Digital Integrated Circuits
(83-313)

Lecture 4:
Technology Scaling

Semester B, 2016-17

Lecturer: Dr. Adam Teman

TAs: Itamar Levi,
Robert Giterman

2 April 2017

Emerging Nanoscaled
Integrated Circuits and Systems Labs

Bar-Ilan University
אוניברסיטת בר-אילן

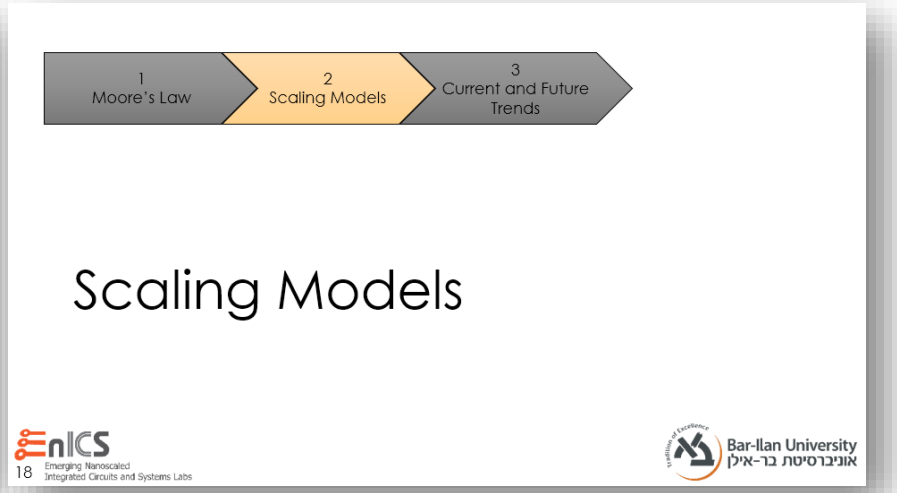
Motivation

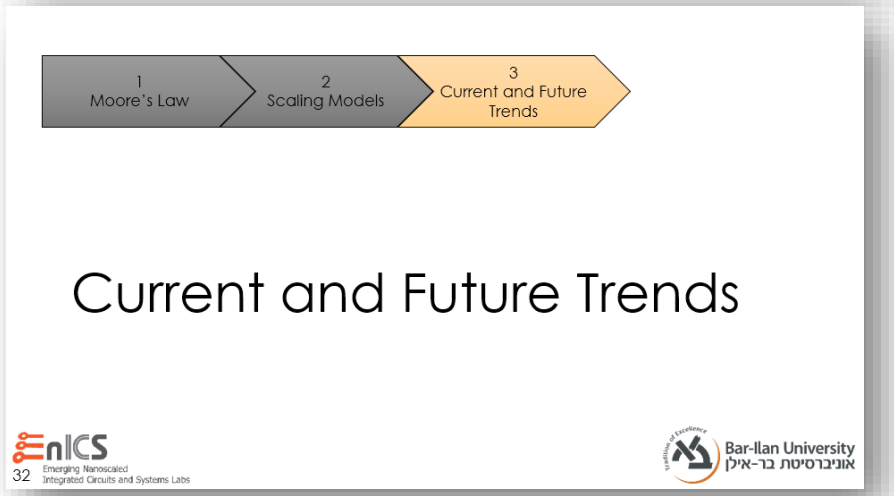
- **If transistors were people…**

Courtesy: Intel 2011

- Now imagine that those 1.3B people could fit onstage in the original music hall.
- That's the scale of Moore's Law.

Lecture Content

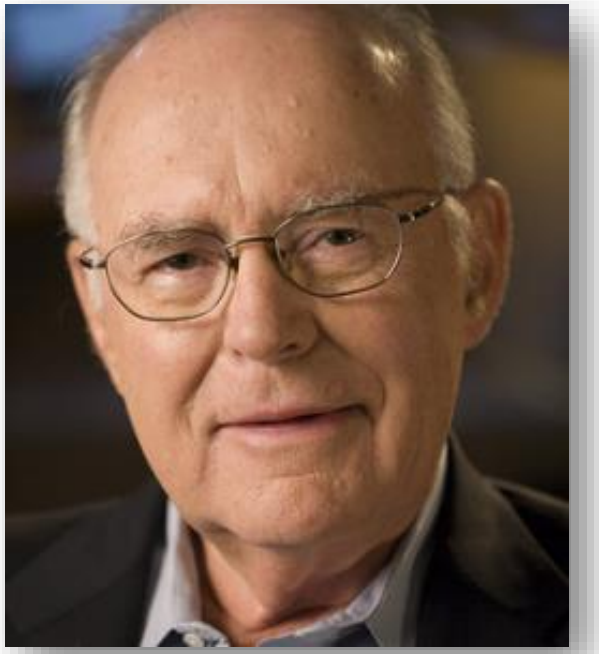


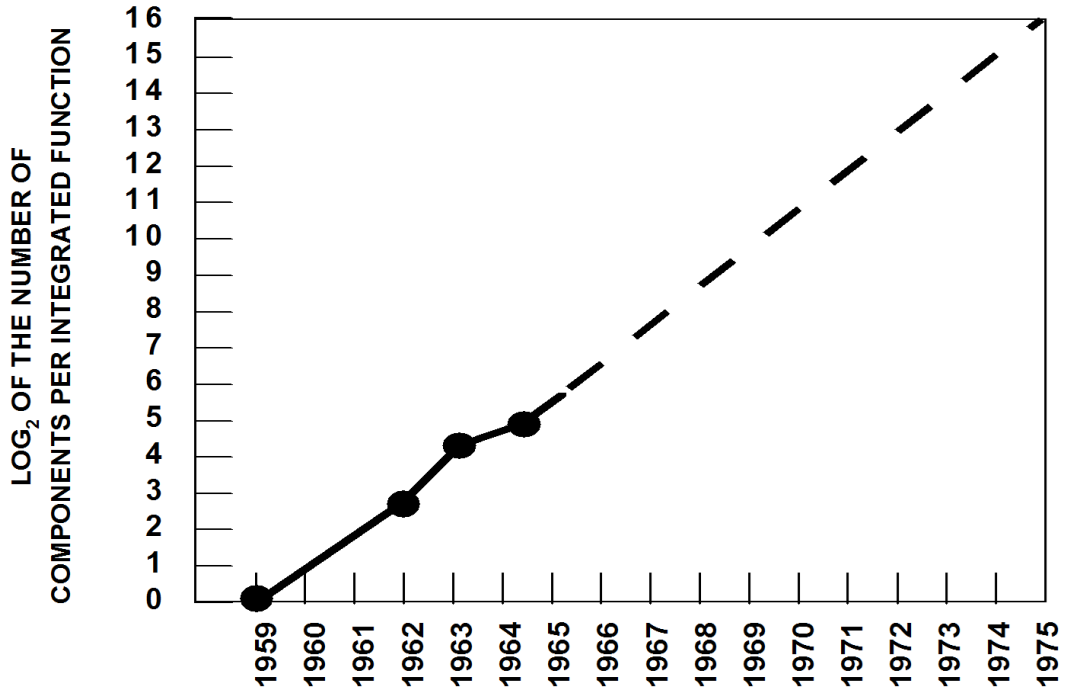


Moore's Law

Moore's Law

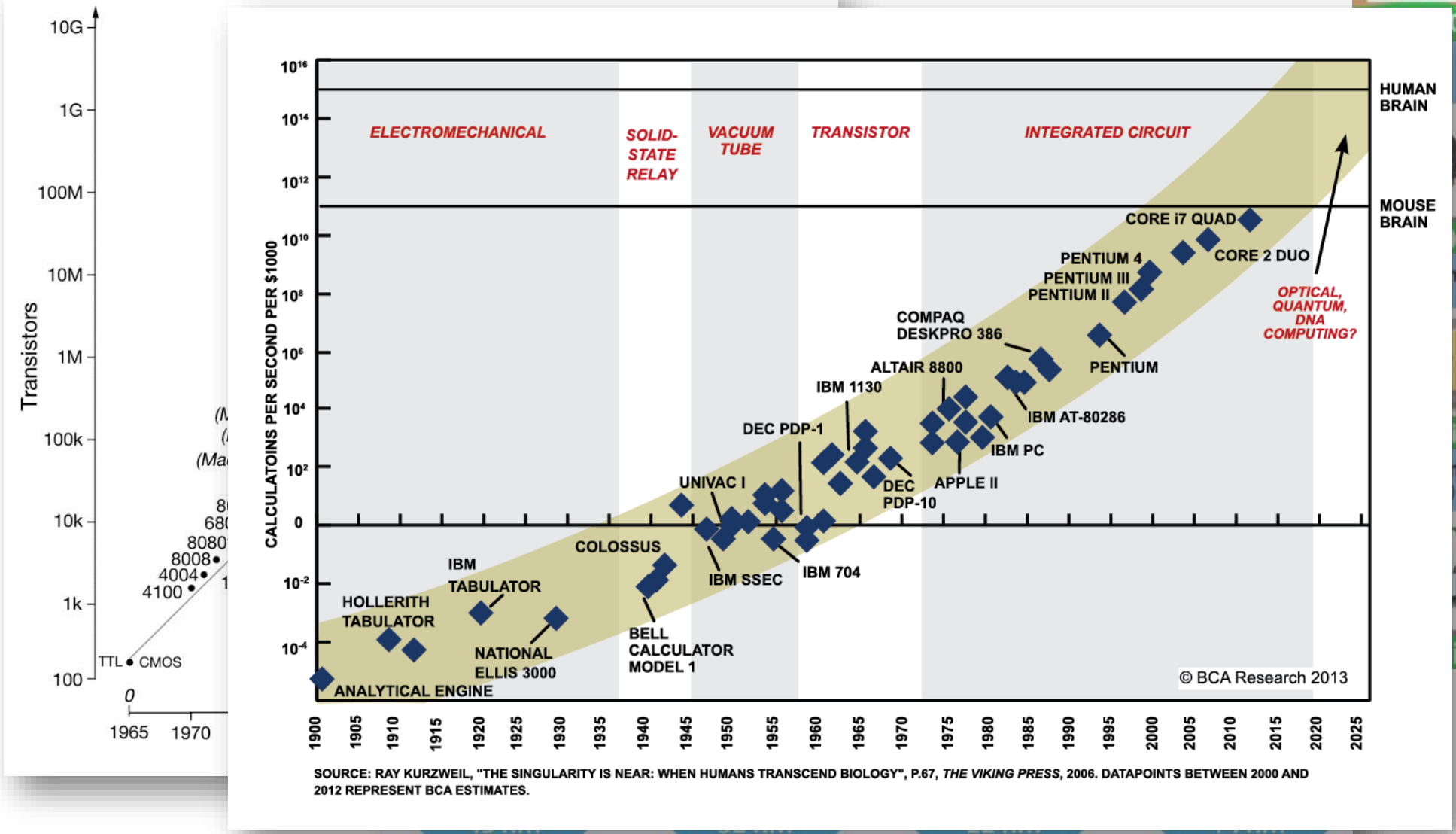
- In 1965, Gordon Moore noted that the number of components on a chip doubled every 18 to 24 months.
- He made a prediction that semiconductor technology will double its effectiveness every 18 months



