CSAIL

MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

FROM

MODERN ALGORITHMS WORKSHOP

Parallel Algorithms

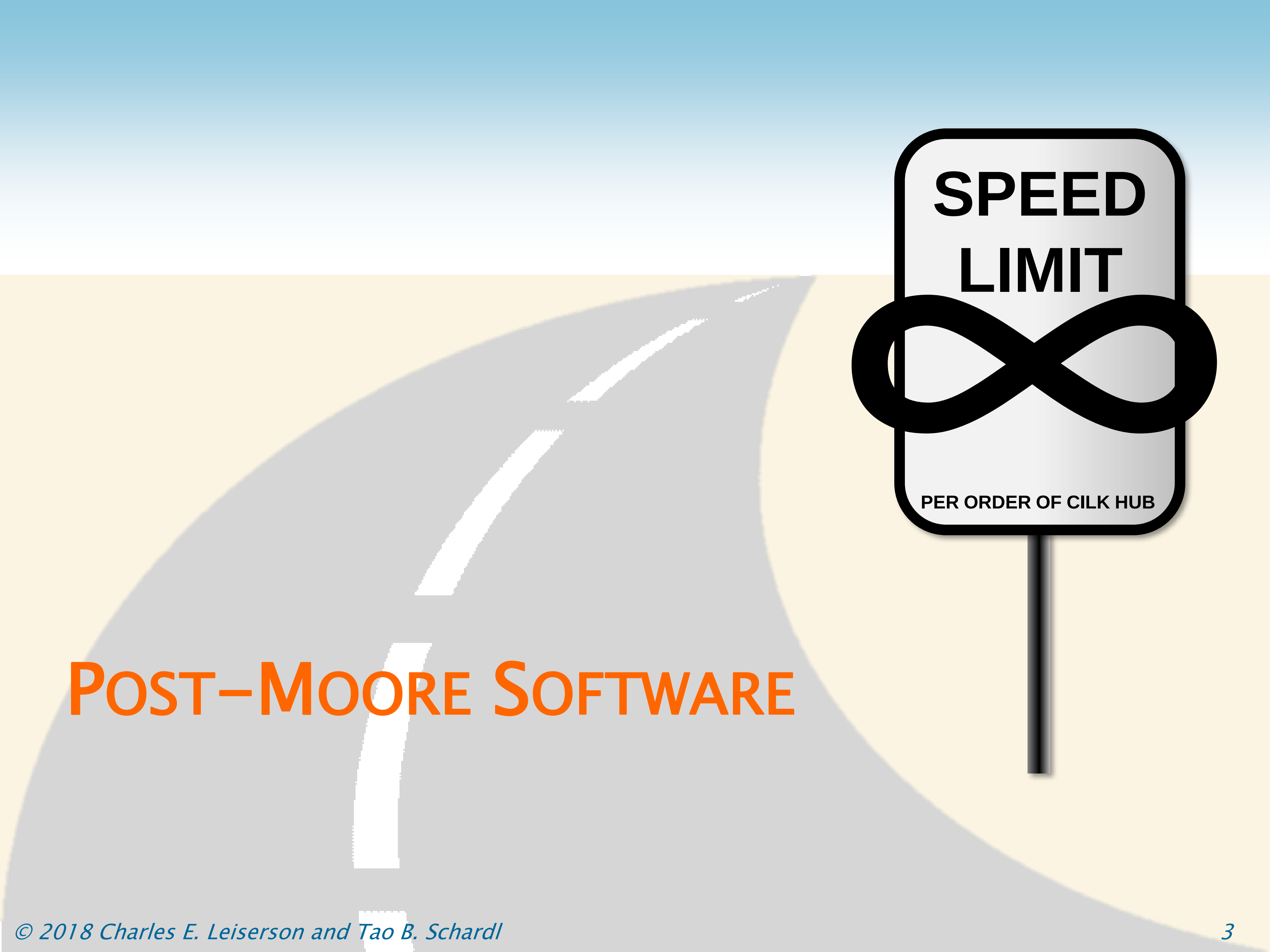
Prof. Charles E. Leiserson

Dr. Tao B. Schardl

September 19, 2018

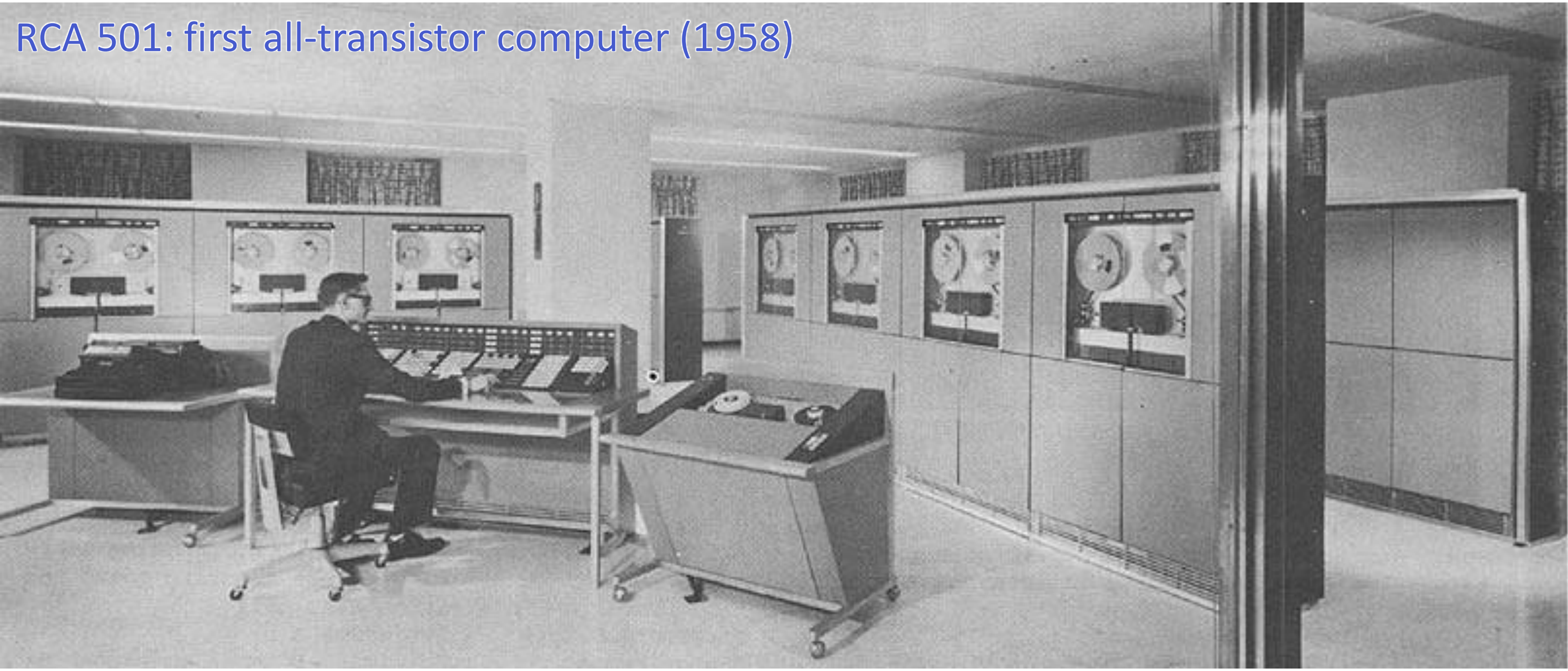
Outline

- Introduction
- Cilk Model
- Detecting Nondeterminism
- What Is Parallelism?
- Scheduling Theory Primer
- *Lunch Break*
- Analysis of Parallel Loops
- Case Study: Matrix Multiplication
- Case Study: Jaccard Similarity
- Post-Moore Software



