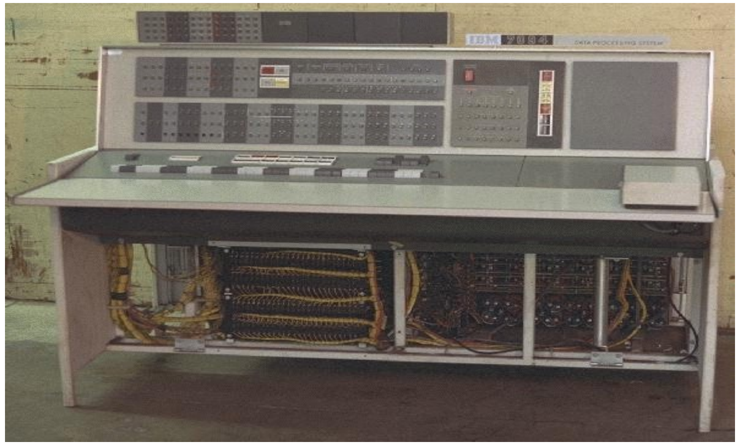
# CM12002
# Computer Systems Architectures

Fabio Nemetz

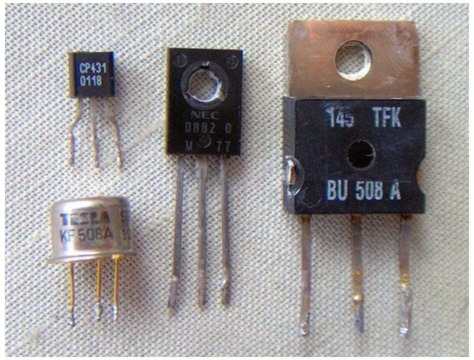
# What is a computer?

## Second Generation – transistors (1950s…)



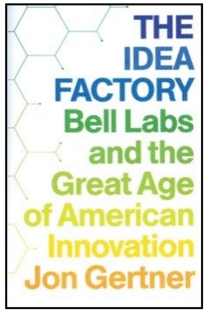
IBM 7094 (1962)

Many consider transistors to be one of the greatest inventions of the 20th century.




Transistors: small devices used to transfer electronic signals across a resistor.

- transistors were smaller than vacuum tubes
- they needed no warm up time
- consumed less energy
- generated much less heat
- faster and more reliable



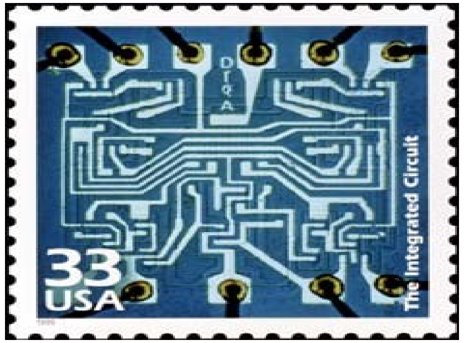

The **Bell Labs** team of John Bardeen, Walter Brattain and William Shockley won the 1956 **Nobel Prize** in Physics for their work in developing transistors.

For info about transistors:
[http://electronics.howstuffworks.com/transistor.htm](http://electronics.howstuffworks.com/transistor.htm)

What is a computer?

**Third generation – Integrated Circuits (chips) (1961)**

- a set of electronic circuits on one small plate ("chip") of semiconductor material, normally silicon
- 10s to 100s of transistors on each chip

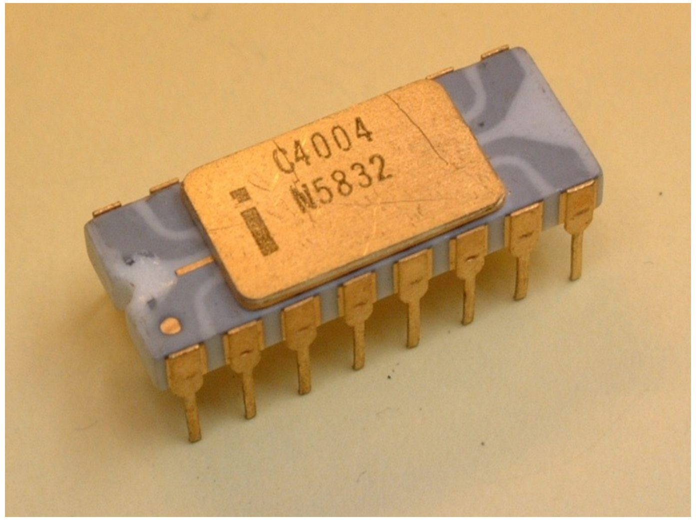



(“Silicon Valley” – started with companies like Fairchild Semiconductor and Intel)

## What is a computer?

**Fourth Generation** - Large Scale Integration (LSI) (70s)
(tens of thousands of transistors per chip)

Intel 4004 (1971) – 1st microprocessor



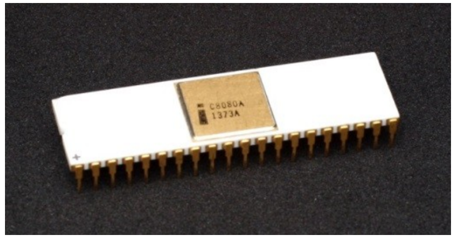

TI TMS 1000 (1971) - 1st microcontroller

# What is a computer?

**January 1975**

January 1975


Intel 8080 (1974)