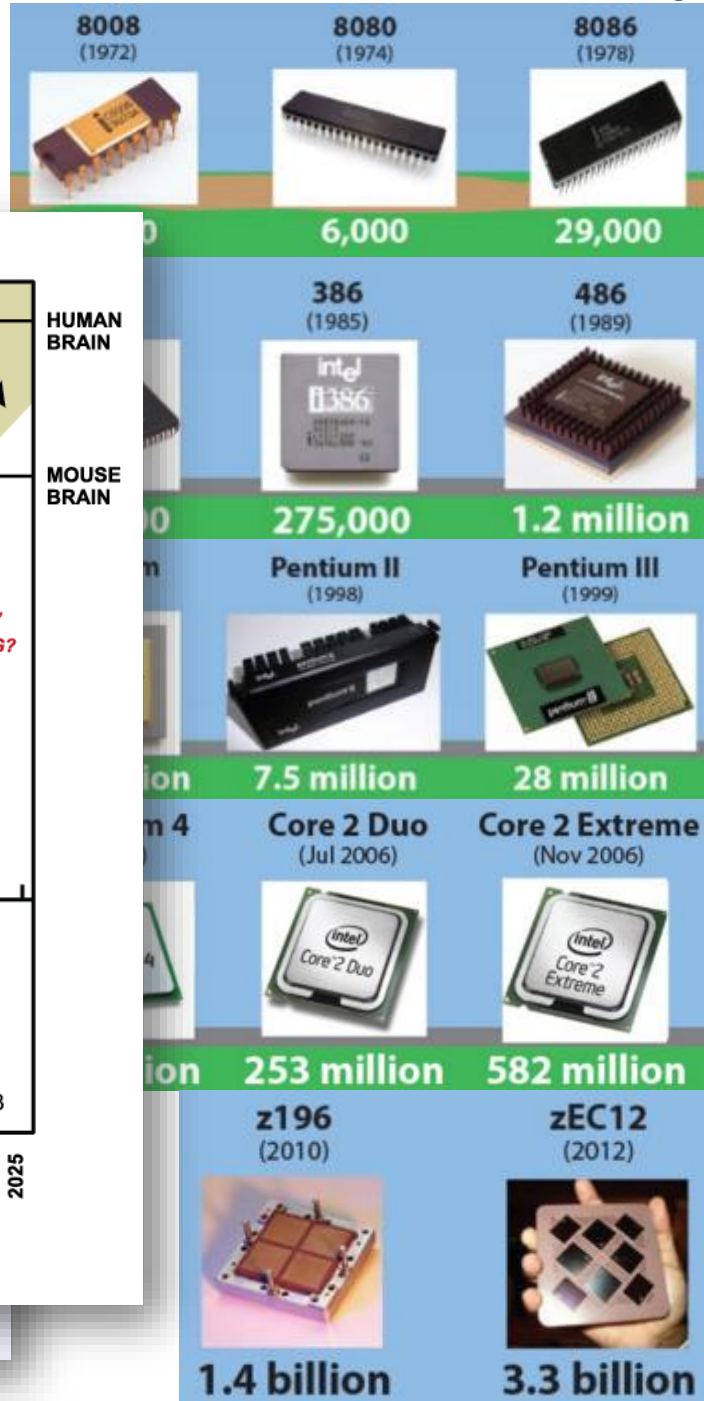


Electronics, April 19, 1965.

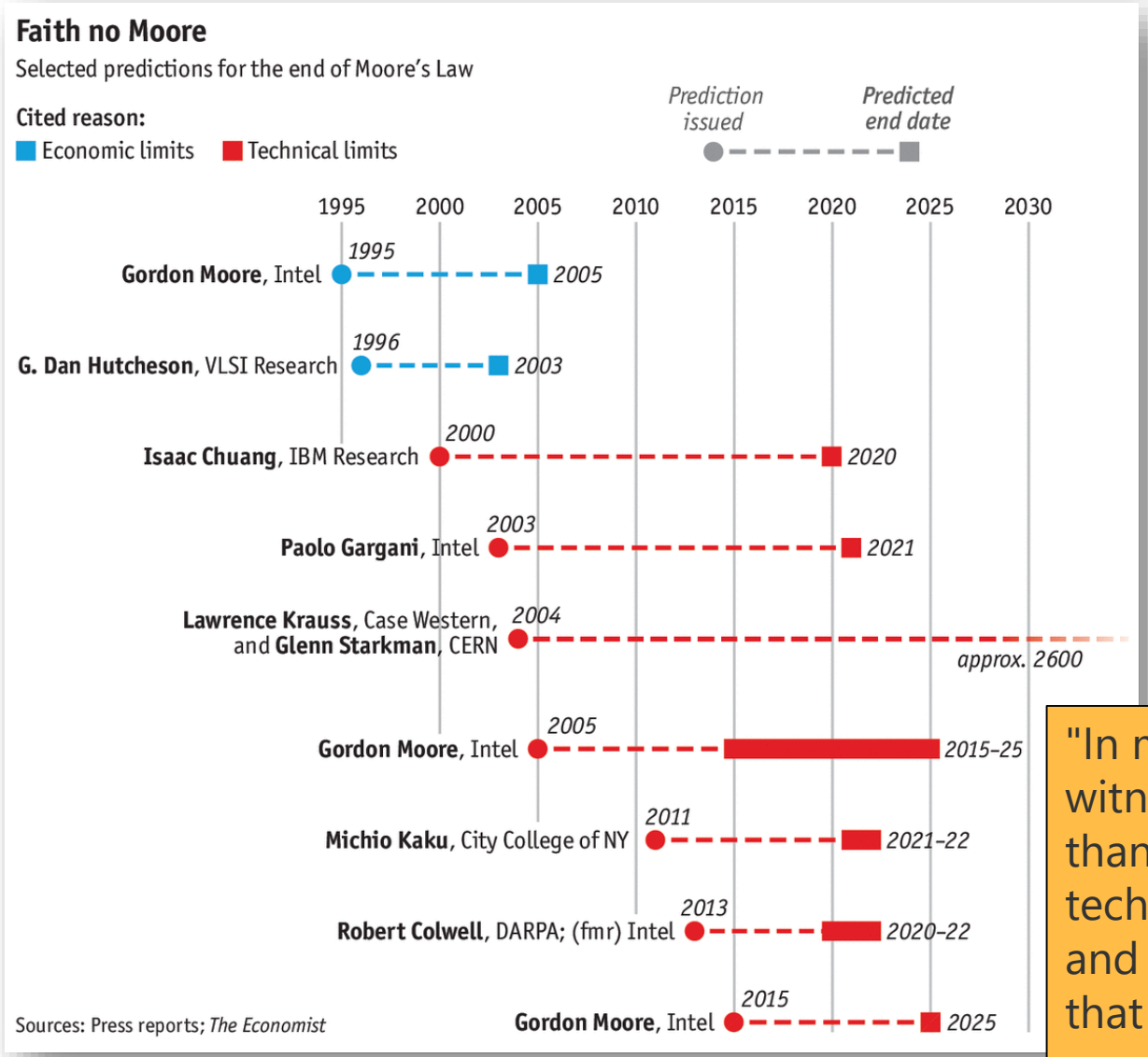
Moore's Law

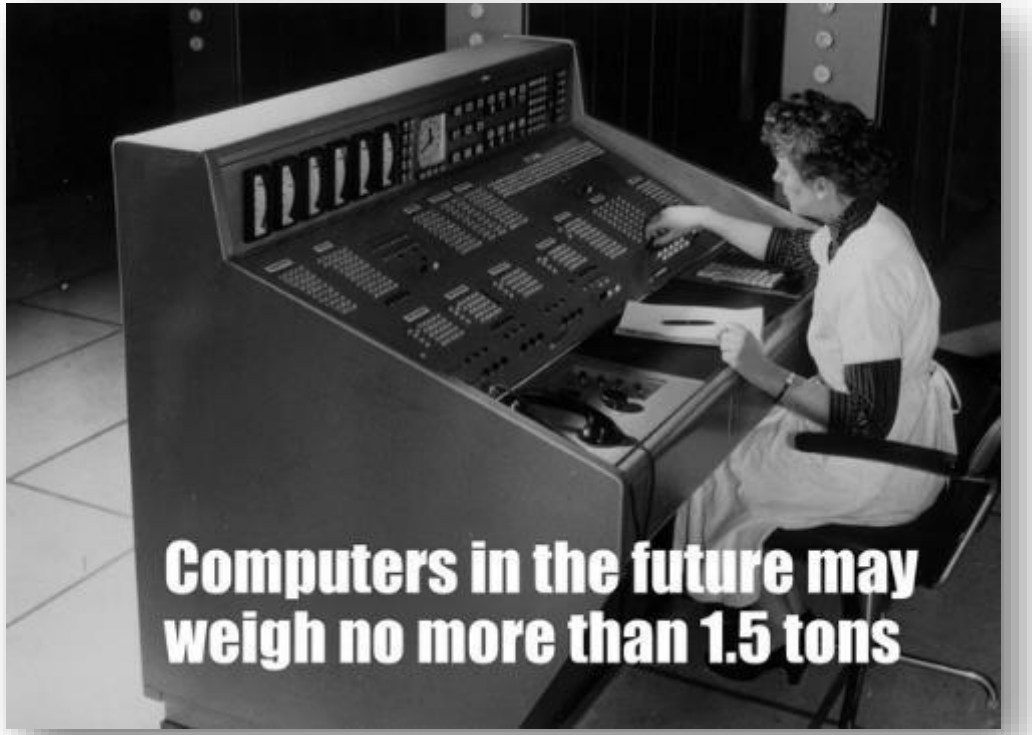


Computersciencezone.org



Reports of my death were greatly exaggerated

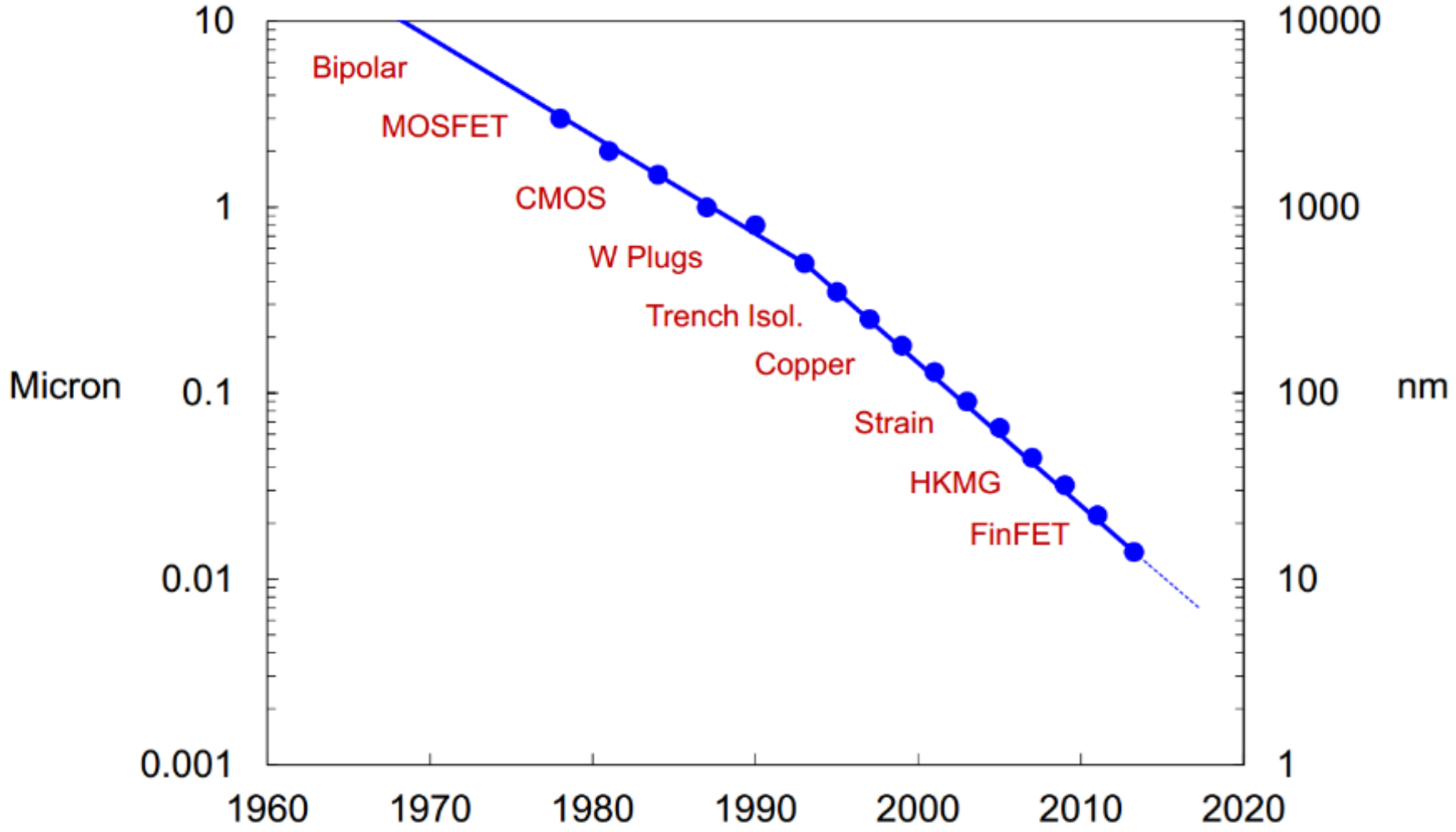




"In my 34 years in the semiconductor industry, I have witnessed the advertised death of Moore's Law no less than four times. As we progress from 14 nanometer technology to 10 nanometer and plan for 7 nanometer and 5 nanometer and even beyond, our plans are proof that Moore's Law is alive and well"