Post-Moore Software

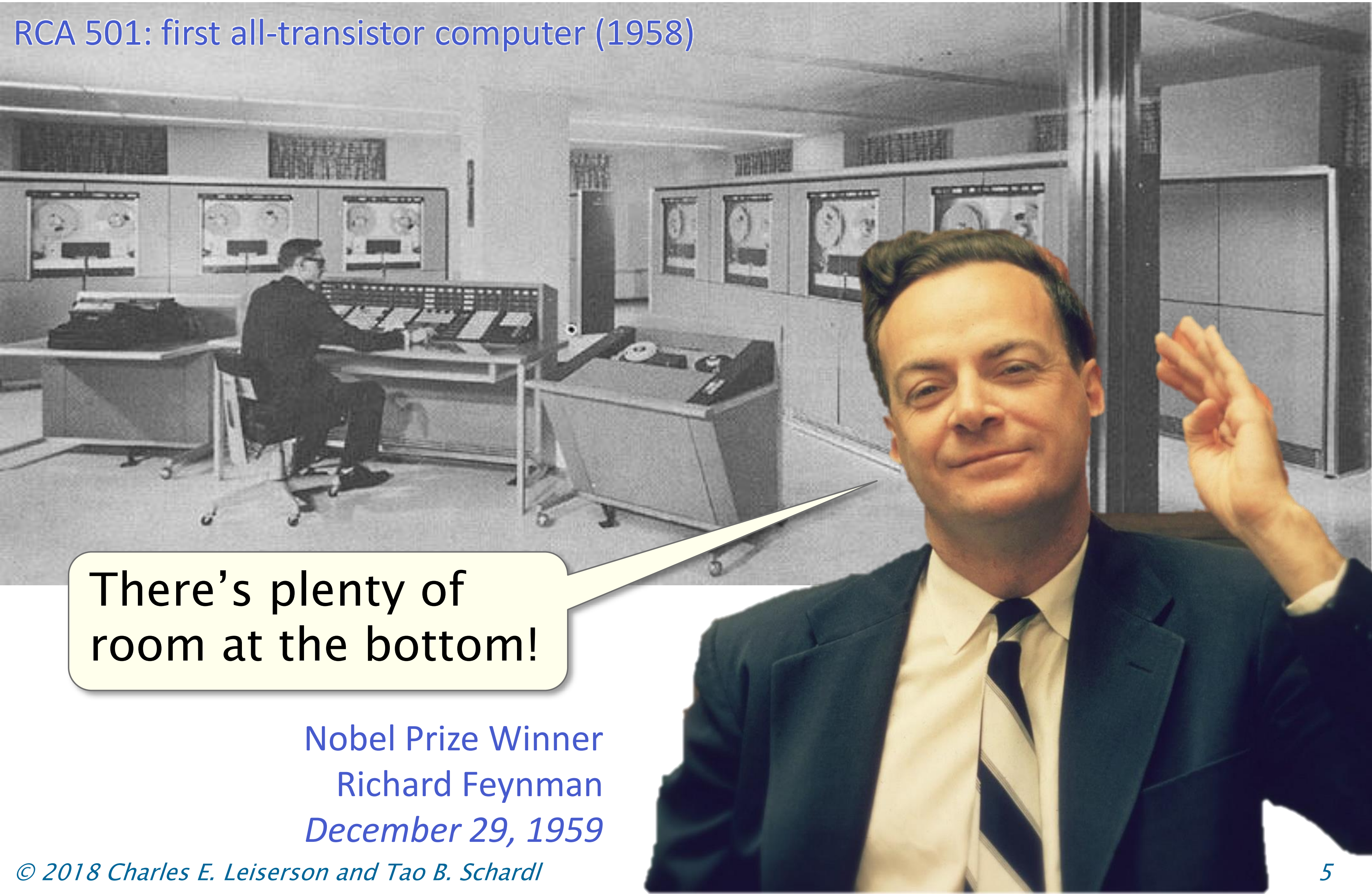
The Early Days of Computing



RCA 501: first all-transistor computer (1958)

The Early Days of Computing

RCA 501: first all-transistor computer (1958)



Nobel Prize Winner
Richard Feynman
December 29, 1959

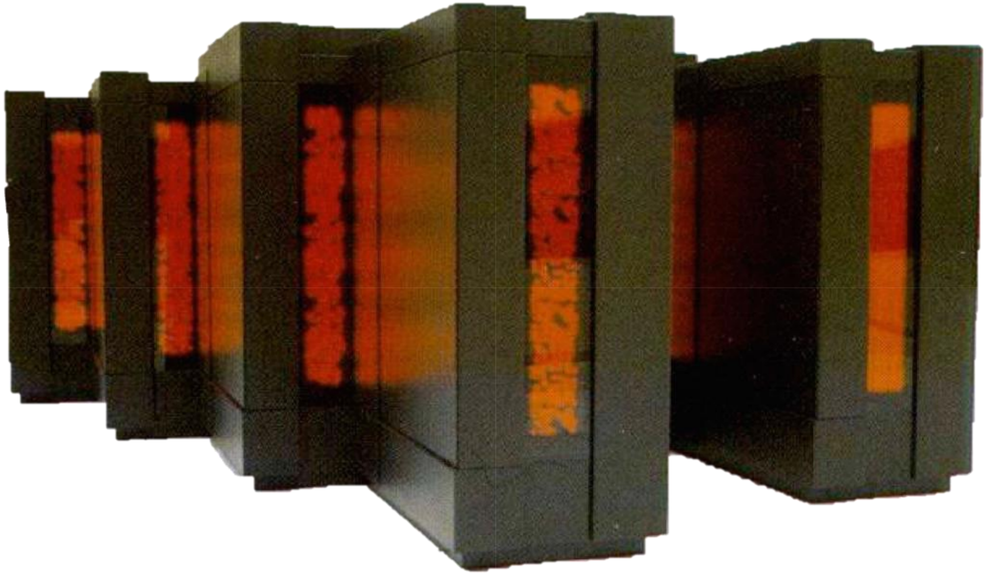
Moore’s Law

Moore’s Law is an economic and technology trend originally articulated in 1965 by Intel founder Gordon Moore.

The trend was christened “MOORE’S LAW” by Caltech professor Carver Mead in 1975.

