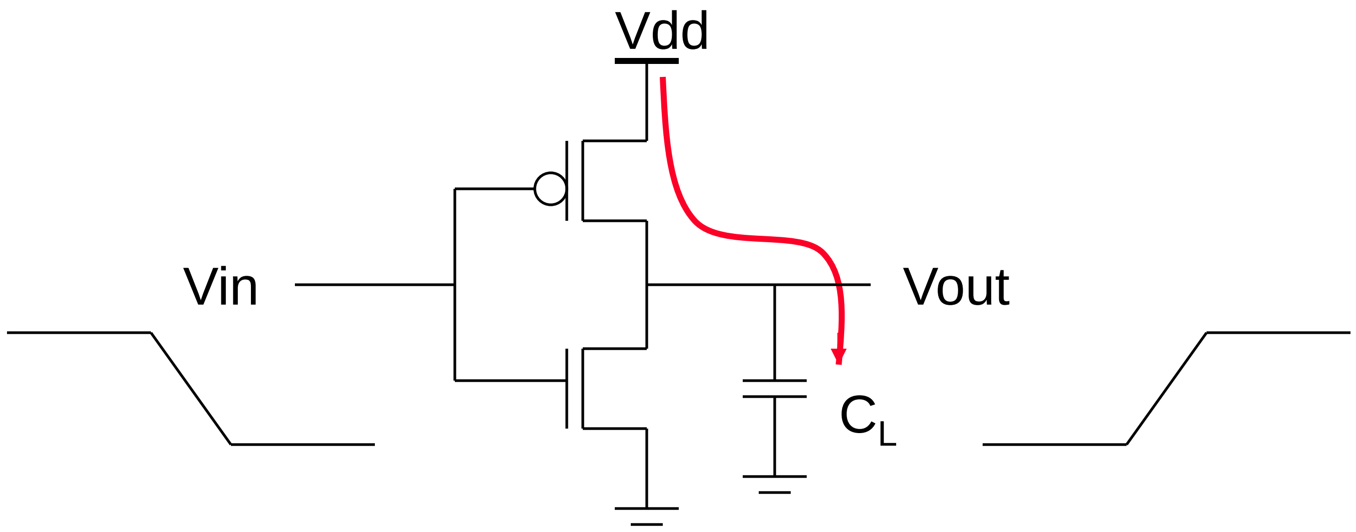
$$P = C_L\ V_{DD}^2\ f \quad + \quad t_{sc} V_{DD}\ I_{peak}\ f \quad + \quad V_{DD}\ I_{leak}$$

Dynamic power

Short-circuit power

Leakage power

Dynamic Power Consumption



$f_{0\rightarrow1}$

$$\text{Energy/transition} = C_L * V_{DD}^2 * P_{0\rightarrow1}$$

$$P_{dyn} = \text{Energy/transition} * f = C_L * V_{DD}^2 * P_{0\rightarrow1} * f$$

$$P_{dyn} = C_{EFF} * V_{DD}^2 * f \quad \text{where } C_{EFF} = P_{0\rightarrow1} C_L$$

Not a function of transistor sizes!

Data dependent - a function of switching activity!

Lowering Dynamic Power

Capacitance:
Function of fan-out, wire length, transistor sizes

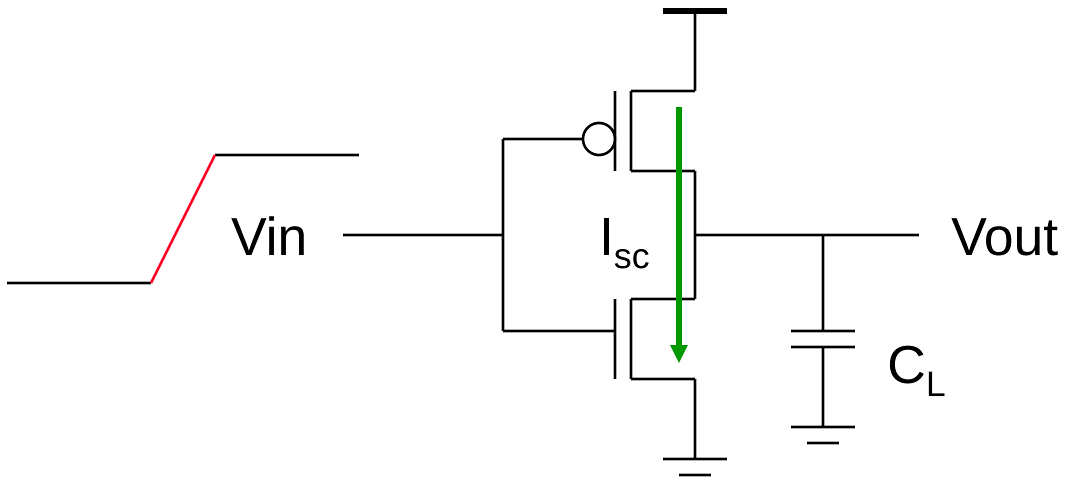
Supply voltage:
Has been dropping with successive generations