Altair 8800 – Intel 8080

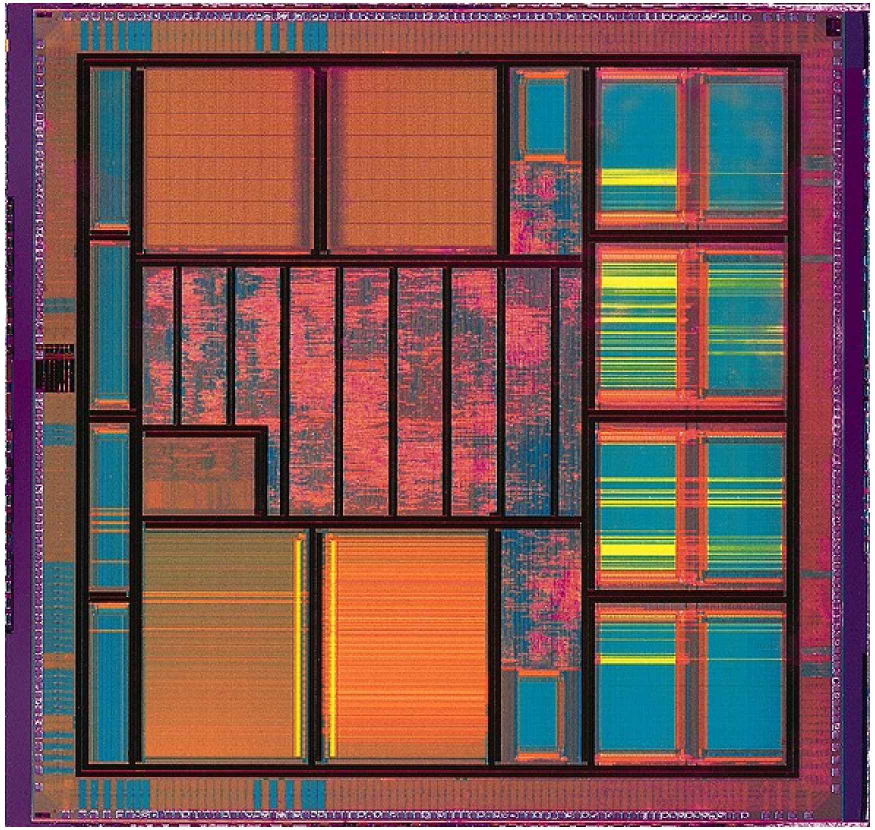
- Bill Gates and Paul Allen: founded **Microsoft** and wrote a Basic language interpreter for the Altair
- Steve Wozniak designed and built a better computer, and showed it to Steve Jobs. They founded **Apple**.

Recommended Film: *Pirates of Silicon Valley (1999)*

## What is a computer?

**Fifth Generation** - Very Large Scale Integration (VLSI) (mid 80s)

(several billion transistors as of 2009 – started to be called Ultra Large Scale Integration (ULSI) – but VLSI is still preferred instead of ULSI)



| Year | Technology used in computers | Relative performance/unit cost |
|---|---|---|
| 1951 | Vacuum tube | 1 |
| 1965 | Transistor | 35 |
| 1975 | Integrated circuit | 900 |
| 1995 | Very large-scale integrated circuit | 2,400,000 |
| 2020 | Ultra large-scale integrated circuit | 500,000,000,000 |

**FIGURE 1.10 Relative performance per unit cost of technologies used in computers over time.** Source: Computer Museum, Boston, with 2020 extrapolated by the authors. See Section 1.13.

From Computer organization and design : the hardware/software interface
David A. Patterson and Hennessy John L. 2021

# Introduction to the x86 Architecture

| Chip | Date | MHz | Trans. | Memory | Notes |
|---|---|---|---|---|---|
| 4004 | 4/1971 | 0.108 | 2300 | 640 | First microprocessor on a chip |
| 8008 | 4/1972 | 0.108 | 3500 | 16 KB | First 8-bit microprocessor |
| 8080 | 4/1974 | 2 | 6000 | 64 KB | First general-purpose CPU on a chip |
| 8086 | 6/1978 | 5–10 | 29,000 | 1 MB | First 16-bit CPU on a chip |
| 8088 | 6/1979 | 5–8 | 29,000 | 1 MB | Used in IBM PC |
| 80286 | 2/1982 | 8–12 | 134,000 | 16 MB | Memory protection present |
| 80386 | 10/1985 | 16–33 | 275,000 | 4 GB | First 32-bit CPU |
| 80486 | 4/1989 | 25–100 | 1.2M | 4 GB | Built-in 8-KB cache memory |
| Pentium | 3/1993 | 60–233 | 3.1M | 4 GB | Two pipelines; later models had MMX |
| Pentium Pro | 3/1995 | 150–200 | 5.5M | 4 GB | Two levels of cache built in |
| Pentium II | 5/1997 | 233–450 | 7.5M | 4 GB | Pentium Pro plus MMX instructions |
| Pentium III | 2/1999 | 650–1400 | 9.5M | 4 GB | SSE Instructions for 3D graphics |
| Pentium 4 | 11/2000 | 1300–3800 | 42M | 4 GB | Hyperthreading; more SSE instructions |
| Core Duo | 1/2006 | 1600–3200 | 152M | 2 GB | Dual cores on a single die |
| Core | 7/2006 | 1200–3200 | 410M | 64 GB | 64-bit quad core architecture |
| Core i7 | 1/2011 | 1100–3300 | 1160M | 24 GB | Integrated graphics processor |


Source: Tanembaum's Structured Computer Organization

Figure 1-11. Key members of the Intel CPU family. Clock speeds are measured in MHz (megahertz), where 1 MHz is 1 million cycles/sec.