The “Popular” Moore’s Law

The popular conception of MOORE’S LAW is that the cost of computing drops exponentially year by year, which is actually an implication of the “real” MOORE’S LAW.



≈

Connection Machine CM-5

- 60 GFLOPS on LINPACK
- $47 Million in 1993

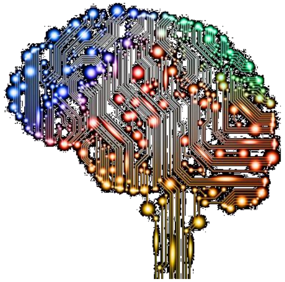
Apple 15” MacBook Pro

- 120 GFLOPS on LINPACK
- $2799 in 2018

Had Moore's Law Ended 15 Years Ago, We Would Not Be Enjoying…

Electric Cars

Innovative Game Technology



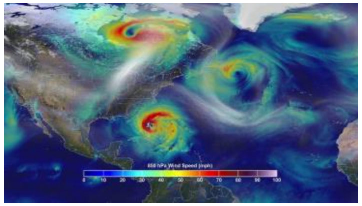
Deep Learning Applications

Digital Photography

High-Resolution Medical Imaging

Electronic Monitors

Inexpensive Robots



Accurate Weather Prediction

Smart Phones

Wearable Cameras

Cloud Computing

Tablets

Minions

Smart Appliances

Etc.

The “Original” Moore’s Law

The complexity for minimum component costs has increased at a rate of roughly a factor of two per year.

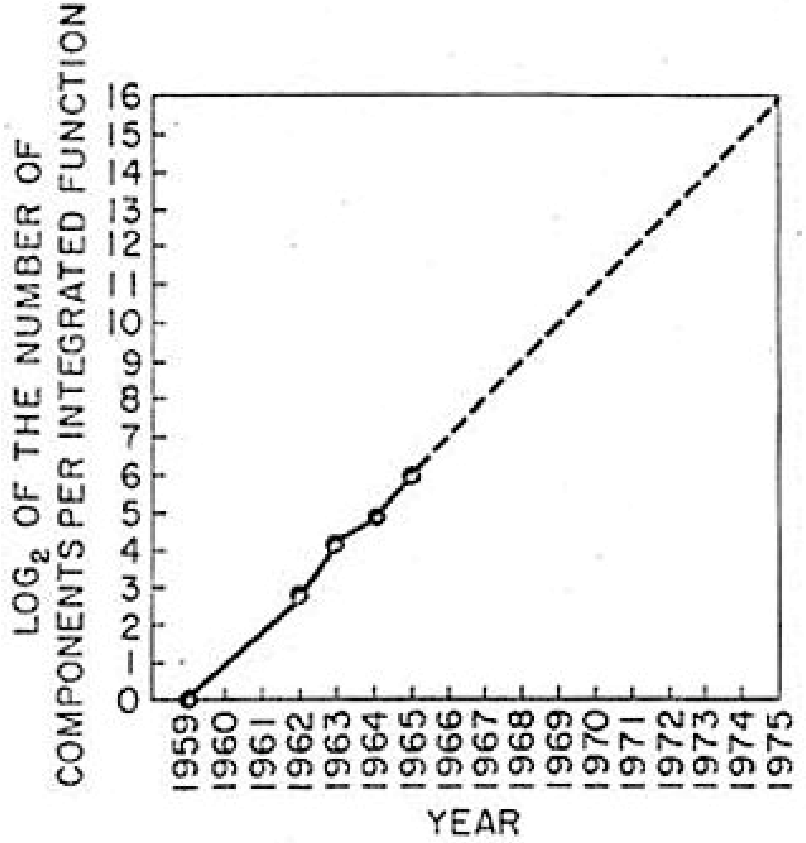
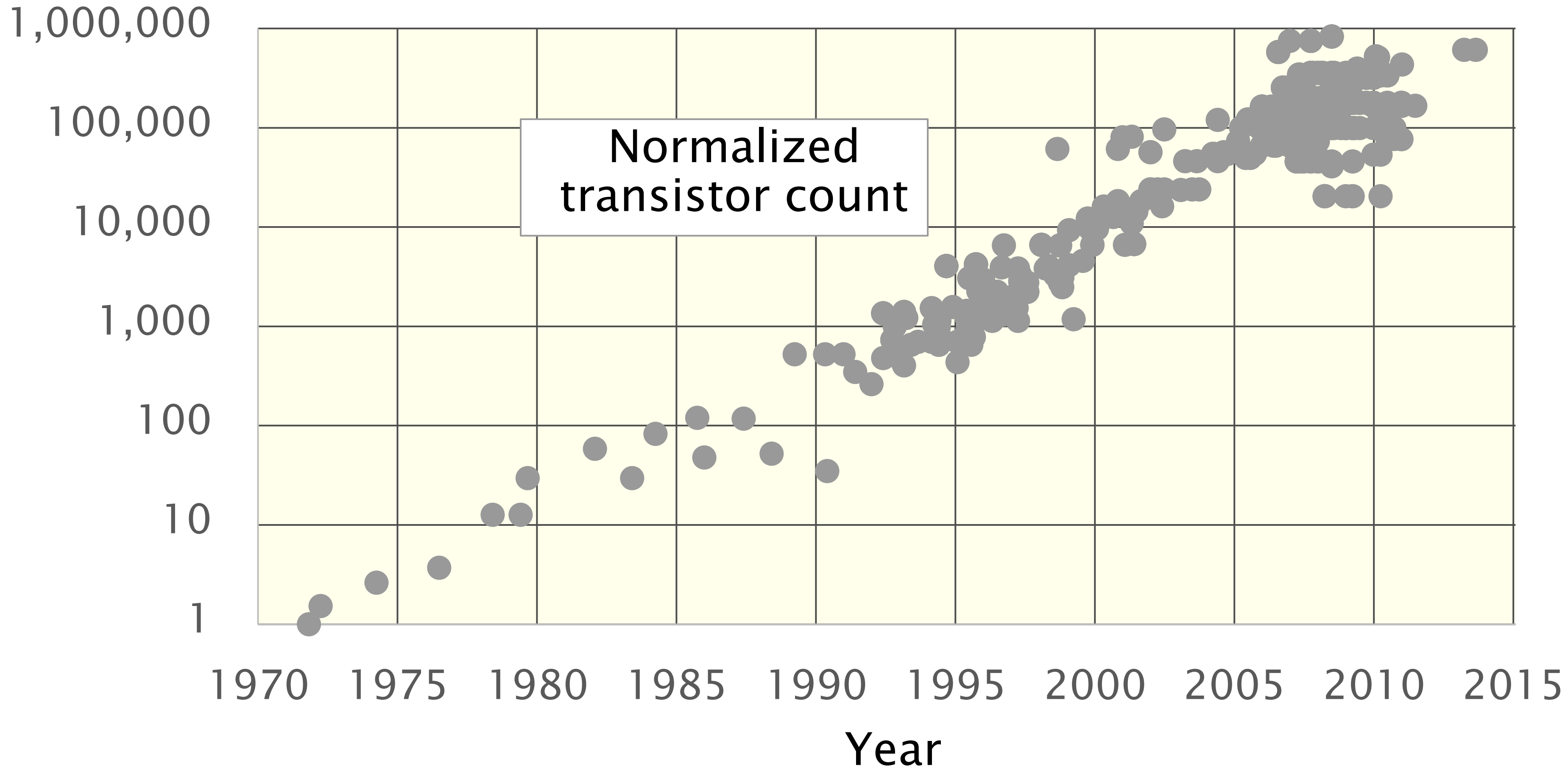


Fig. 2 Number of components per Integrated function for minimum cost per component extrapolated vs time.