$$P_{dyn} = C_L \, V_{DD}^2 \, P_{0\rightarrow 1} \, f$$

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

Clock frequency:
Increasing…

Short Circuit Power Consumption



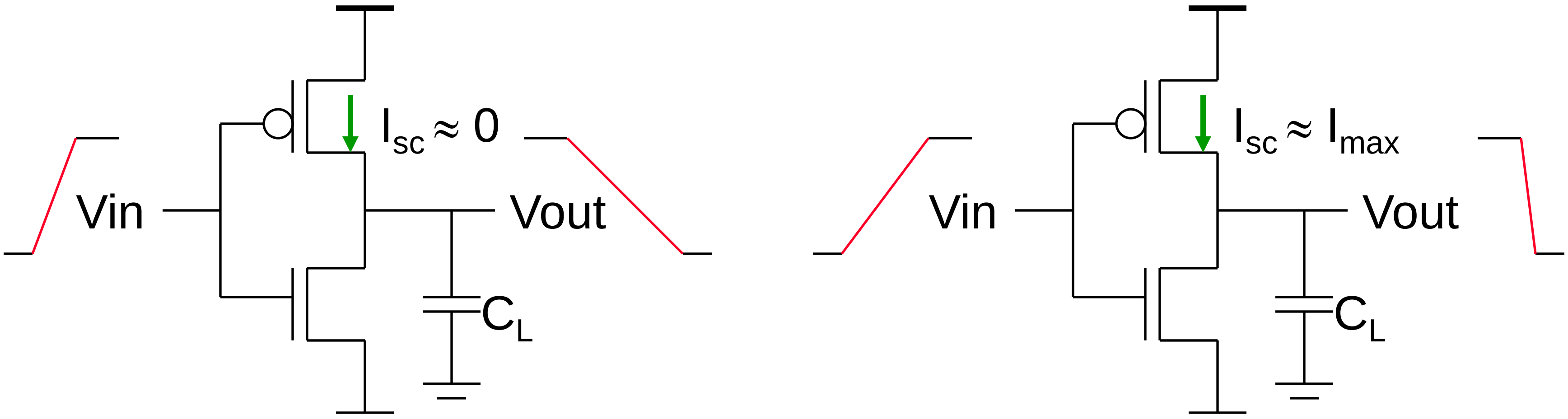
Finite slope of the input signal causes a direct current path between V_{DD} and GND for a short period of time during switching when both the NMOS and PMOS transistors are conducting.

Short Circuit Currents Determinates

$$P_{sc} = t_{sc} \, V_{DD} \, I_{peak} \, f_{0 \rightarrow 1}$$

- Duration and slope of the input signal, t_{sc}
- I_{peak} determined by
  - the saturation current of the P and N transistors which depend on their sizes, process technology, temperature, etc.
  - strong function of the ratio between input and output slopes
    - a function of C_L