# The present

The main hardware developments since the 70s have been in the *scale* of processors: billions of transistors per chip (VLSI &ULSI - very/ultra large scale integration), and consequently engineered on smaller and smaller scale.

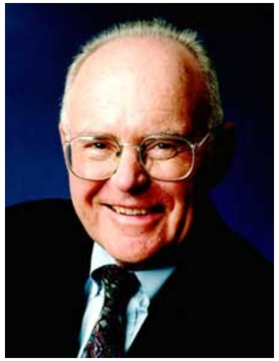
These have required/enabled many software, conceptual and connectivity developments:

- The rise of the operating system.
- Networking, distributed computing.
- Mobile and ubiquitous computing.

# The future

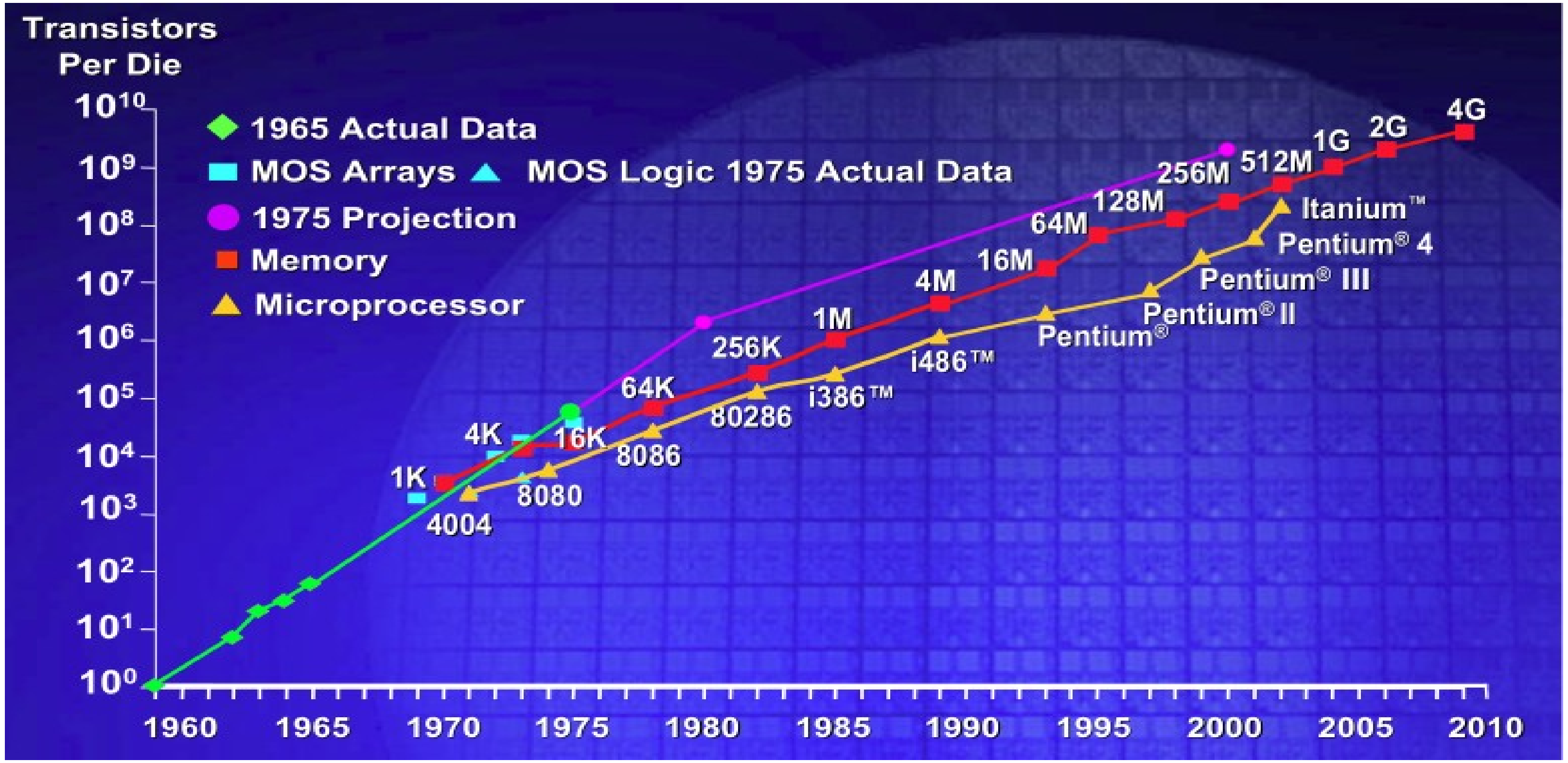
- **Moore's Law (1965)***
  - computing power (relative to component cost) **doubles** over a fixed time (about 2 years).