(“The number of transistors has been doubling every year.”)

The “Real” Moore’s Law [M75]

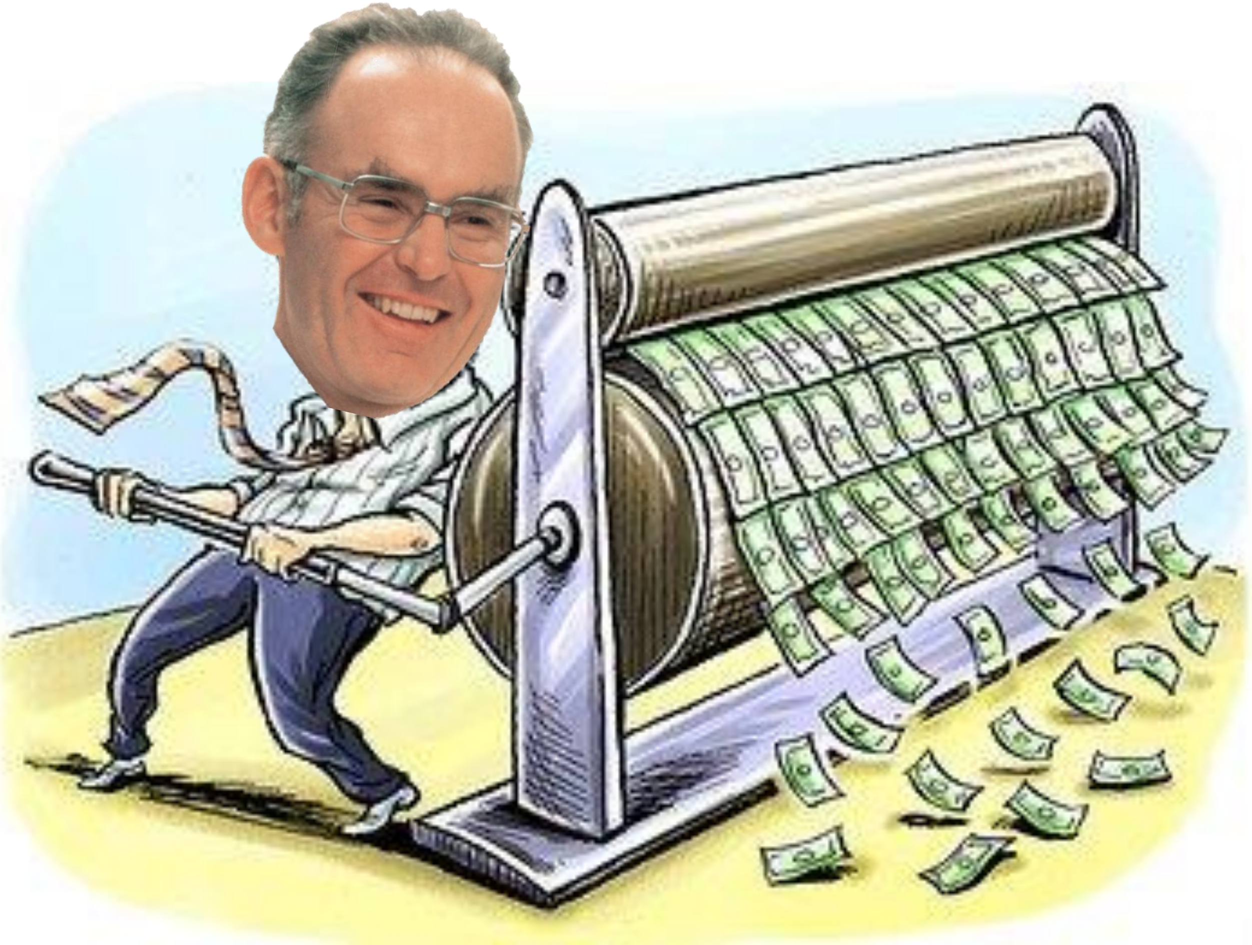
“The new slope might approximate a doubling every two years, rather than every year, by the end of the decade.”



Processor data from Stanford’s CPU DB [DKM12].

50-Year Impact of Moore's Law

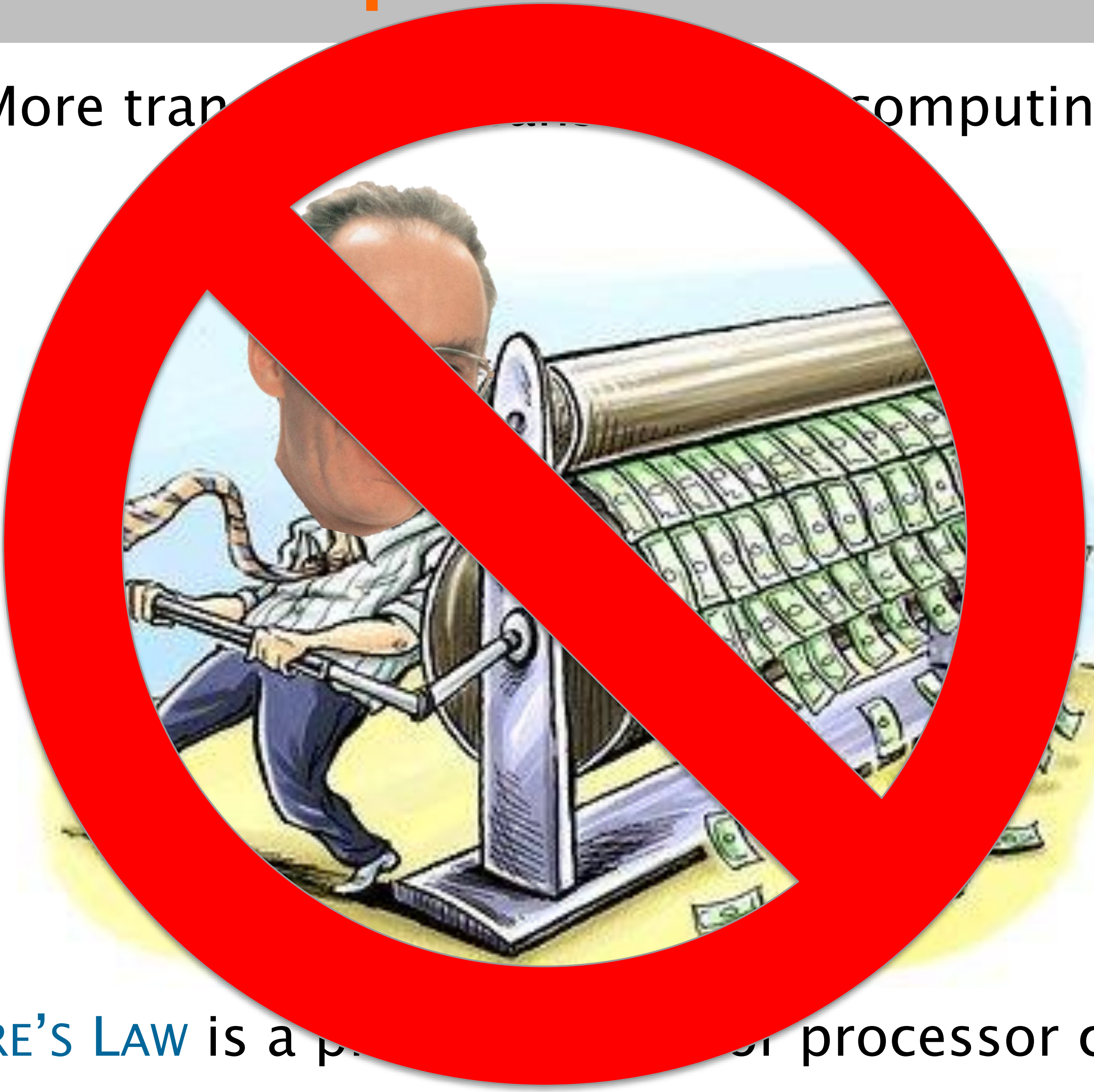
More transistors means cheaper computing.