Bryan Krzanich, CEO Intel, April 2016

Technology supporting Moore's Law

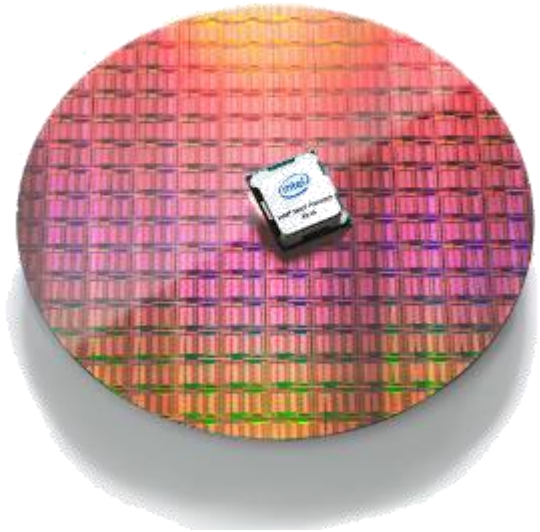


Process/device innovation has always been an indispensable part of scaling

Courtesy: Intel

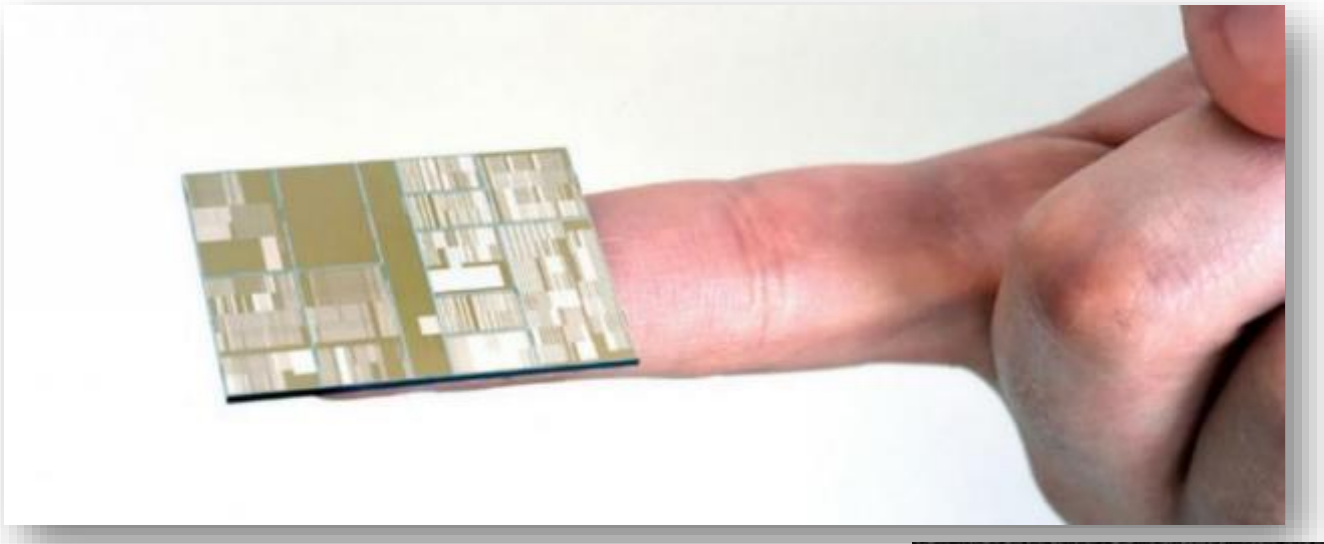
Moore's Law Today (2016)

Intel Xeon E5-2600 V4

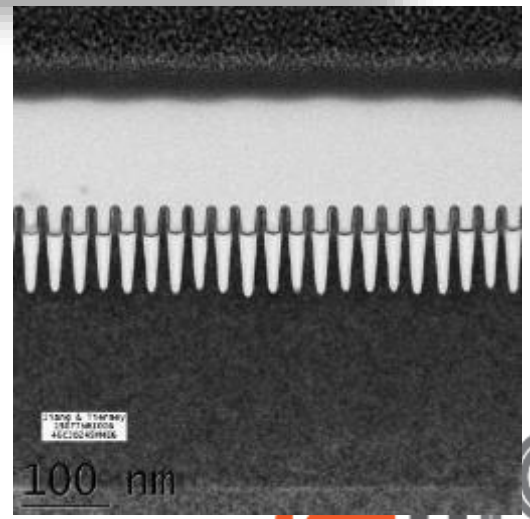


- 14nm "Broadwell"
- 22 Cores
- 2.2 GHz
- 55MB Cache
- 416 mm^2
- 7.2 Billion Transistors
- 456 mm^2 Die size
- Introduced March 31, 2016

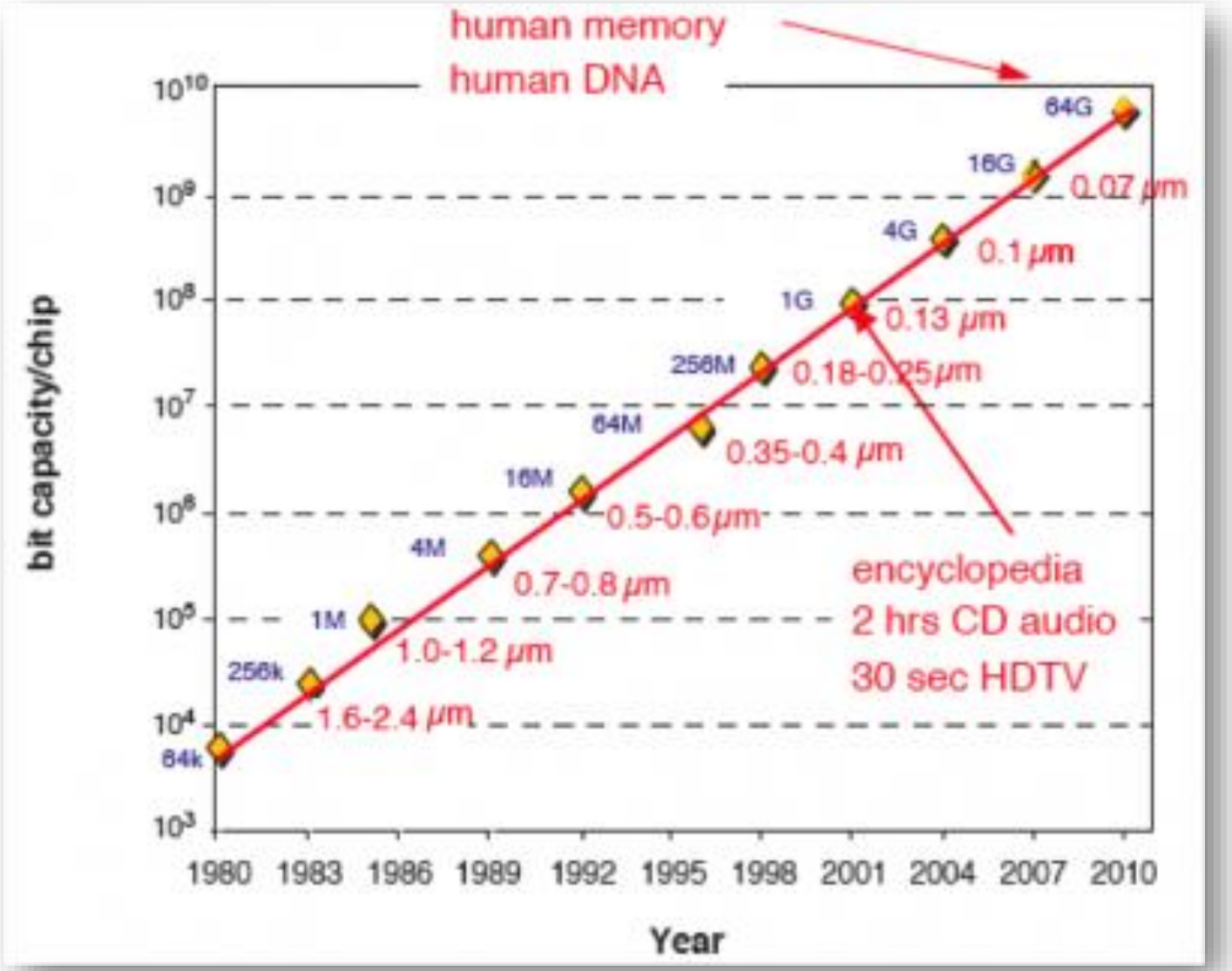
IBM 7nm Test Chip



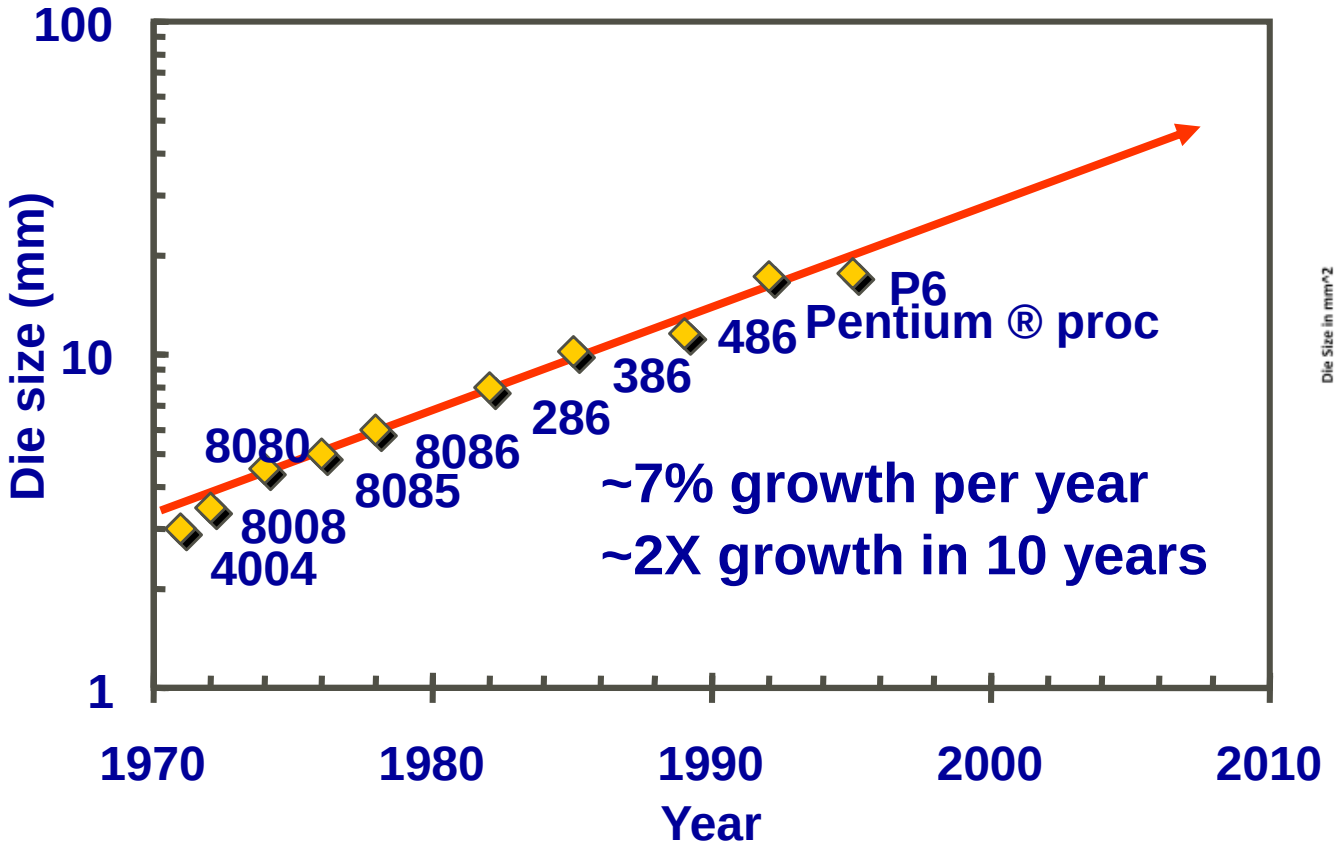
- 7nm
- EUV Photolithography
- SiGe channels
- Introduced July 2015