- This has been sustained for 40 years, but the *physical properties* of silicon mean that it cannot continue indefinitely: it already slowed down.
- Really?



*(Gordon Moore co-founded Intel in 1968)

# Moore’s Law

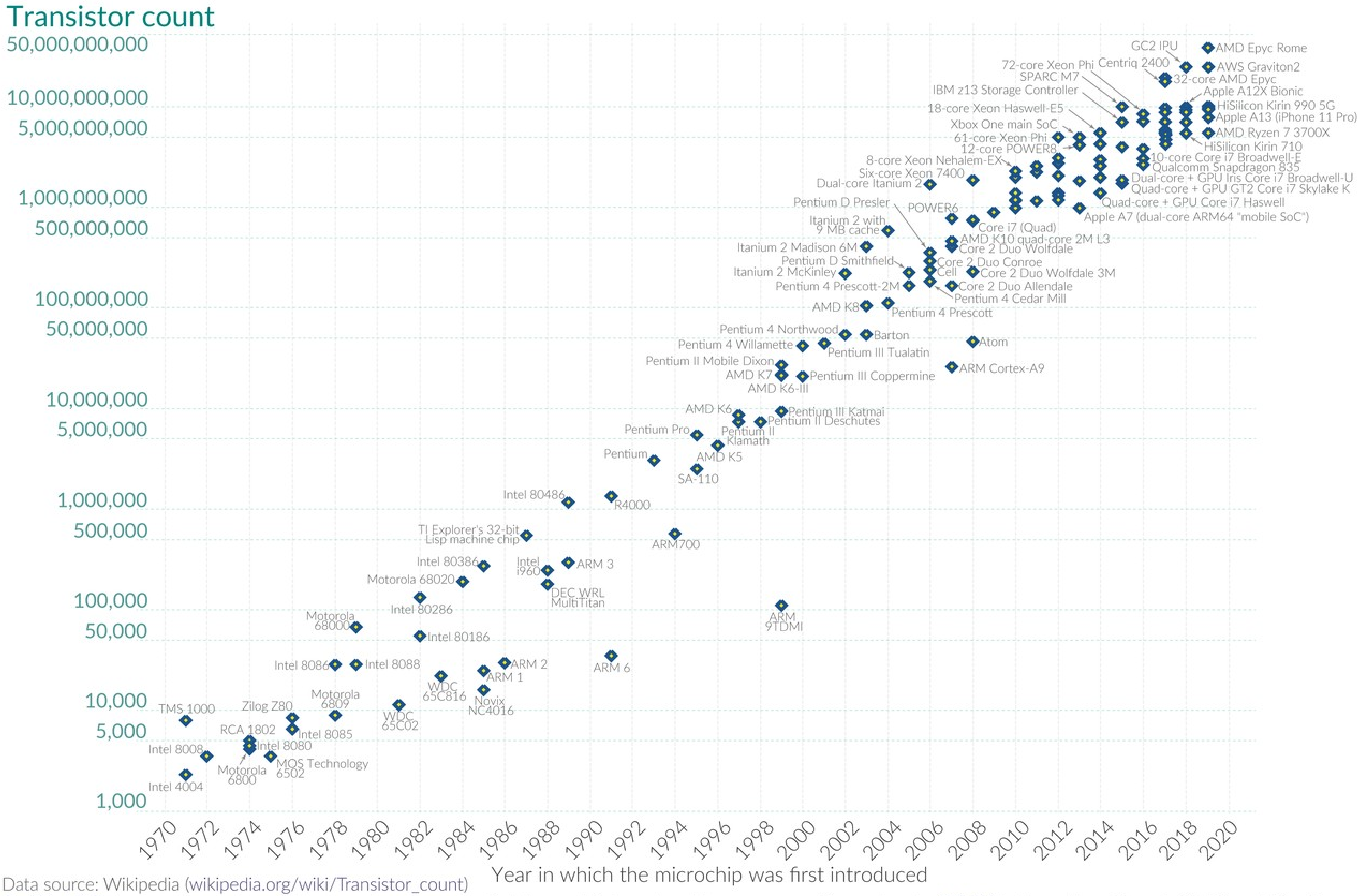


Transistors Per Die
1965 Actual Data
MOS Arrays
MOS Logic 1975 Actual Data
1975 Projection
Memory
Microprocessor
$10^{10}$
$10^{9}$
$10^{8}$
$10^{7}$
$10^{6}$
$10^{5}$
$10^{4}$
$10^{3}$
$10^{2}$
$10^{1}$
$10^{0}$
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
1K
4K
16K
64K
256K
1M
4M
16M
64M
128M
256M
512M
1G
2G
4G
4004
8080
8086
80286
i386™
i486™
Pentium®
Pentium® II
Pentium® III
Pentium® 4
Itanium™

# Moore's Law: The number of transistors on microchips doubles every two years

Our World in Data

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.




Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)

OurWorldinData.org – Research and data to make progress against the world's largest problems.

# The future

- The 2010 update to the International Technology Roadmap for Semiconductors predicted that after 2013 transistor counts and densities were to double only every three years (approx. tenfold increase in the next 10 years)

- Really?


Huawei's Kirin 980 (7 nanometre):

"The advent of 10nm and 7nm chips means you are unlikely to see anyone suggesting again that we are reaching our limits. Physicists always come up with new physics.