MOORE'S LAW is a printing press for processor cycles.

50-Year Impact of Moore's Law

More tran   computing.



MOORE'S LAW is a p   r processor cycles.

STOP
THE END IS
~~1980~~
~~1988~~
~~1999~~
~~2000~~
NEAR
ROBERTSON '10

Stand-up Comedian Steven Wright

Why Must the Party End?

Because of Physics

- It's implausible that semiconductor technologists can make wires thinner than atoms, which are at most a few angstroms across.
- The silicon lattice constant is 0.543 nanometers = 5.43 angstroms.

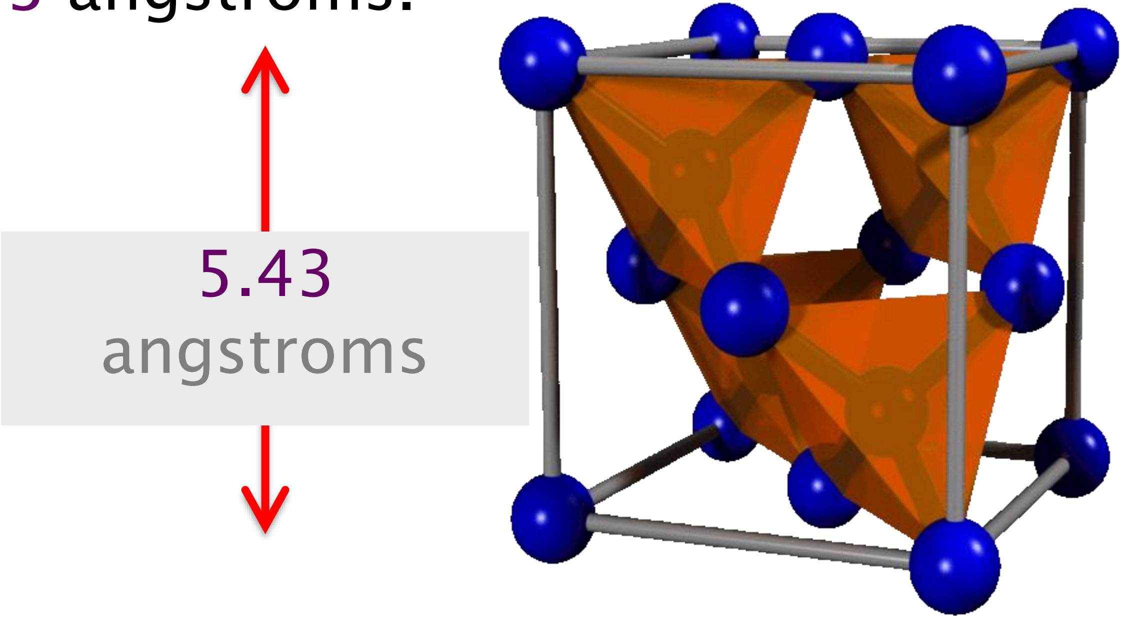
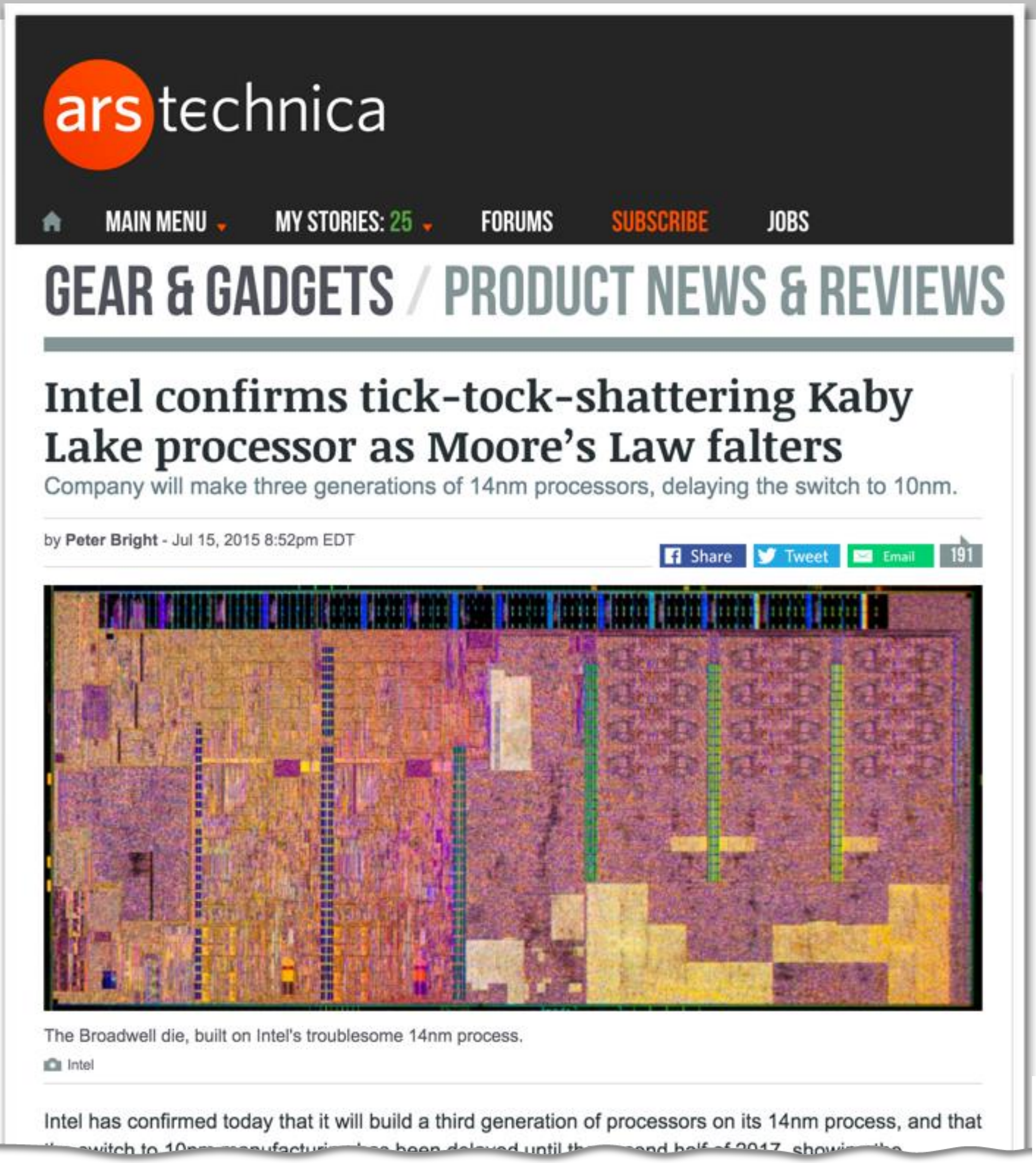