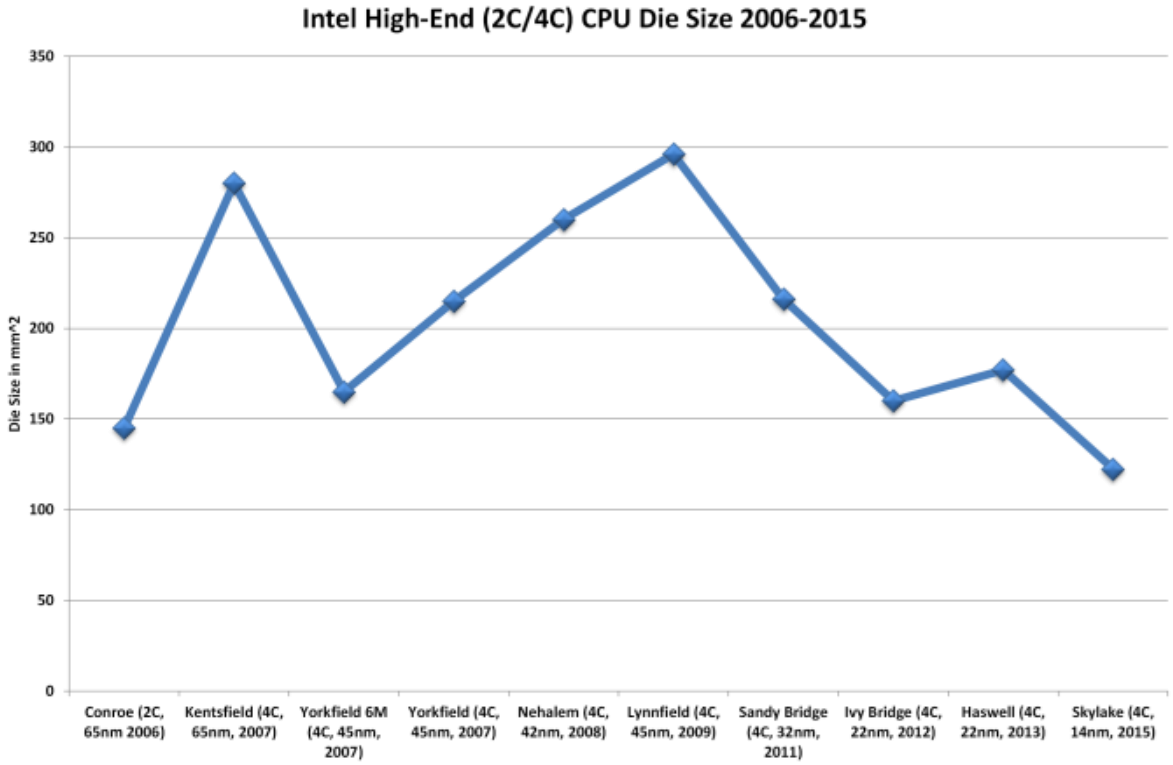


Evolution in Memory Complexity



Die Size Growth



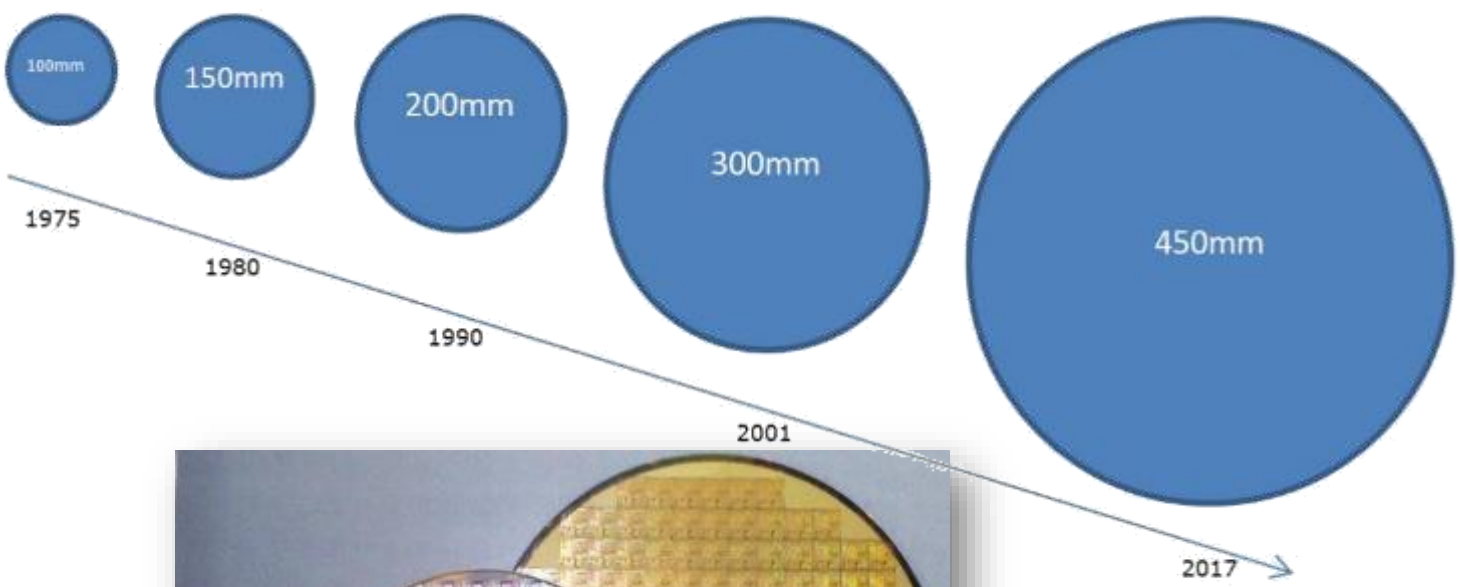


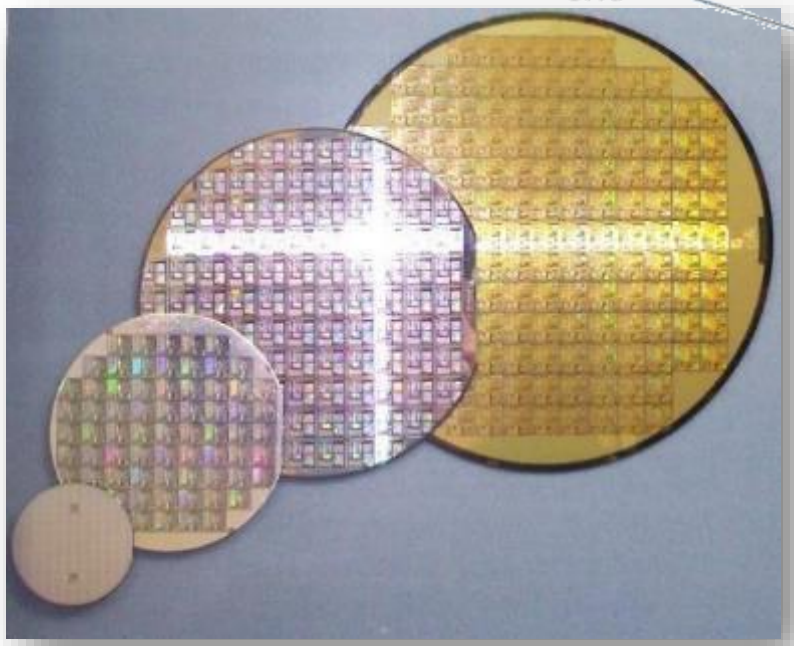
Apparently, that doesn't apply anymore…

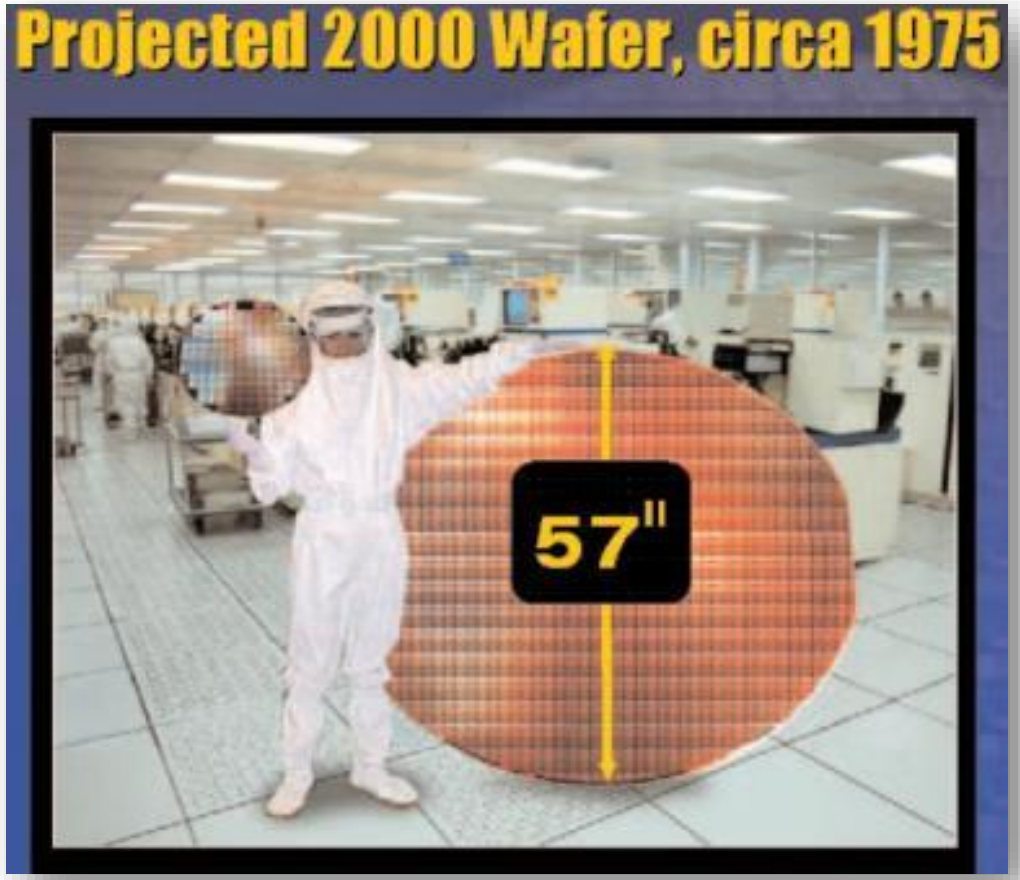
Die size grows by 14% to satisfy Moore's Law