Moore's law states that each year you would have the size reducing by half but resistor count and performance improvements, etc, increasing by double, so we are actually beating that, as we are going from 10nm to 7nm in the space of a couple months." (Akhram Mohamed, CTO of Huawei, 2018)

1nm = 1 billionth of a meter

Same year (later on)

17 Sep 2018 | 20:09 GMT

# David Patterson Says It’s Time for New Computer Architectures and Software Languages

Moore’s Law is over, ushering in a golden age for computer architecture, says RISC pioneer

By **Tekla S. Perry**

“*there’s no better time than now to be a computer architect*.”

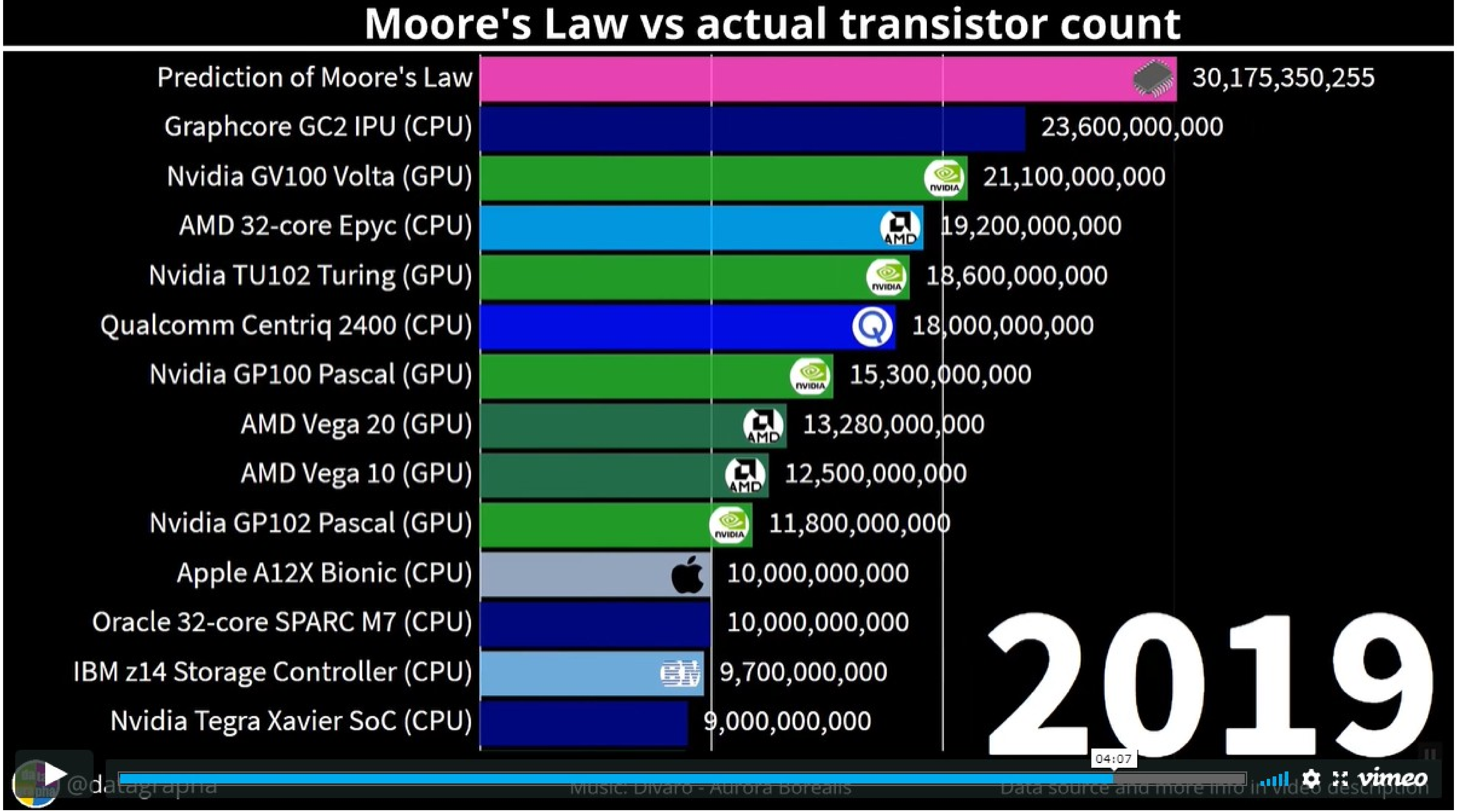
That’s because Moore’s Law really is over, he says: “*We are now a factor of 15 behind where we should be if Moore’s Law were still operative. We are in the post–Moore’s Law era*.”

This means, Patterson told engineers attending the 2018 @Scale Conference held in San Jose last week, that “*we’re at the end of the performance scaling that we are used to. When performance doubled every 18 months, people would throw out their desktop computers that were working fine because a friend’s new computer was so much faster*.”