Image by Pieter Kuiper, Wikipedia Commons.

- The IRDS technology roadmap sees miniaturization ending around 5–7nm — at most 3 shrinks from 14 nm — but the historic Moore rate has already attenuated.

Tick-Tock-Tock



GEAR & GADGETS / PRODUCT NEWS & REVIEWS

Intel confirms tick-tock-shattering Kaby Lake processor as Moore's Law falters

Company will make three generations of 14nm processors, delaying the switch to 10nm.

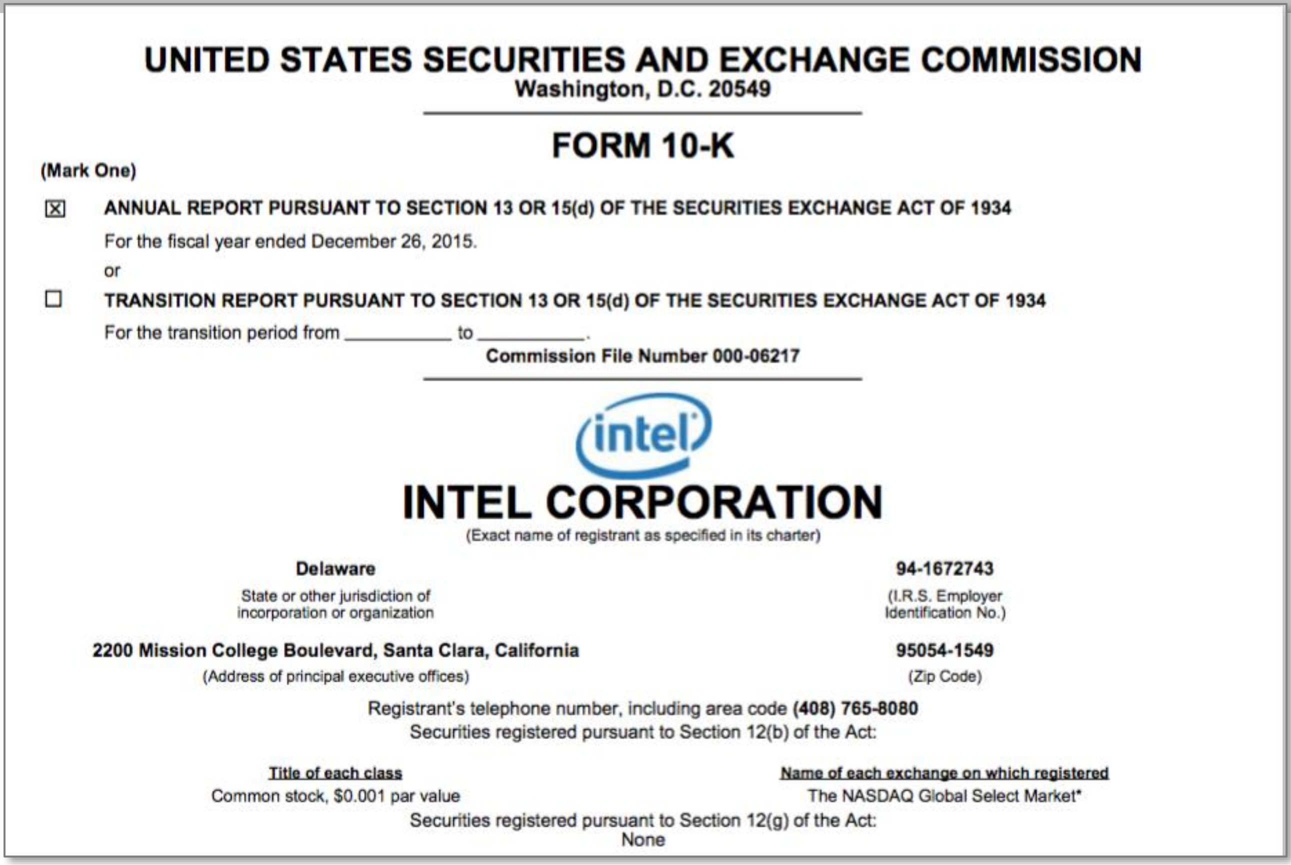
by **Peter Bright** - Jul 15, 2015 8:52pm EDT

The Broadwell die, built on Intel's troublesome 14nm process.

Intel

Intel has confirmed today that it will build a third generation of processors on its 14nm process, and that the switch to 10nm manufacturing has been delayed until the second half of 2017, showing the

February 2016 Intel SEC Filing



UNITED STATES SECURITIES AND EXCHANGE COMMISSION
Washington, D.C. 20549

FORM 10-K

(Mark One)

☒ ANNUAL REPORT PURSUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934

For the fiscal year ended December 26, 2015.

or

□ TRANSITION REPORT PURSUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934

For the transition period from ____________ to ____________.