Impact of C_L on P_{sc}



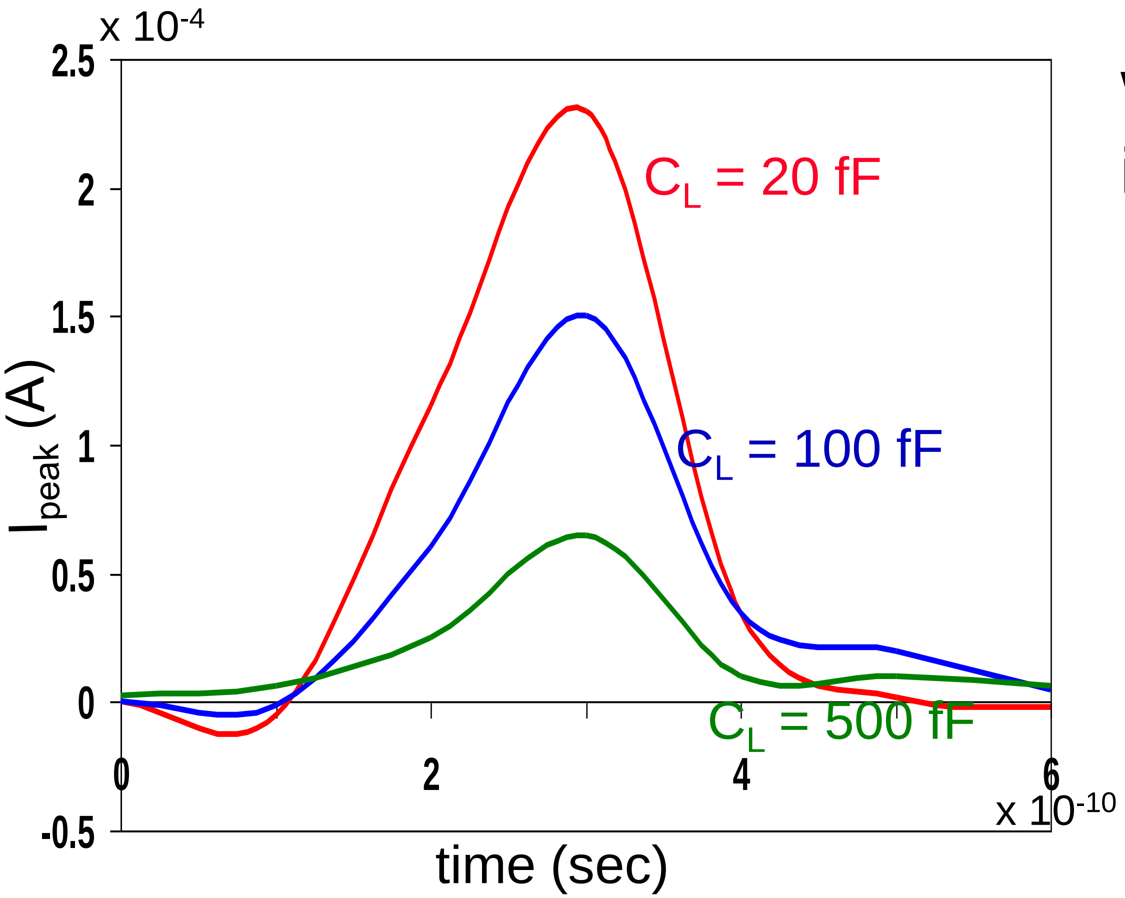
Large capacitive load

Output fall time significantly larger than input rise time.

Small capacitive load

Output fall time substantially smaller than the input rise time.