Courtesy, Intel

Moore was not always accurate

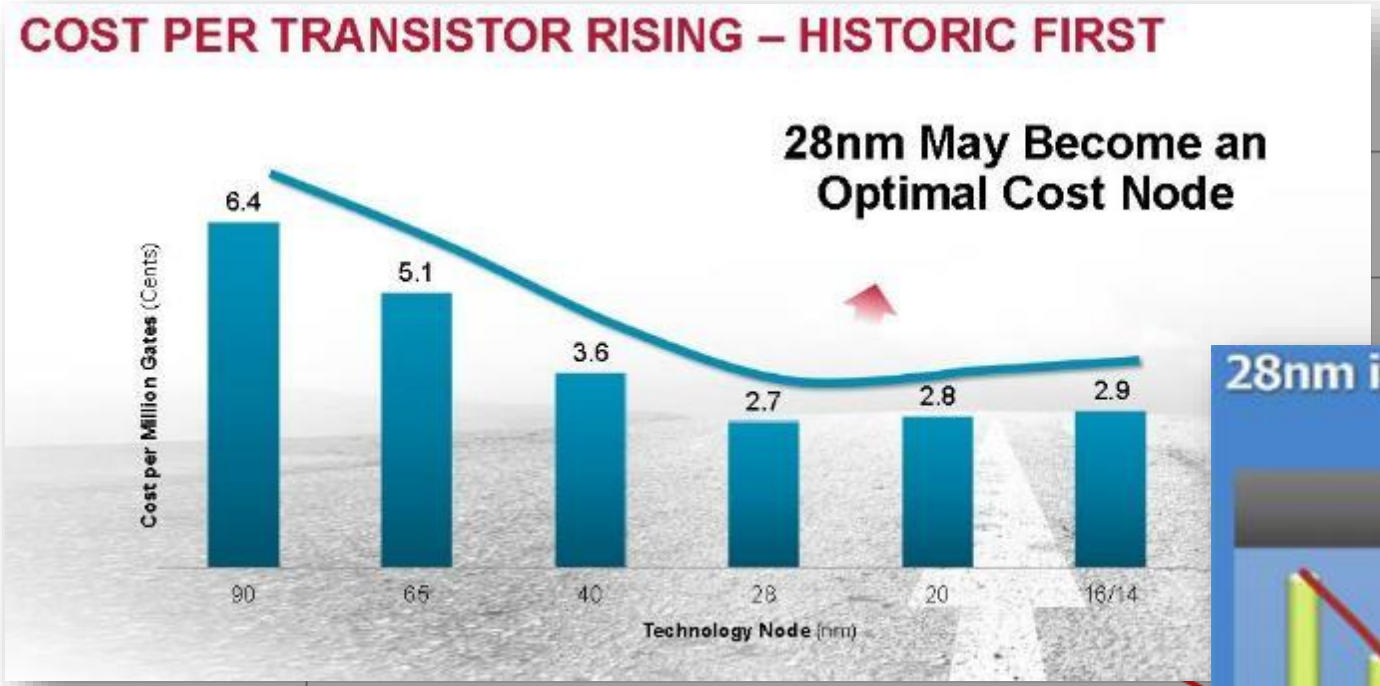


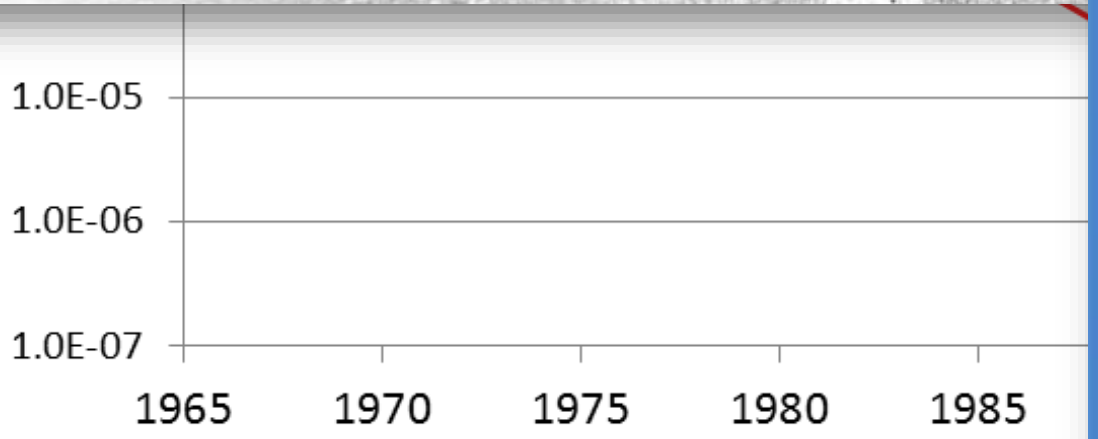




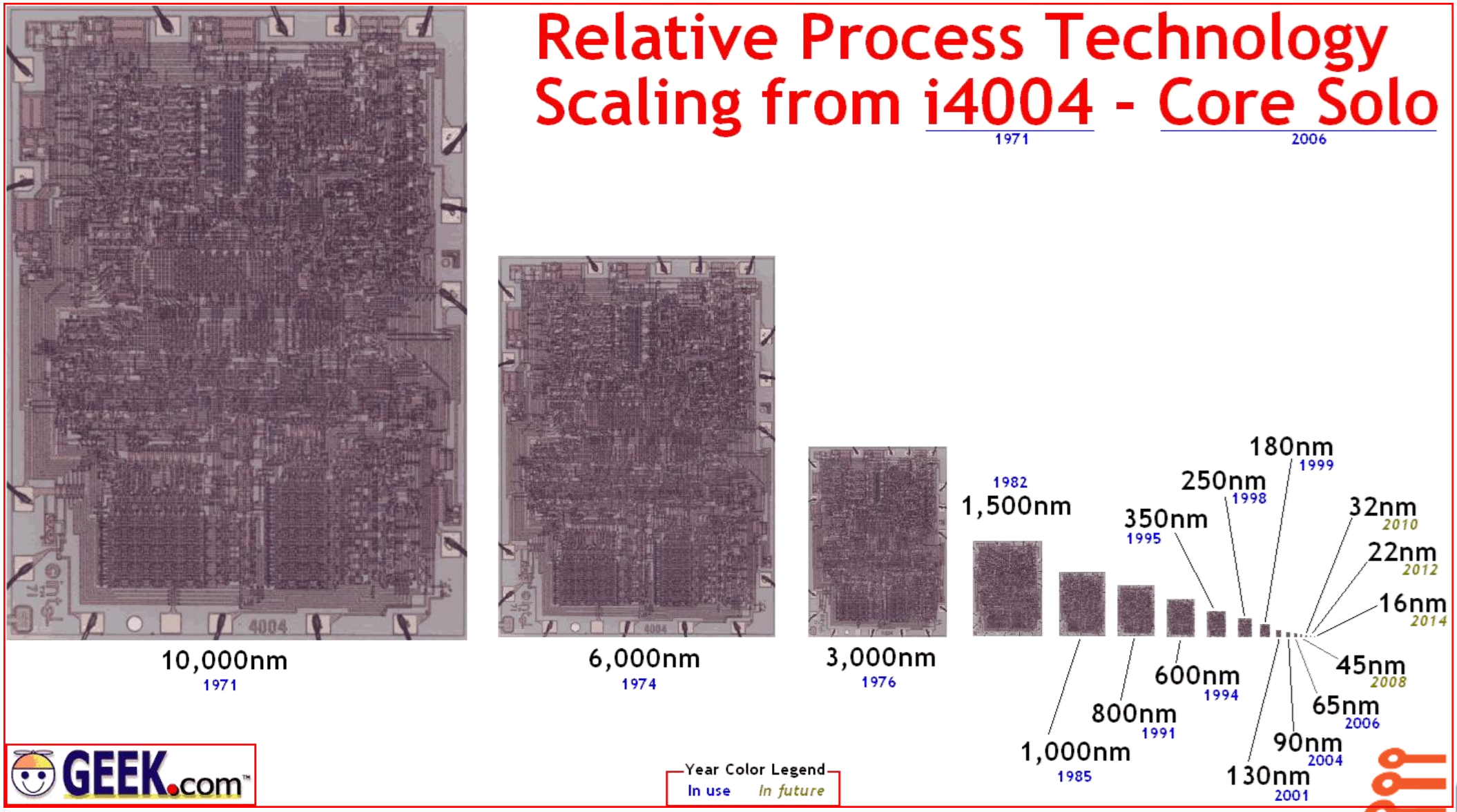
Teman's Law

Cost per Transistor





Scaling…


Relative Process Technology
Scaling from i4004 - Core Solo
1971
2006
10,000nm
1971
6,000nm
1974
3,000nm
1976
1982
1,500nm
1,000nm
1985
800nm
1991
600nm
1994
350nm
1995
250nm
1998
180nm
1999
130nm
2001
90nm
2004
65nm
2006
45nm
2008
32nm
2010
22nm
2012
16nm
2014
GEEK.com
Year Color Legend
In use
In future

Goals of Technology Scaling

- **Make things cheaper:**
  - Want to sell more functions (transistors) per chip for the same money
  - Build same products cheaper, sell the same part for less money
  - Price of a transistor **has to be reduced**
- **But also want to be faster, smaller, lower power**

Rabaey's Law of Playstations

Technology Scaling – Dennard's Law

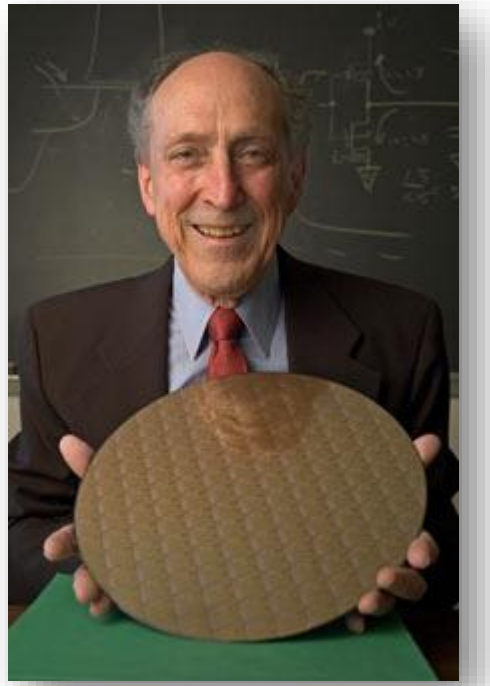
- **Benefits of scaling the dimensions by 30% (Dennard):**
  - Double transistor density
  - Reduce gate delay by 30%
    (increase operating frequency by 43%)
  - Reduce energy per transition by 65%
    (50% power savings @ 43% increase in frequency

- **Die size used to increase by 14% per generation**

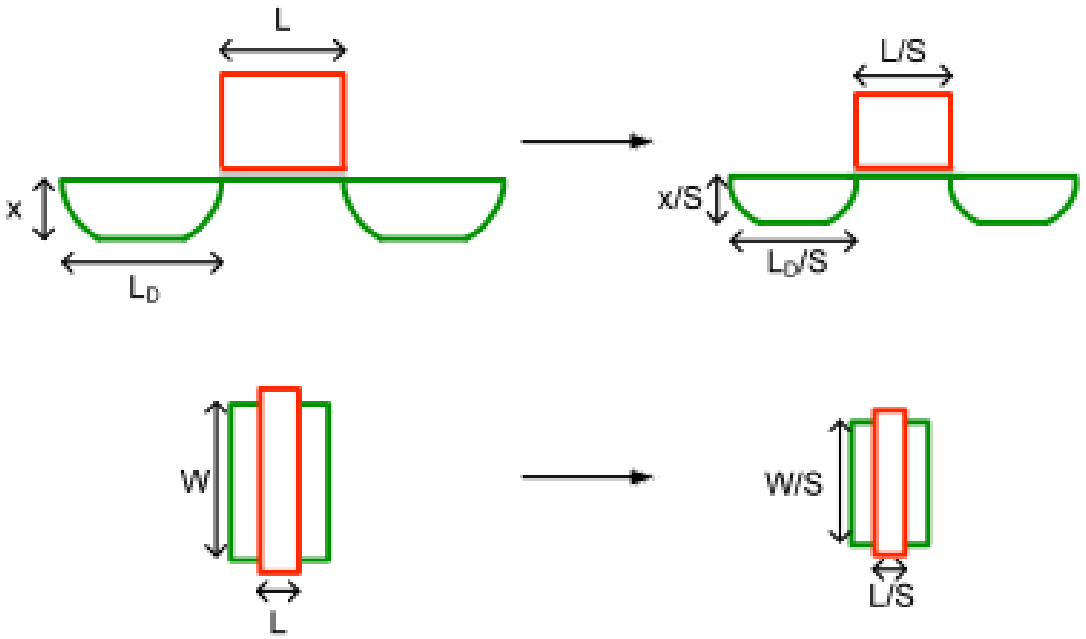
- **Technology generation spans 2-3 years**

Scaling Models

Dennard Scaling



- **In 1974, Robert Dennard of IBM described the MOS scaling principles that have accompanied us for forty years.**
- **As long as we scale all dimensions of a MOSFET by the same amount (*S*), we will arrive at better devices and lower cost:**
  - L – 1/S
  - W – 1/S
  - t_{ox} – 1/S
  - Na – S
  - Vdd – 1/S
  - V_T – 1/S