Commission File Number 000-06217

INTEL CORPORATION

(Exact name of registrant as specified in its charter)

Delaware	94-1672743
State or other jurisdiction of incorporation or organization	(I.R.S. Employer Identification No.)
2200 Mission College Boulevard, Santa Clara, California	95054-1549
(Address of principal executive offices)	(Zip Code)

Registrant's telephone number, including area code (408) 765-8080

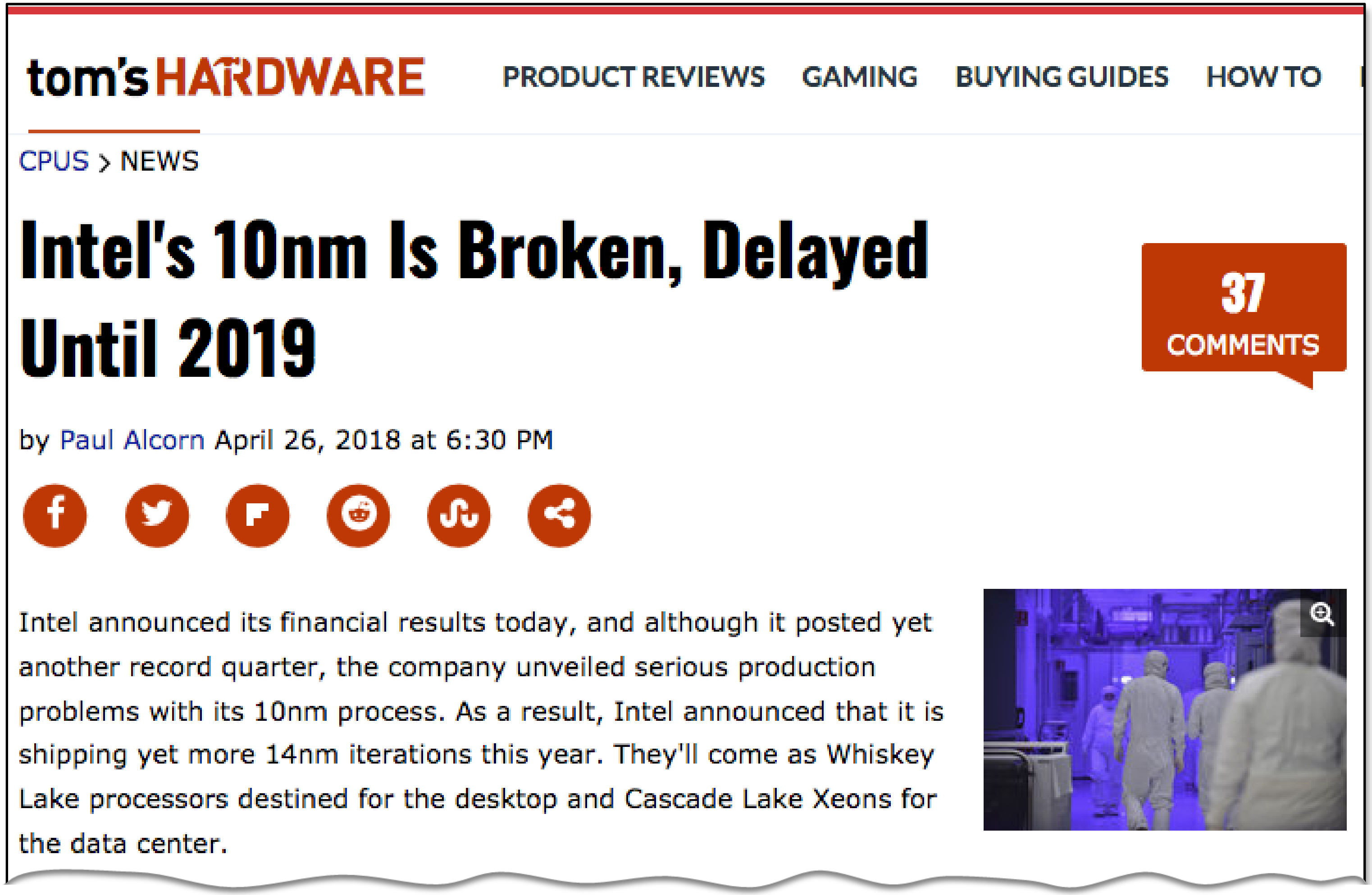
Securities registered pursuant to Section 12(b) of the Act:

Title of each class	Name of each exchange on which registered
Common stock, $0.001 par value	The NASDAQ Global Select Market*

Securities registered pursuant to Section 12(g) of the Act:
None

"We expect to lengthen the amount of time we will utilize our 14nm and our next-generation 10nm process technologies, further optimizing our products and process technologies while meeting the yearly market cadence for product introductions."

Tock–Tock–Tock…



Intel's 10nm Is Broken, Delayed Until 2019

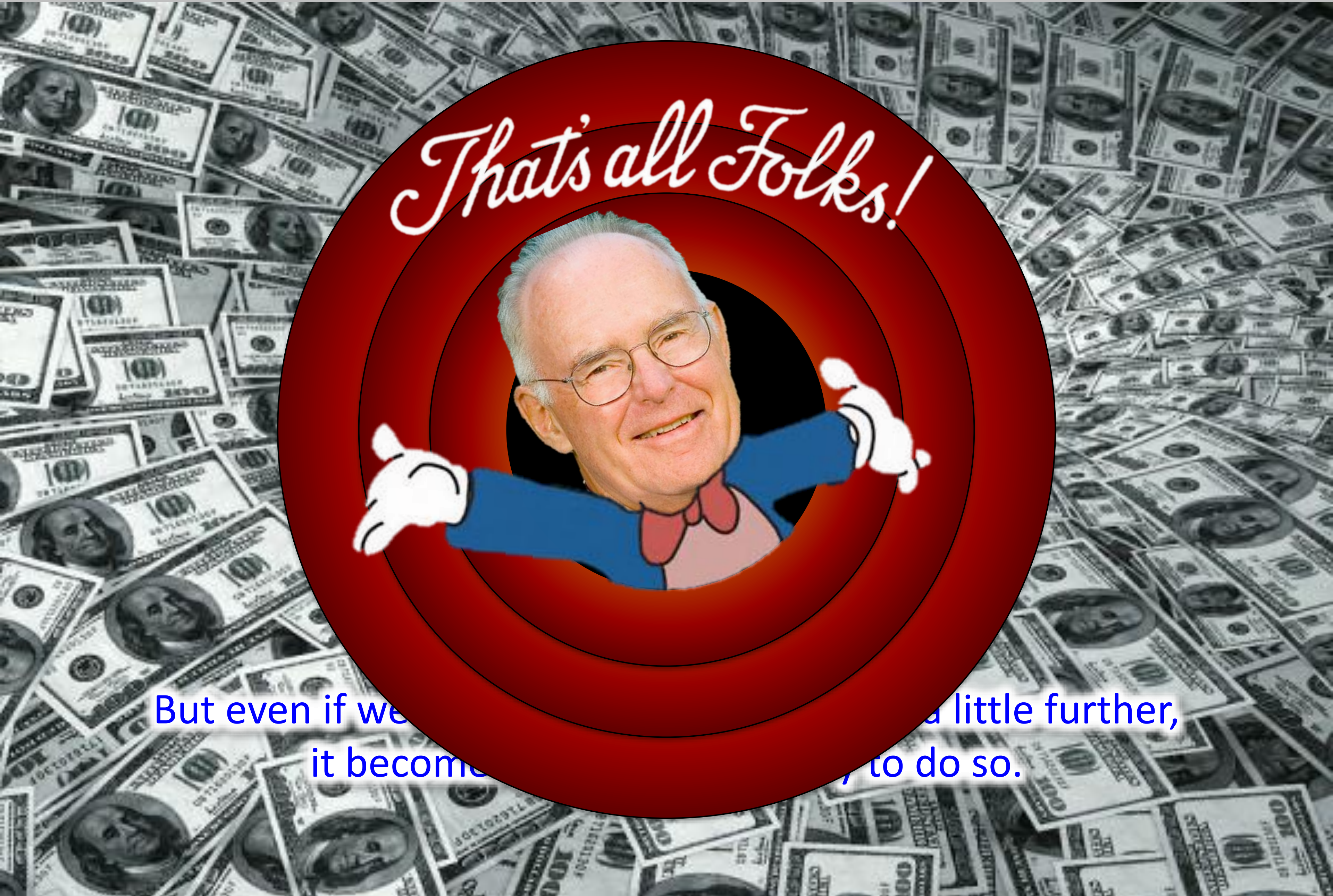
by Paul Alcorn April 26, 2018 at 6:30 PM

Intel announced its financial results today, and although it posted yet another record quarter, the company unveiled serious production problems with its 10nm process. As a result, Intel announced that it is shipping yet more 14nm iterations this year. They'll come as Whiskey Lake processors destined for the desktop and Cascade Lake Xeons for the data center.