I_{peak} as a Function of C_L

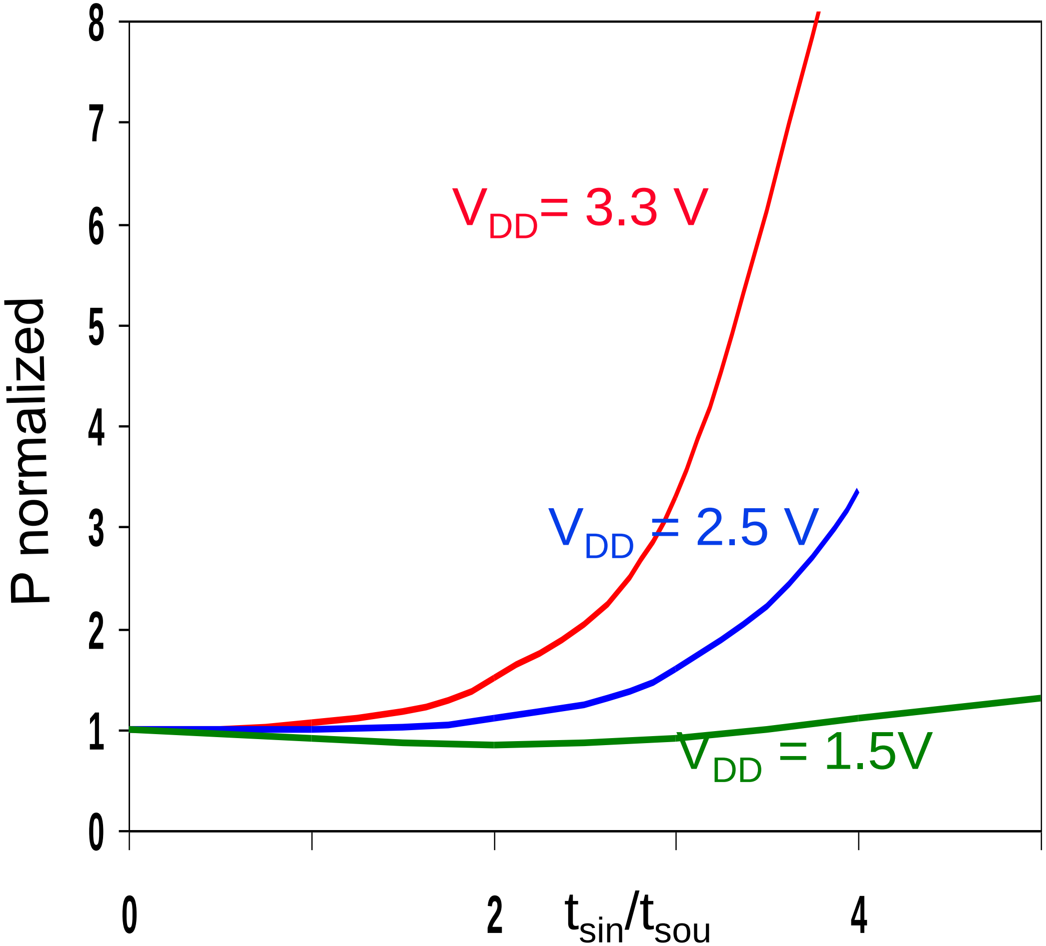


When load capacitance is small, I_{peak} is large.

Short circuit dissipation is minimized by matching the rise/fall times of the input and output signals - slope engineering.