Reminder – our simple timing/power models

- **In our previous course, we developed the *unified model* for MOS transistor conduction:**

$$I_{DS} = K\left(V_{GT}V_{DSeff} - 0.5V_{DSeff}^2\right)\left(1+\lambda V_{DS}\right)$$

$$V_{DSeff} = \min\left(V_{GT}, V_{DS}, V_{DSAT}\right)$$

$$K = \mu_n C_{ox} W/L$$

$$C_{ox} = \varepsilon_{ox}/t_{ox}$$

$$I_{on} = K_n V_{GT}^2$$

$$R_{on} = V_{DD}/I_{on}$$

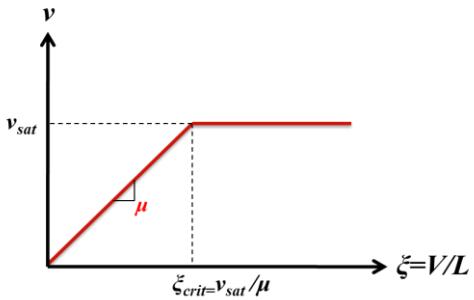
$$t_{pd} = R_{on} C_g$$

$$P_{dyn} = f \cdot C \cdot V_{DD}^2$$

Dennard (Full) Scaling for Long Transistors

$$L \propto S^{-1}$$
$$W \propto S^{-1}$$
$$t_{ox} \propto S^{-1}$$
$$V_{DD} \propto S^{-1}$$
$$V_T \propto S^{-1}$$
$$N_A \propto S$$

$$V_{DSat} = \xi_{crit} L$$



Property	Sym	Equation	Calculation	Scaling	Good?
Oxide Capacitance	C_{ox}	ε_{ox}/t_{ox}	$1/S^{-1}$	S	
Device Area	A	$W \cdot L$	$S^{-1} \cdot S^{-1}$	$1/S^2$	☺
Gate Capacitance	C_g	$C_{ox} \cdot W \cdot L$	$S \cdot S^{-1} \cdot S^{-1}$	$1/S$	☺
Transconductance	K_n	$\mu_n C_{ox} W/L$	$S \cdot S^{-1}/S^{-1}$	S	☺
Saturation Current	I_{on}	$K_n V_{DSat}(W_{GT} - \frac{1}{2}V_{DSat})$	$S \cdot S^{-1} \cdot S^{-1}$	$1/S$	
On Resistance	R_{on}	V_{DD}/I_{on}	S^{-1}/S^{-1}	1	
Intrinsic Delay	t_{pd}	$R_{on} C_g$	$1 \cdot S^{-1}$	$1/S$	☺
Power	P_{av}	$f \cdot C \cdot V_{DD}^2$	$S \cdot S^{-1} \cdot S^{-2}$	$1/S^2$	☺ ☺
Power Density	PD	P_{av}/A	S^{-2}/S^{-2}	1	☺

Dennard Scaling

- **This previous slide showed the principal that has led to scaling for the last 50 years.**
  - Assume that we scale our process by 30% every generation.

$$1/S = 0.7 \longrightarrow S = \sqrt{2}$$

  - Therefore, if the area scales by $1/S^2 = 1/2$, our ***die size*** goes down by **2X** every generation!
  - In addition, our ***speed*** goes up by 30%!
  - And our ***power*** also gets cut in half, without any increase in power density.

Sorry… I couldn't resist!

- **We have hit one of those rare win-win *free lunch* situations!**

But what if we want more speed?

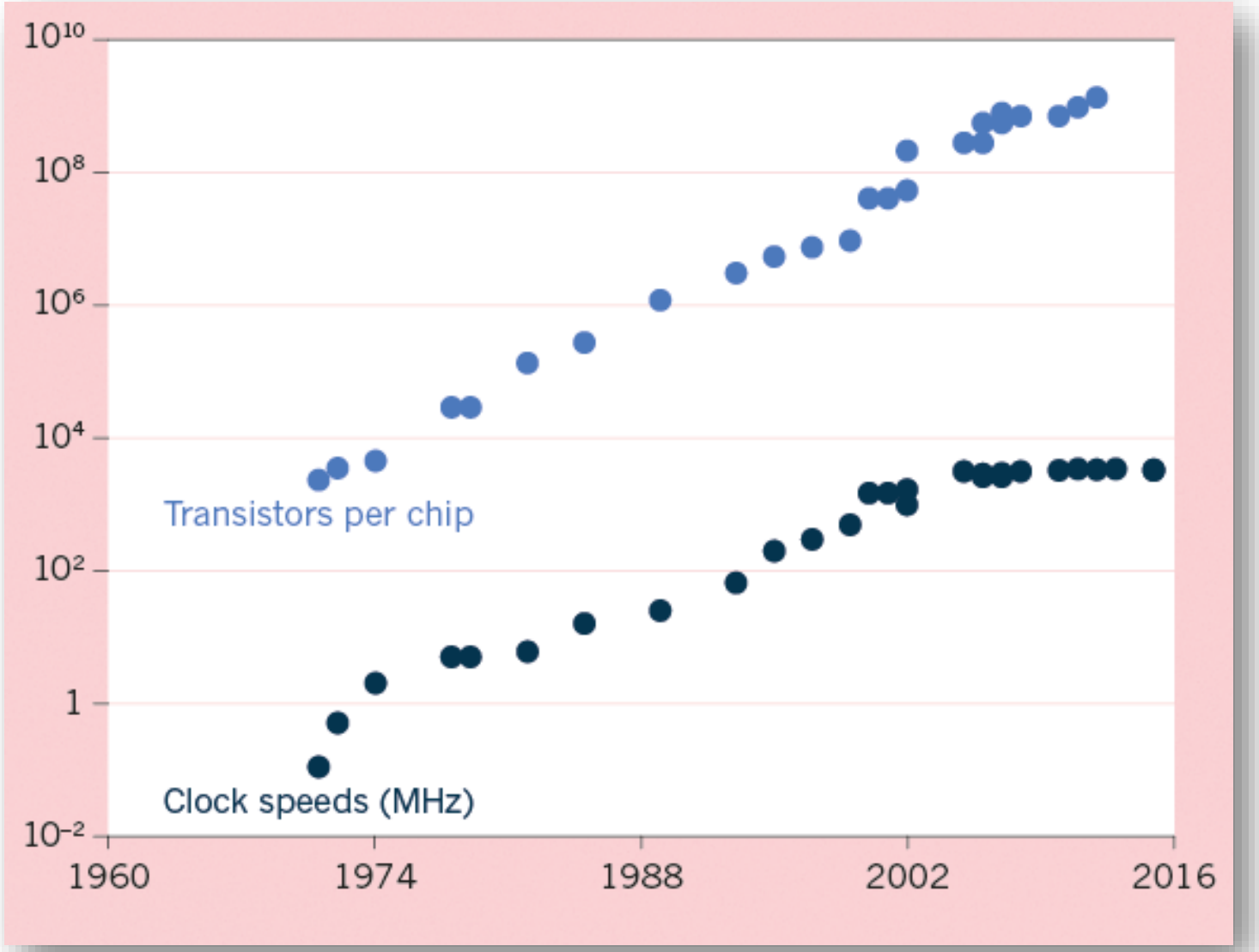
- We saw that $$t_{pd} \propto C_g \cdot V_{DD} / I_{on}$$

- We can aggressively increase the speed by keeping the voltage constant.

$$I_{on} \propto K_n V_{GT}^2 \propto S \Rightarrow t_{pd} \propto S^{-1} \cdot 1/S = 1/S^2$$

- This led to the ***Fixed Voltage Scaling Model,*** which was used until the 1990s (V_{DD}=5V)

Moore's Law in Frequency



Nature

Fixed Voltage Scaling

$V_{DD} \propto 1$

$L \propto S^{-1}$

$W \propto S^{-1}$

$t_{ox} \propto S^{-1}$

$V_T \propto S^{-1}$

$N_A \propto S$

Property	Sym	Equation	Calculation	Scaling	Good?
Oxide Capacitance	C_{ox}	ε_{ox}/t_{ox}	$1/S^{-1}$	S	
Device Area	A	$W \cdot L$	$S^{-1} \cdot S^{-1}$	$1/S^2$	☺
Gate Capacitance	C_g	$C_{ox} \cdot W \cdot L$	$S \cdot S^{-1} \cdot S^{-1}$	$1/S$	☺
Transconductance	K_n	$\mu_n C_{ox} W/L$	$S \cdot S^{-1}/S^{-1}$	S	☺
Saturation Current	I_{on}	$K_n V_{GT}^2$	$S \cdot 1$	S	☺
On Resistance	R_{on}	V_{DD}/I_{on}	$1/S$	$1/S$	
Intrinsic Delay	t_{pd}	$R_{on} C_g$	$S^{-1} \cdot S^{-1}$	$1/S^2$	☺ ☺
Power	P_{av}	$f \cdot C \cdot V_{DD}^2$	$S^2 \cdot S^{-1} \cdot 1$	S	✖
Power Density	PD	P_{av}/A	S/S^{-2}	S^3	✖ ✖

Fixed Voltage Scaling – Short Channel

- **What happens with velocity saturated devices?**

$$I_{on} \propto K_n V_{DSat}\left(V_{GT} - V_{DSat}\right) \propto S \cdot S^{-1} \cdot 1 = 1$$

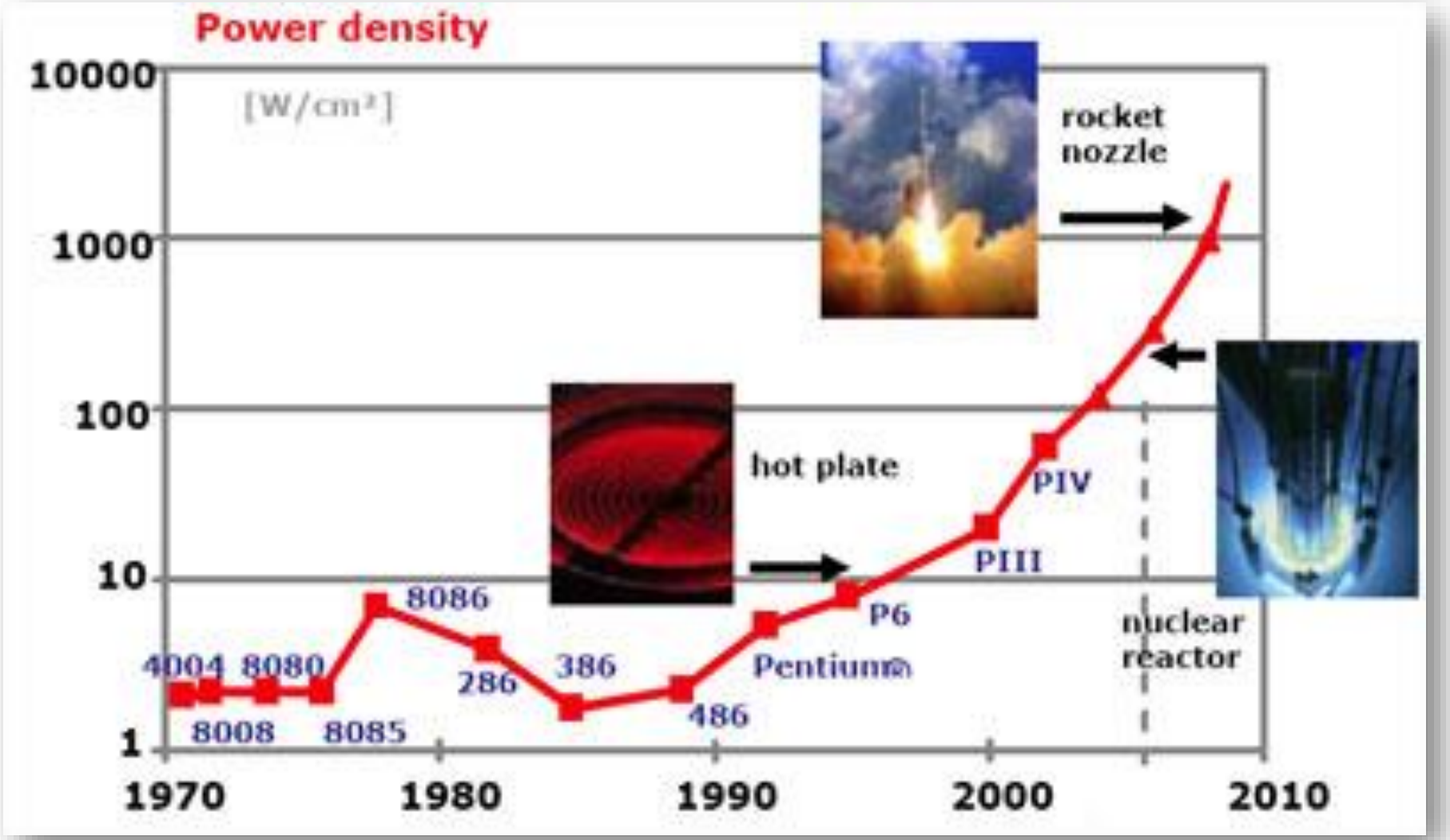
- **So the on current doesn't increase leading to less effective speed increase.**