Oh, My! Economics!

But even if we could push semiconductors a little further, it becomes prohibitively costly to do so.

Oh, My! Economics!



But even if we a little further,
it become to do so.

Software Performance Engineering to the Rescue!

Redress a Legacy of Excess

Considerable performance can be mined from many existing software applications by ridding them of their

- Performance improvements will no longer be broad-based and come on a predictable schedule, as with MOORE'S LAW.
- Instead, software performance engineering will produce opportunistic, uneven, and sporadic gains.

Algorithmic Tailoring of Matrix Multiplication

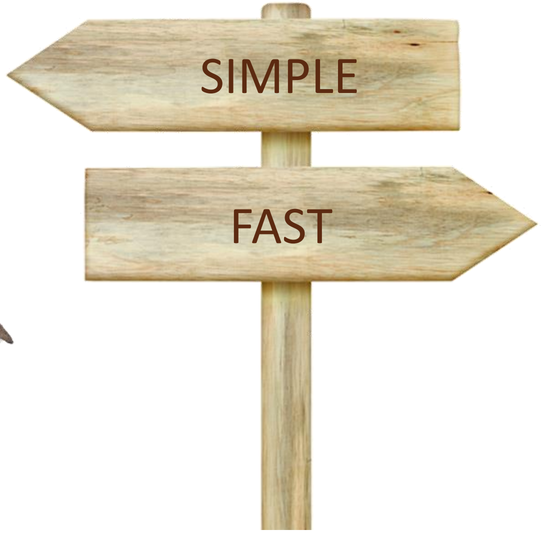
Version	Implementation	Running time (s)	Relative speedup	Absolute Speedup	GFLOPS	Percent of peak
1	Python	21041.67	1.00	1	0.006	0.001
2	Java	2387.32	8.81	9	0.058	0.007
3	C	1155.77	2.07	18	0.118	0.014
4	+ interchange loops	177.68	6.50	118	0.774	0.093
5	+ optimization flags	54.63	3.25	385	2.516	0.301
6	Parallel loops	3.04	17.97	6,921	45.211	5.408
7	+ tiling	1.79	1.70	11,772	76.782	9.184
8	Parallel divide-and-conquer	1.30	1.38	16,197	105.722	12.646
9	+ compiler vectorization	0.70	1.87	30,272	196.341	23.486
10	+ AVX intrinsics	0.39	1.76	53,292	352.408	41.677

Machine: Amazon AWS c4.8xlarge

- Dual-socket Intel Xeon E5-2666 v3 (Haswell)
- 18 cores, 2.9 GHz, 60 GiB DRAM

Simple versus Fast

But the fully optimized matrix-multiplication program contains over 20 times more lines of source code than the original Python program!



Simple code is slow.

Fast code is complicated.

Research Agenda

Let's make a world where it's easy to write fast code!

Strategy: Remedy the *ad hoc* nature of software performance engineering by enabling programmers to write fast code as a principled, scientific process.

Four Cornerstones of Science-Based Performance Engineering

- ✓ Systems you can reason about because performance obeys simple mathematical properties, such as monotonicity and composability.
  - ➢ E.g., Cilk runtime system
- ✓ Theories of performance that work in practice.
  - ➢ E.g., asymptotic analysis, work-span analysis, ideal-cache model, BOTEC
- ✓ Diagnostic tools for correctness and performance whose efficacy is mathematically grounded.
  - ➢ E.g., Cilkscale, Cilkprof, Cilksan
- ✓ Reliable measurement and ubiquitous instrumentation.
  - ➢ E.g., CilkCloud, CSI

Four Cornerstones of Science-Based Performance Engineering

✓Systems you can reason about because performance obeys simple mathematical properties, such as monotonicity and composability.

- ➢E.g., Cilk runtime system

✓Theories of performance that work in practice.

- ➢E.g., asymptotic analysis, work-span analysis, ideal-ca

✓Diagnostic tools for corre s and perf whose efficacy is mathematically grounde

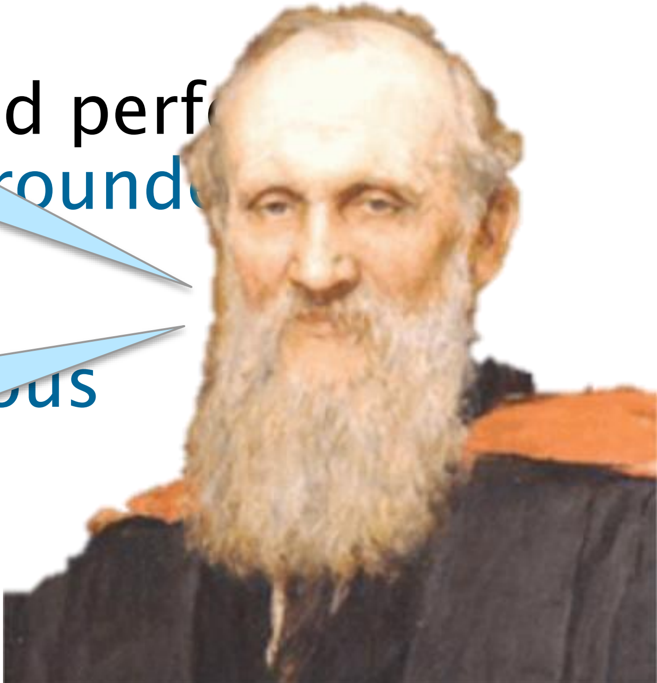
- ➢E.g., Cilkscale, Cilkprof, Cilksan

✓Reliable ous instrume

- ➢E.g., CilkCloud, CSI

To measure is to know.

If you cannot measure it, you cannot improve it.



Lord Kelvin

Life after Moore's Law?

- Cloud computing and the end of MOORE'S LAW will elevate the importance of performance engineering.
- We can make a world in which performance engineering is tedious, unpleasant, and boring, practiced by high priests with arcane knowledge.
- Or we can make a world in which performance engineering is engaging, creative, and fun, practiced even by average programmers.
- Software engineers must become educated in the principles and practices of science-based performance engineering, which too few people today understand or appreciate.

Thank You!