500 psec input slope

P_{sc} as a Function of Rise/Fall Times



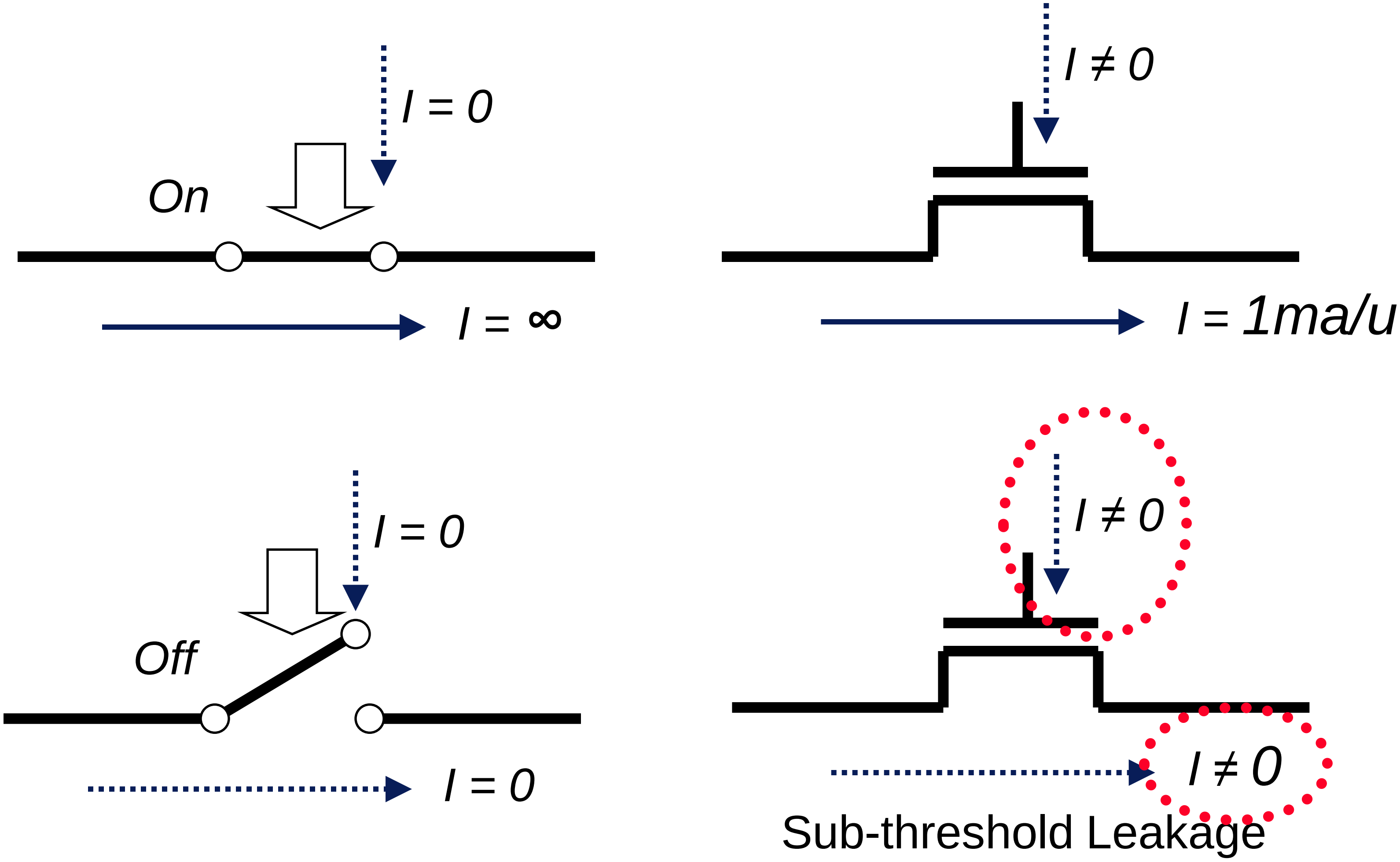
When load capacitance is small (t_{sin}/t_{sout} > 2 for V_{DD} > 2V) the power is dominated by P_{sc}

If $V_{DD} < V_{Tn} + |V_{Tp}|$ then P_{sc} is eliminated since both devices are never on at the same time.

W/L_p = 1.125 μm/0.25 μm
W/L_n = 0.375 μm/0.25 μm
C_L = 30 fF

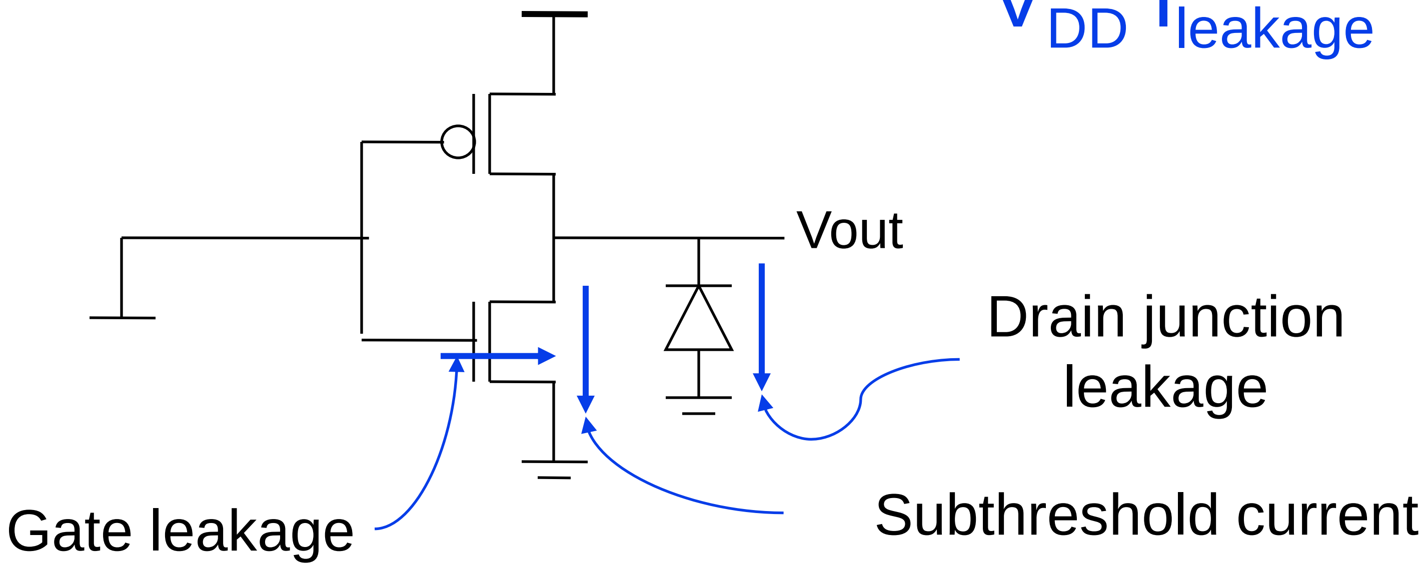
normalized wrt zero input rise-time dissipation