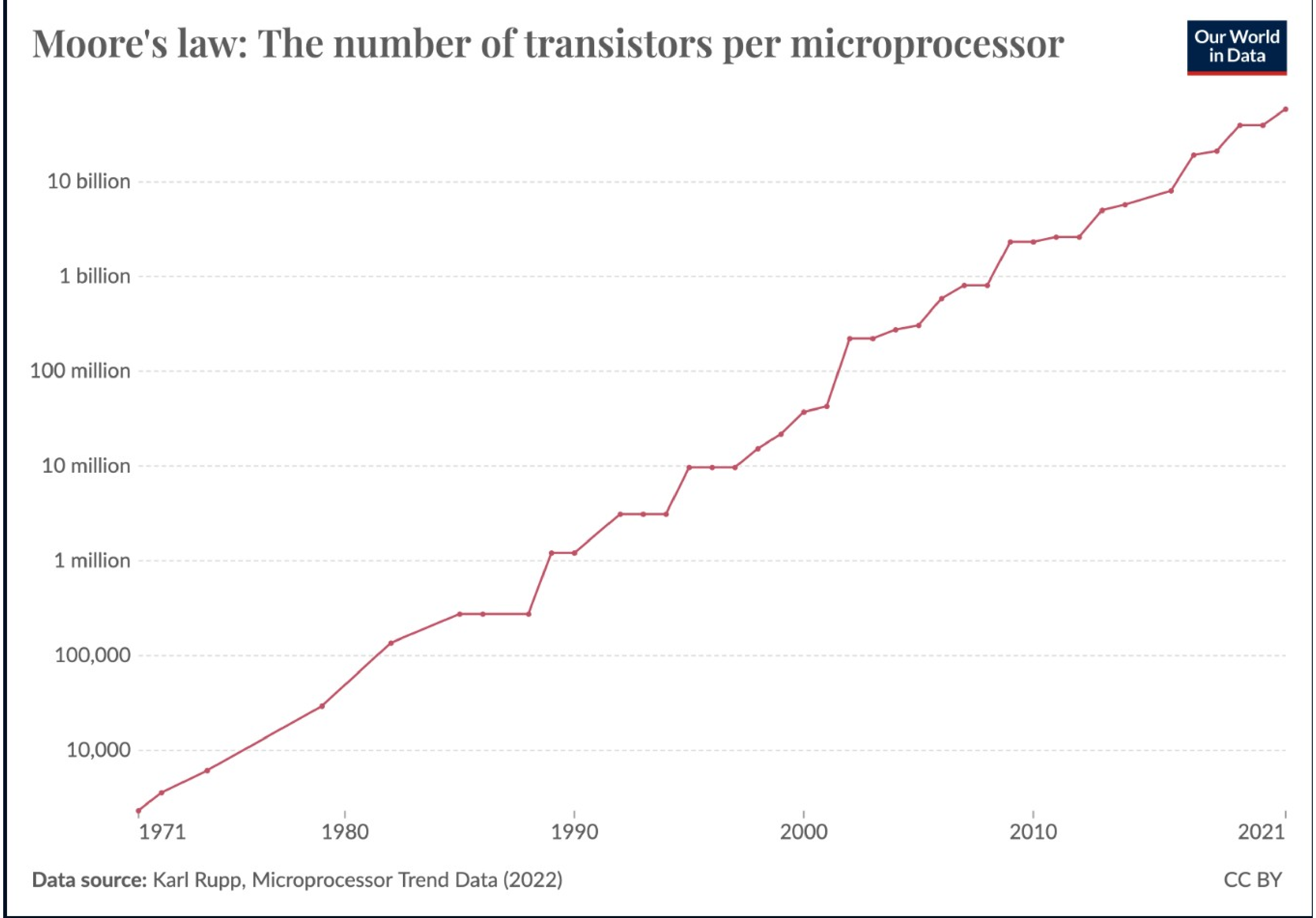
https://goo.gl/4n699j


"Calling the end of Moore's Law is a bit of a tradition. As Peter Lee, vice president at Microsoft Research, quipped to The Economist a few years ago, "*The number of people predicting the death of Moore's Law doubles every two years*." To date, prophets of doom have been premature, and though the pace is slowing, the industry continues to dodge death with creative engineering.

Source:
https://singularityhub.com/2020/08/23/moores-law-lives-intel-says-chips-will-pack-50-times-more-transistors/


Moore's Law vs actual transistor count
Prediction of Moore's Law
30,175,350,255
Graphcore GC2 IPU (CPU)
23,600,000,000
Nvidia GV100 Volta (GPU)
21,100,000,000
AMD 32-core Epyc (CPU)
19,200,000,000
Nvidia TU102 Turing (GPU)
18,600,000,000
Qualcomm Centriq 2400 (CPU)
18,000,000,000
Nvidia GP100 Pascal (GPU)
15,300,000,000
AMD Vega 20 (GPU)
13,280,000,000
AMD Vega 10 (GPU)
12,500,000,000
Nvidia GP102 Pascal (GPU)
11,800,000,000
Apple A12X Bionic (CPU)
10,000,000,000
Oracle 32-core SPARC M7 (CPU)
10,000,000,000
IBM z14 Storage Controller (CPU)
9,700,000,000
Nvidia Tegra Xavier SoC (CPU)
9,000,000,000
2019
04:07
vimeo

Animation:


Moore's law: The number of transistors per microprocessor
Our World in Data
10 billion
1 billion
100 million
10 million
1 million
100,000
10,000
1971
1980
1990
2000
2010
2021
Data source: Karl Rupp, Microprocessor Trend Data (2022)
CC BY

# What is a computer?

What is a computer? Two kinds of answers

- A computer is a “black box” for processing data
  - all I need to know is: what outputs I will get from a given input (functional)

- A computer is a device with internal structure :
  - different parts which interact to turn input into output. To understand what a computer is, I need to know what these parts are, and how they work.