$$t_{pd} \propto R_{on} C_g \propto 1 \cdot S^{-1} = 1/S$$

- **The power density still increases quadratically!**

$$PD \propto fCV_{DD}^2 / A \propto S \cdot S^{-1} \cdot 1/S^{-2} = S^2$$

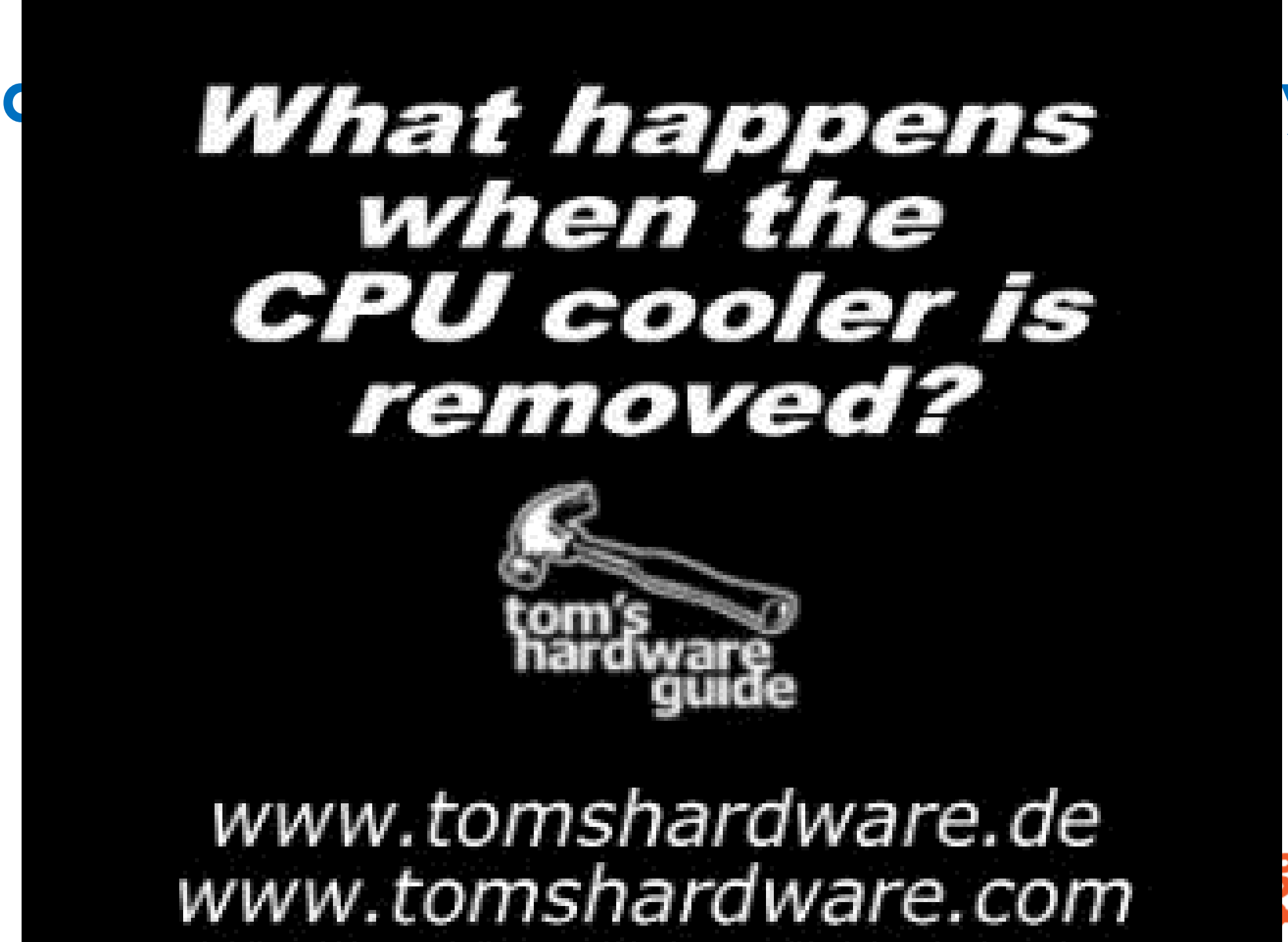
Power density (2004 expectation)



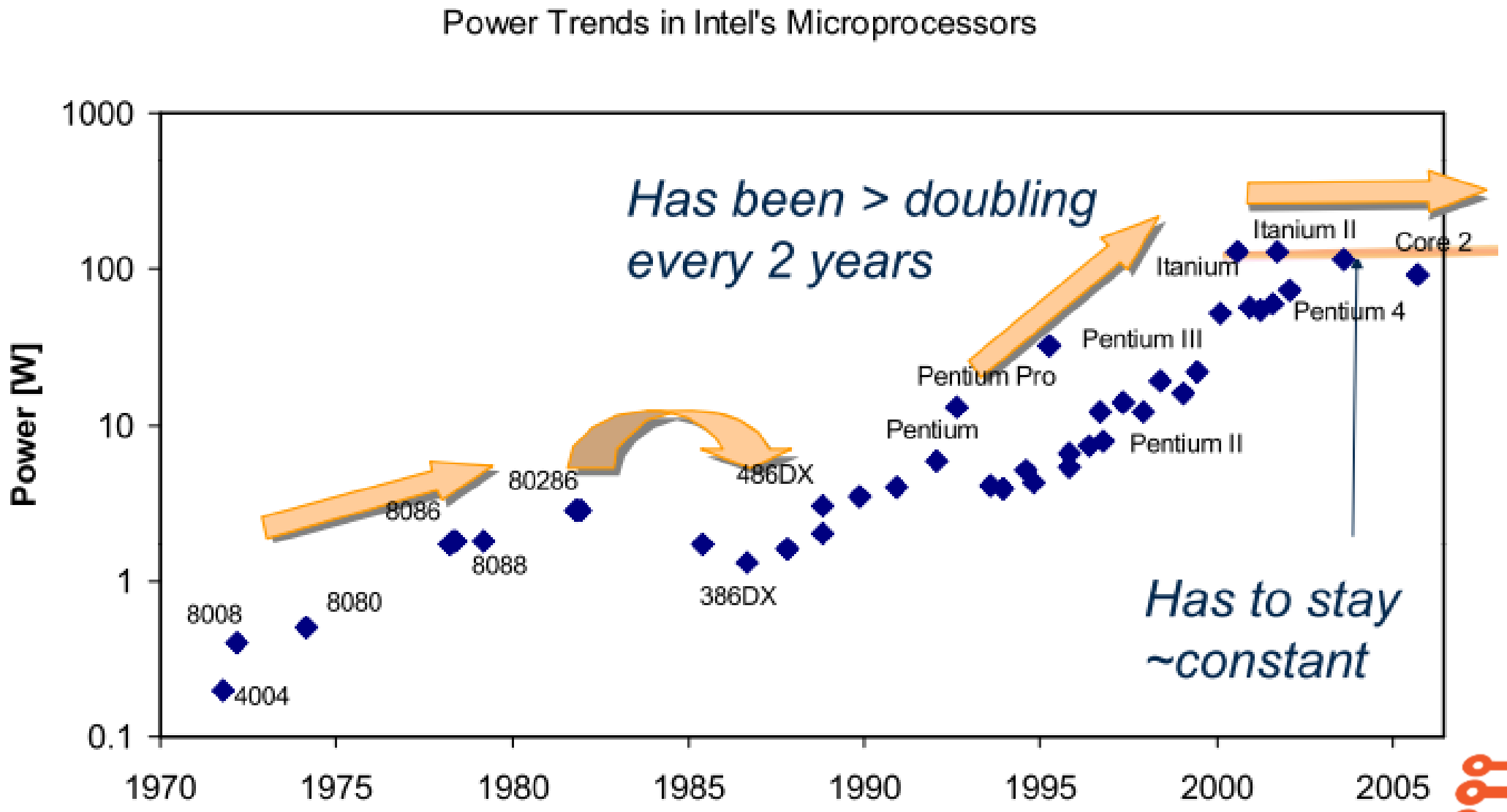
The Power Density Crisis

Patrick Gelsinger, Intel
ISSCC 2001

Who y…?


What happens when the CPU cooler is removed?
tom's hardware guide
www.tomshardware.de
www.tomshardware.com

What actually happened?



Technology Scaling Models

- **Fixed Voltage Scaling**
  - Supply voltages have to be similar for all devices (one battery)
  - Only device dimensions are scaled.
  - 1970s-1990s
- **Full "Dennard" Scaling (Constant Electrical Field)**
  - Scale both device dimensions and voltage by the same factor, *S*.
  - Electrical fields stay constant, eliminates breakdown and many secondary effects.
  - 1990s-2005
- **General Scaling –**
  - Scale device dimensions by *S* and voltage by *U*.
  - Now!

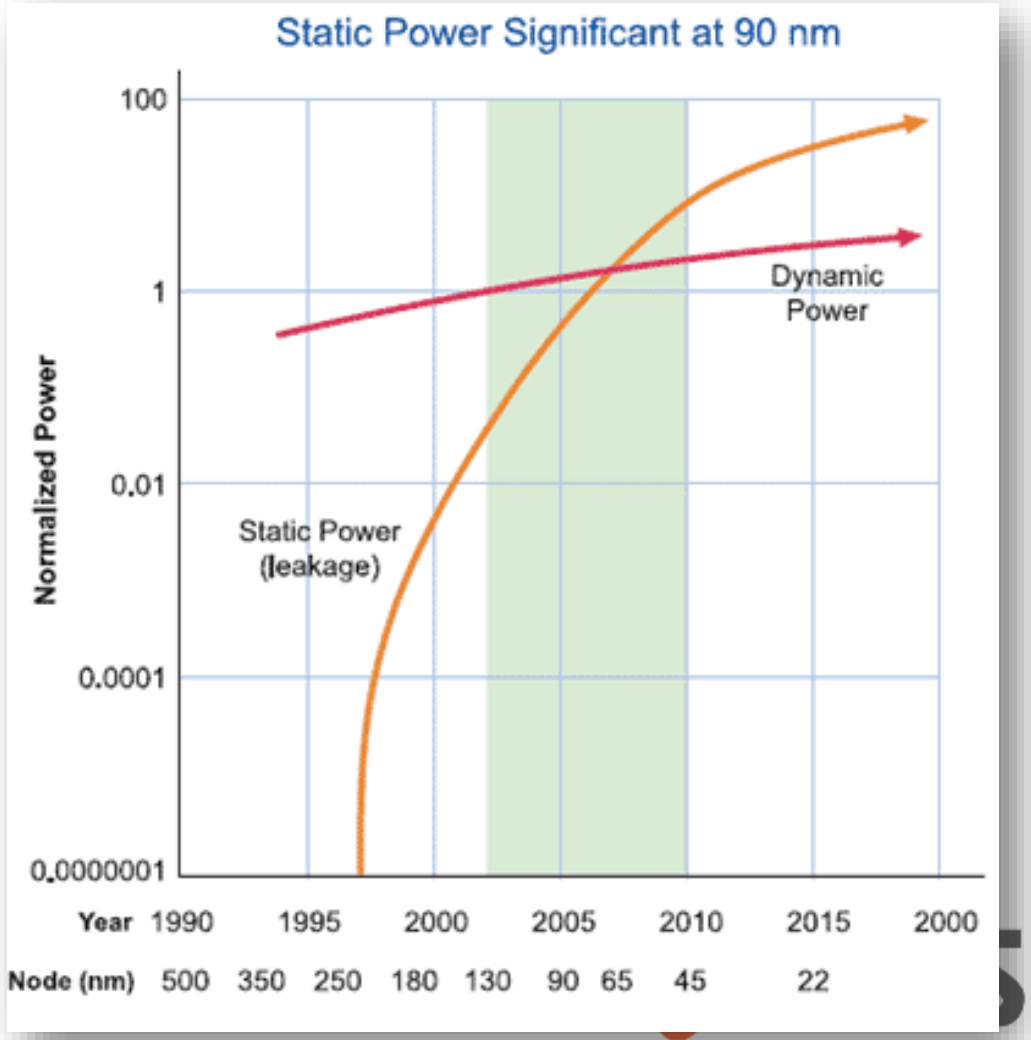
How about Leakage Power?