Is Transistor a Good Switch?



Leakage (Static) Power Consumption

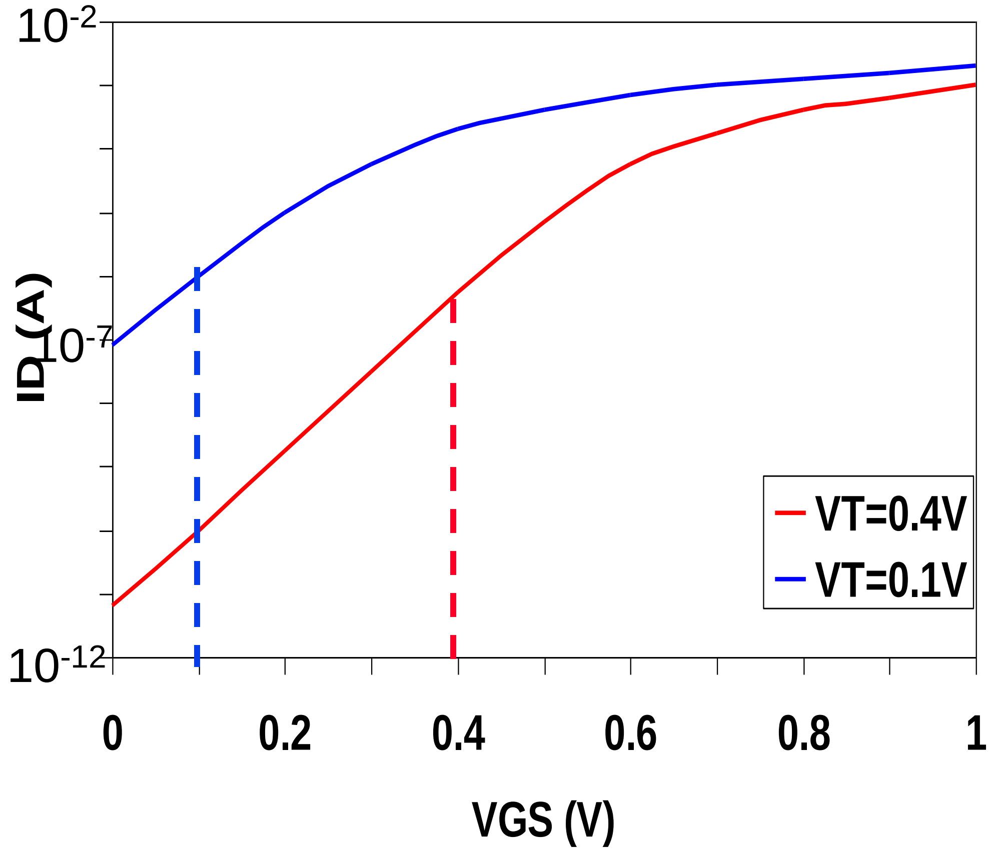
$V_{DD} \, I_{leakage}$



Sub-threshold current is the dominant factor.