# What is a computer?

Why study computer architecture if you don’t want to be a hardware designer ?

- Learning to program: having a  mental model of the implementation can be useful when solving computational problems
- Efficiency and performance: get the most out of your machine.
- Avoiding errors and failures: understand what could go wrong.
- Recovering from errors: identify the problem, and know how to fix it.

# What is a computer?

## A more subtle approach

•Computers are *complex and varied, but they are designed according to some common, basic principles.*

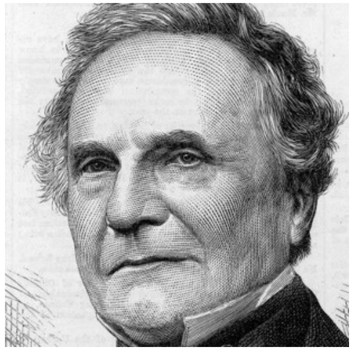
•To describe how a computer works, we need to find the right *level of abstraction.* We can distinguish several of these, corresponding to different answers to the question: What is a computer?

1. A black box.
2. A collection of parts (memory, control, data processing, etc.) each of which is a black box. (ISA: Instruction Set Architecture)
3. A collection of parts, each composed of electronic circuits (storage cells, circuits for control and arithmetic, etc.) each of which is a black box. (Microarchitecture)
4. A collection of parts, each composed of electronic circuits, which are made of electronic components (logic gates, wires, etc.) (Electronics)

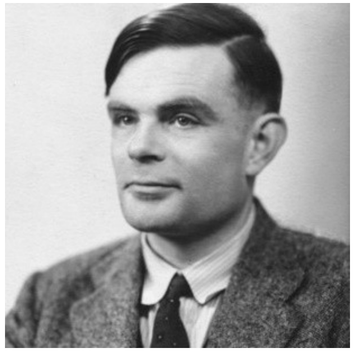
In this course we shall study 2-4 (in that order).

Some things you can do (if you’re curious enough)

- Learn more about:


Charles Babbage

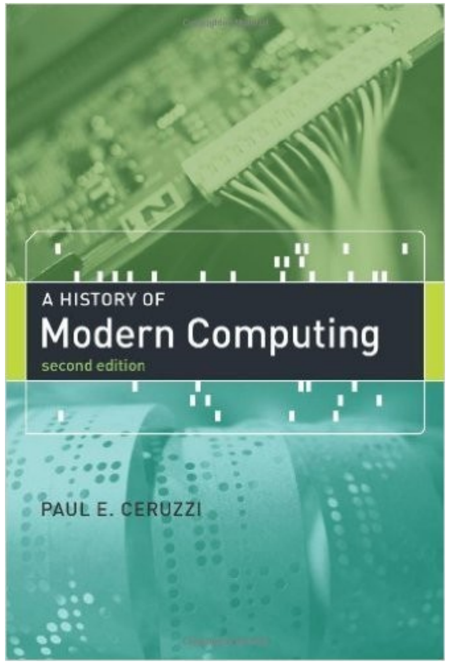

Ada Lovelace


Alan Turing

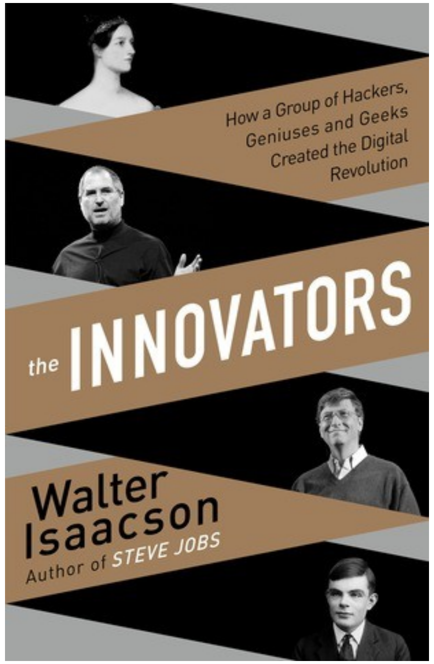

John von Neumann

Previous winners of Turing awards and von Neuman medals

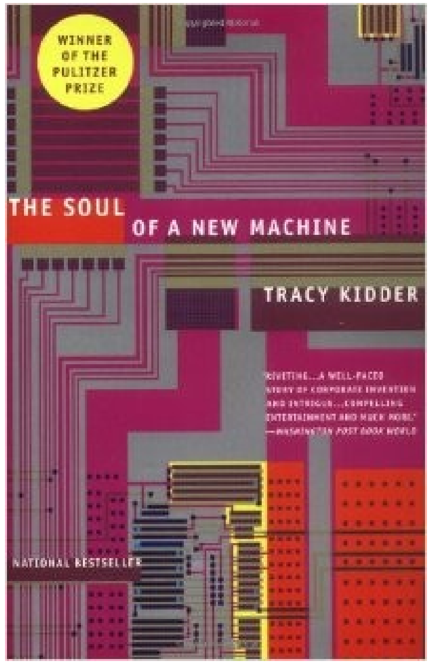

# Interested in the fascinating history of Computers?



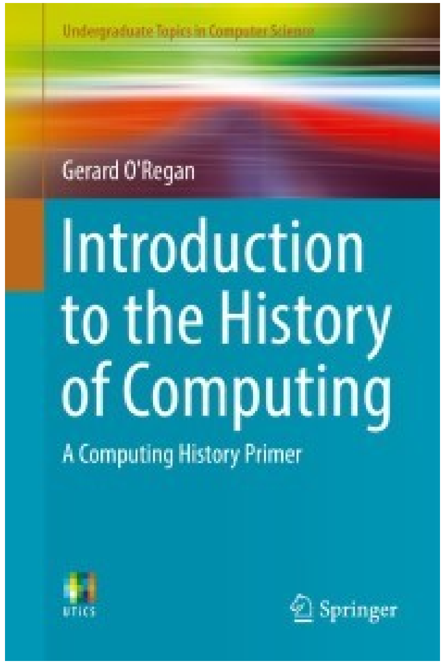

A History of **Modern** Computing (History of Computing) – 2nd ed.
Paul Ceruzzi (Cruzi)



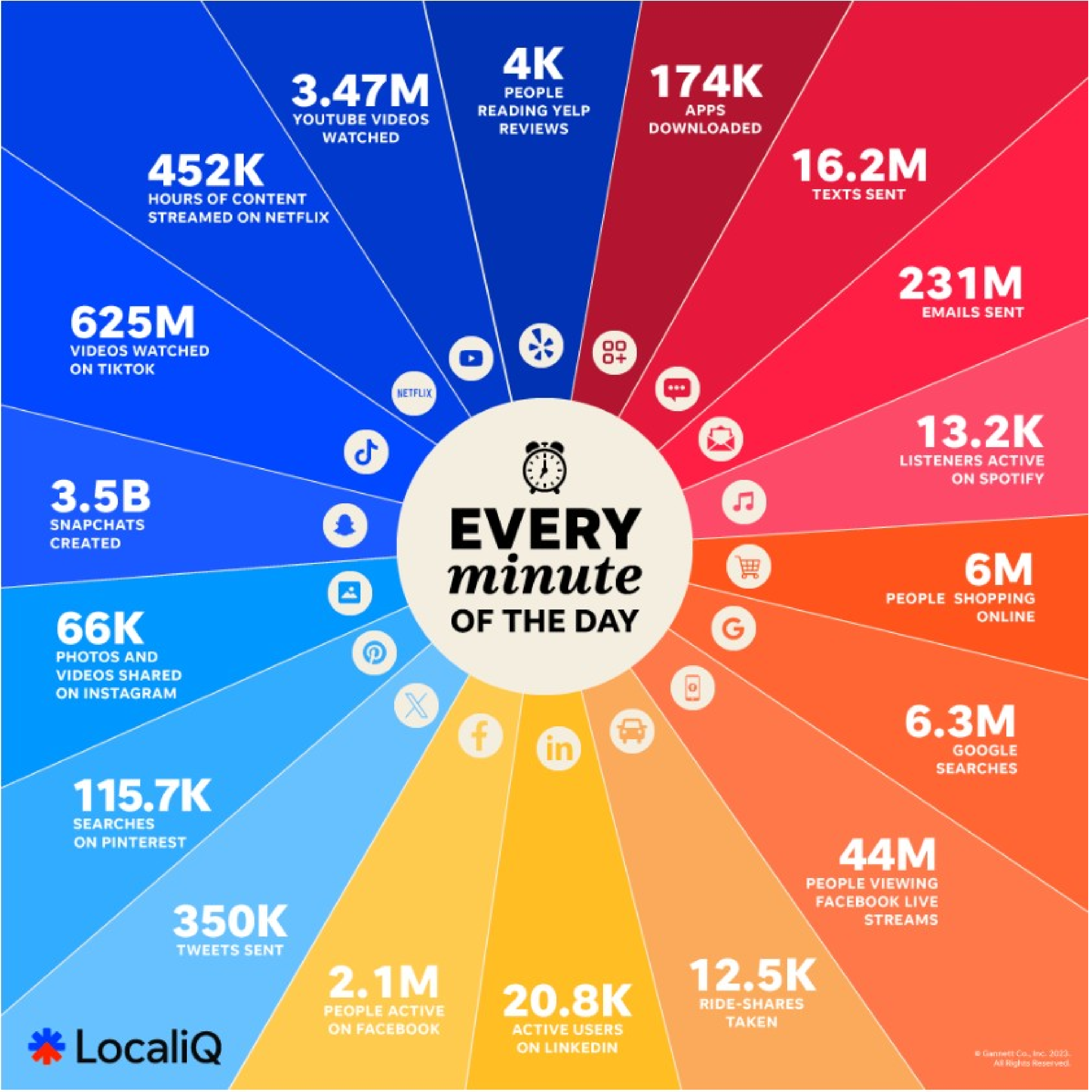

The Innovators
Walter Isaacson




The Soul of A New Machine
Tracy Kidder




Introduction to the History of Computing
Gerard O'Regan
Library link - online access




# Here's what happens every minute on the internet in 2023

Source: https://localiq.com/blog/what-happens-in-an-internet-minute/

# More interesting stuff coming next