- Later in the semester, we will see that the off current is exponentially dependent on the threshold voltage.

$$I_{off} \propto e^{-V_T / n\phi_T}$$

- In the case of *Full Scaling*, the leakage current *increases exponentially* as V_T is decreased!

- Since the 90nm node, static power is one of the major problems in ICs.



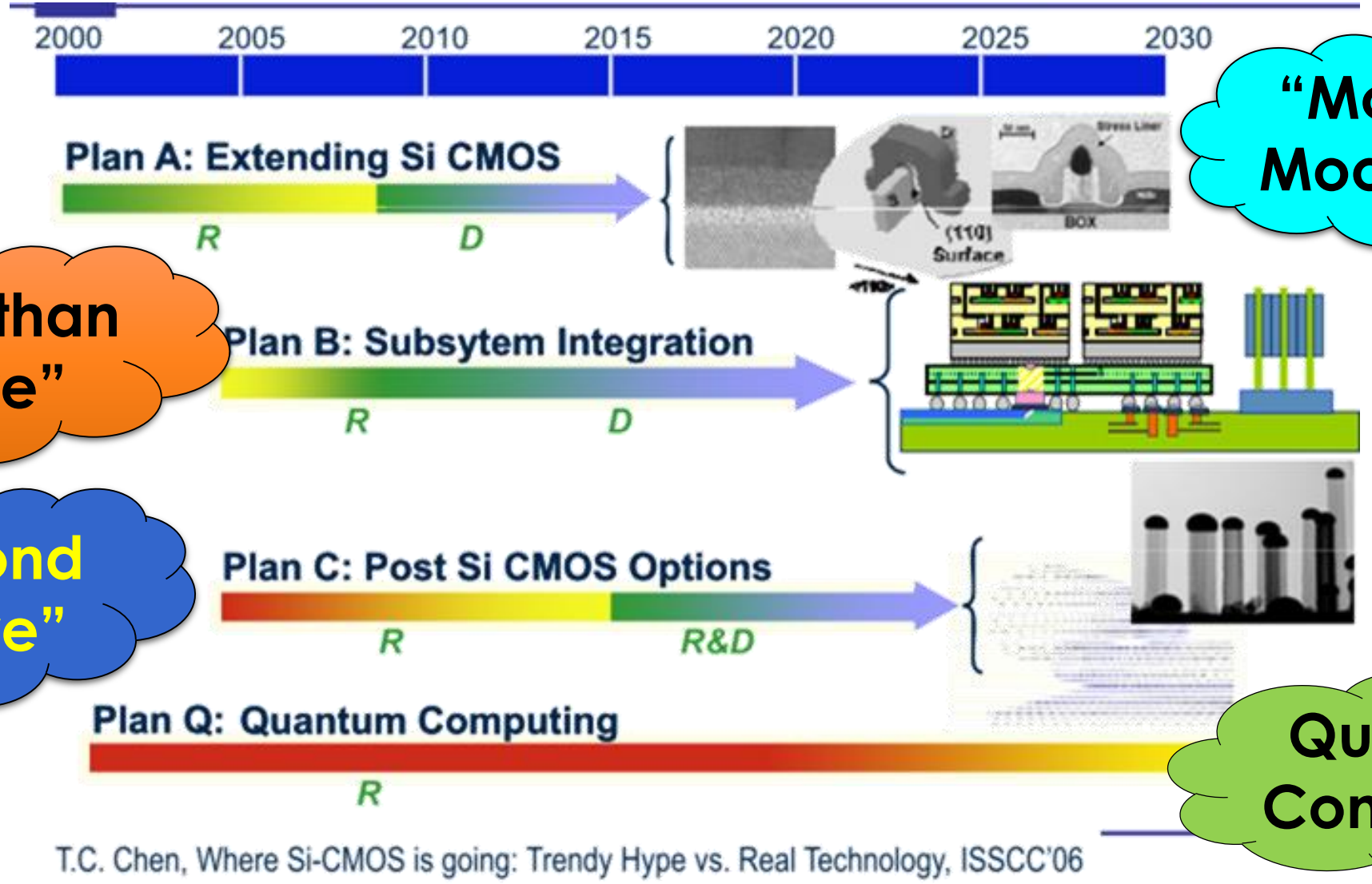
Current and Future Trends

ITRS

- **International Technology Roadmap for Semiconductors**

Year	2009	2012	2015	2018	2021
Feature size (nm)	34	24	17	12	8.4
L_{gate} (nm)	20	14	10	7	5
V_{DD} (V)	1.0	0.9	0.8	0.7	0.65
Billions of transistors/die	1.5	3.1	6.2	12.4	24.7
Wiring levels	12	12	13	14	15
Maximum power (W)	198	198	198	198	198
DRAM capacity (Gb)	2	4	8	16	32
Flash capacity (Gb)	16	32	64	128	256

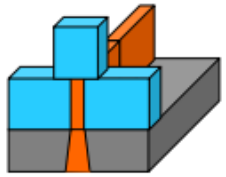
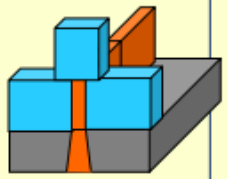
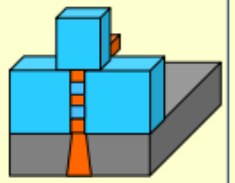
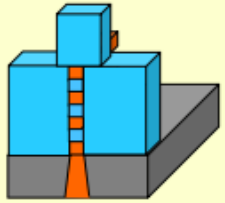
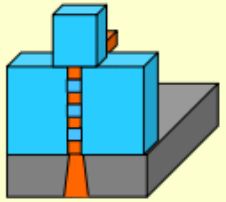
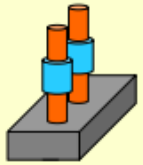
Technology Strategy Roadmap

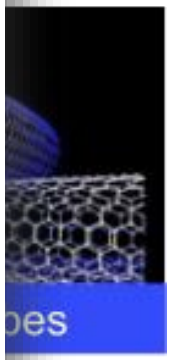


T.C. Chen, Where Si-CMOS is going: Trendy Hype vs. Real Technology, ISSCC'06

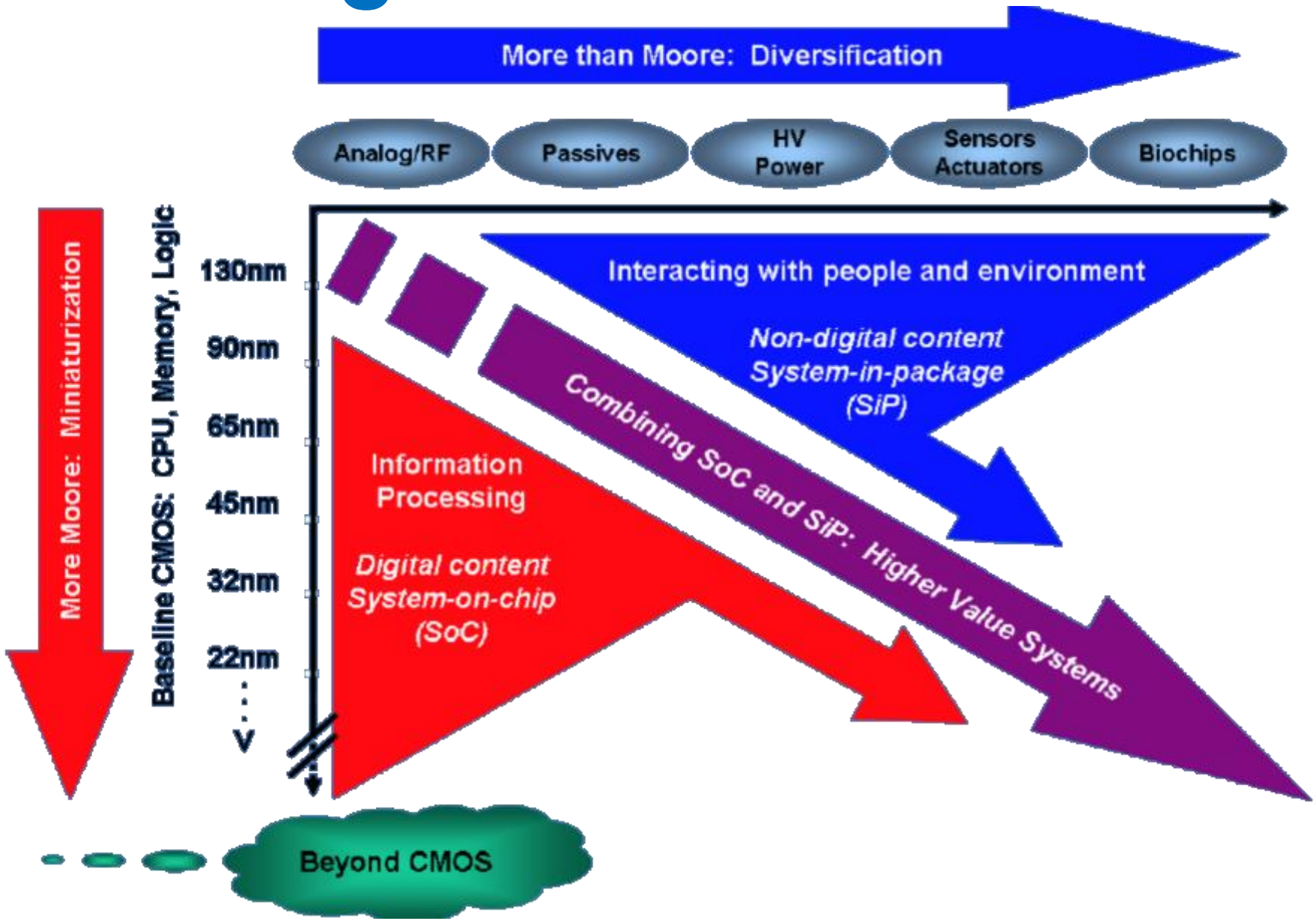
When will Moore's Law End?

IMEC view of logic technology roadmap

Early production	2014 N14	2016 N10	2017-2018 N7	2018-2019 N5	>2020 N3
	 *FinFET*	 *FinFET*	 *FinFET*  *Lateral nanowire*	 *Lateral nanowire*	 *Lateral nanowire*  *Vertical nanowire*
Vdd (V)	0.8	*0.8-0.7*	*0.7-0.6*	*0.7-0.5*	*0.6-0.5*
Gate Pitch (nm)	70-90, 193i	*52-64, 193i*	*36-46, 193i*	*26-36, EUV, 193i*	*18-28, EUV, 193i*
Device	FinFET	*FinFET*	*FinFET {lateral NW}*	*Lateral NW*	*Lateral NW {verical NW}*
Channel nfet/pfet	Si / Si	*Si / Si {SiGe}*	*Si / SiGe*	*Si/ SiGe*	*High mobility*
Metal Pitch (nm)	52-64, 193i	*36-48, 193i*	*26-40, EUV, 193i*	*18-28, EUV, 193i*	*10-20, EUV, 193i*
Low k dielectric	2.5-2.4	*2.5-2.3 {airgaps 1.3x}*	*2.4-2.1 airgaps 1.3x*	*2.4-1.9 airgaps 1.3x*	*2.4-1.9 airgaps 1.3x*
Metallization	TaN/Ta or CuMn ECD Cu	*TaN/Co or Ru ECD Cu*	*MnN/Co or Ru ECD Cu*	*Adv. Barrier/liner ELD or CVD Cu*	*Alternative metals*



Current Strategies



Further Reading

- J. Rabaey, "*Digital Integrated Circuits*" 2003, Chapter 1.3
- E. Alon, Berkeley *EE-141*, Lecture 2 (Fall 2009)
  http://bwrc.eecs.berkeley.edu/classes/icdesign/ee141_f09/

- …a number of years of experience!