Leakage as a Function of V_T

- Continued scaling of supply voltage and the subsequent scaling of threshold voltage will make subthreshold conduction a dominate component of power dissipation.



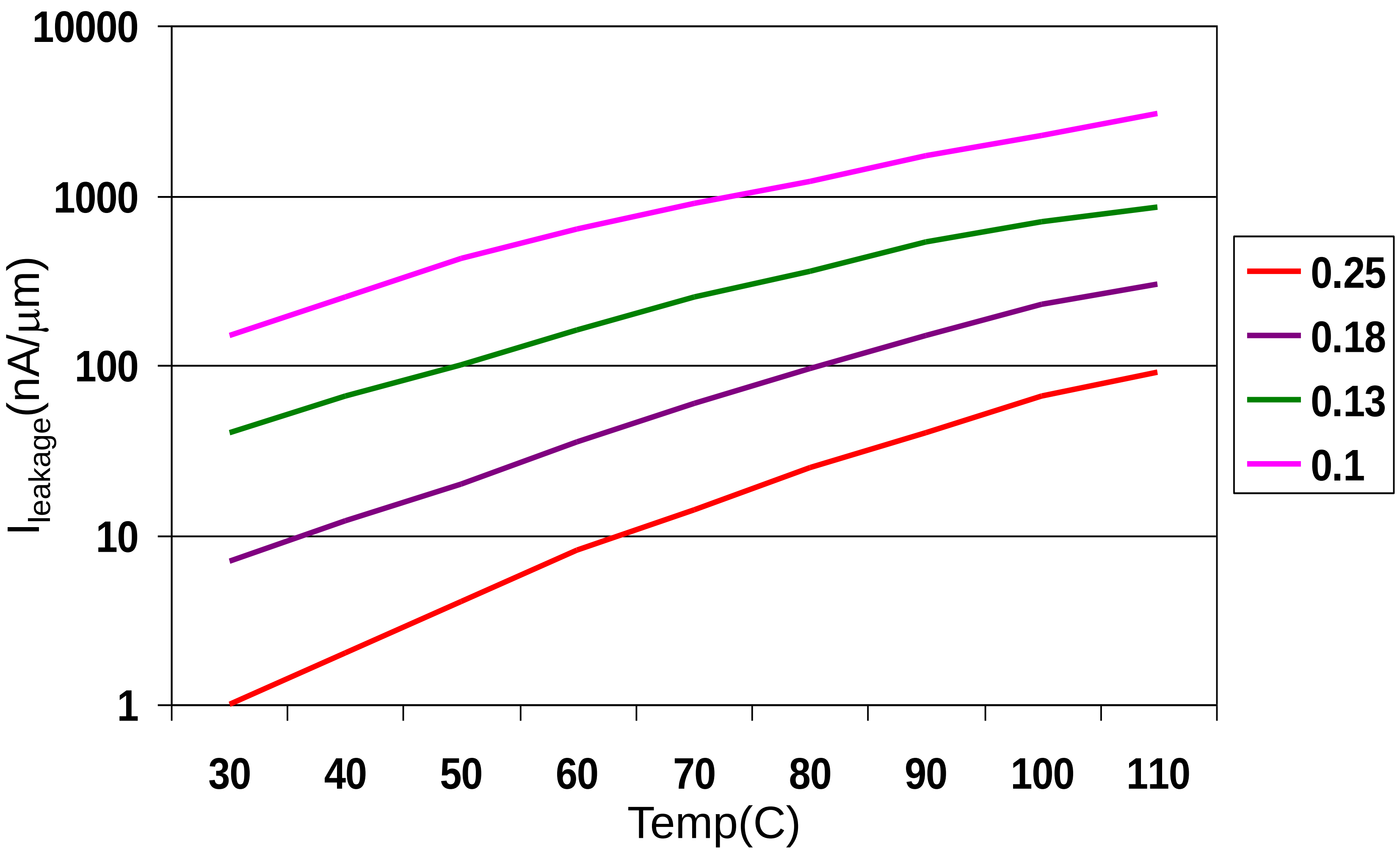
- An 90mV/decade V_T roll-off - so each 270mV increase in V_T gives 3 orders of magnitude reduction in leakage (but adversely affects performance)

TSMC Processes Leakage and V_T

	CL018 G	CL018 LP	CL018 ULP	CL018 HS	CL015 HS	CL013 HS
V_{dd}	1.8 V	1.8 V	1.8 V	2 V	1.5 V	1.2 V
T_{ox} (effective)	42 Å	42 Å	42 Å	42 Å	29 Å	24 Å
L_{gate}	0.16 μm	0.16 μm	0.18 μm	0.13 μm	0.11 μm	0.08 μm
I_{DSat} (n/p) (μA/μm)	600/260	500/180	320/130	780/360	860/370	920/400
I_{off} (leakage) (ρA/μm)	20	1.60	0.15	300	1,800	13,000
V_{Tn}	0.42 V	0.63 V	0.73 V	0.40 V	0.29 V	0.25 V
FET Perf. (GHz)	30	22	14	43	52	80

From MPR, 2000

Exponential Increase in Leakage Currents

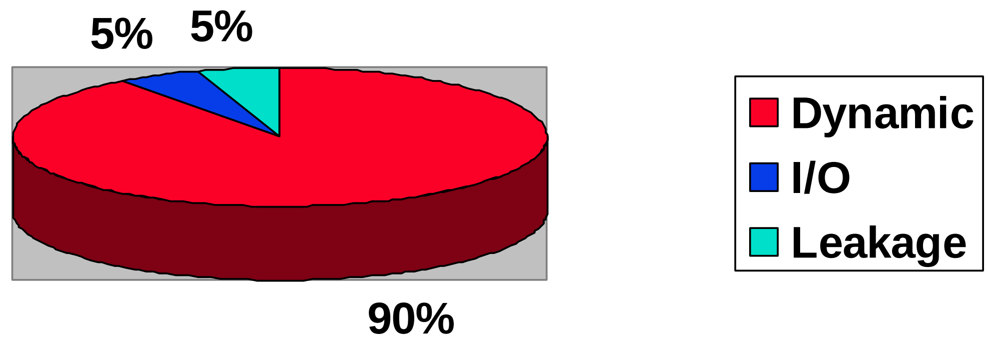


From De,1999

Itanium example

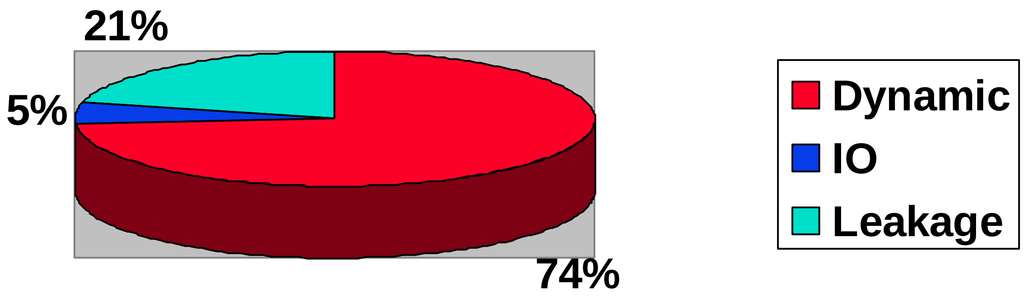
Itanium 2 (Intel)

0.18 um, 1.5V, 1Ghz, 221M transistors



Latest generation Itanium (Intel)

0.13 um, 1.3V, 1.5Ghz, 410M transistors



example

Next Lecture

	Constant Throughput/Latency		Variable Throughput/Latency
Energy	Design Time	Non-active Modules	Run Time
Active (Dynamic)	Logic design Reduced V_{dd} TSizing Multi-V_{dd}	Clock Gating	DFS, DVS (Dynamic Freq, Voltage Scaling)
Leakage (Standby)	Multi-V_T Stack effect Pin ordering	Sleep Transistors Multi-V_{dd} Variable V_T Input control	Variable V_T