Advanced Computer Architecture

Discussion topic: fundamental limits

Feynmann: plenty of room at the bottom

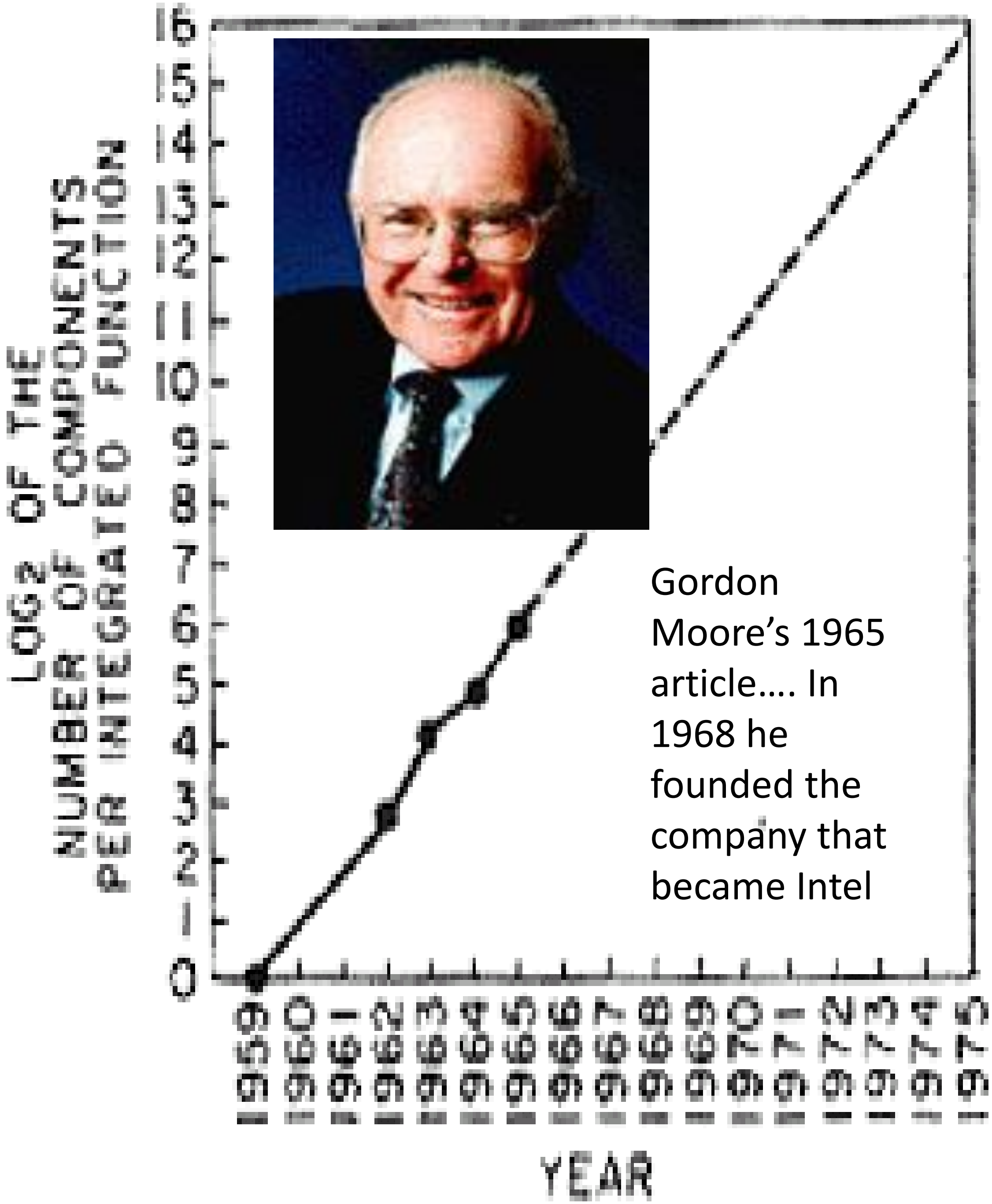
..... **Miniaturizing the computer**

I don't know how to do this on a small scale in a practical way, but I do know that computing machines are very large; they fill rooms. Why can't we make them very small, make them of little wires, little elements—and by little, I mean little. For instance, the wires should be 10 or 100 atoms in diameter, and the circuits should be a few thousand angstroms across.

.....

But I would like to discuss, just for amusement, that there are other possibilities. Why can't we manufacture these small computers somewhat like we manufacture the big ones? Why can't we drill holes, cut things, solder things, stamp things out, mold different shapes all at an infinitesimal level? What are the limitations as to how small a thing has to be before you can no longer mold it?

December 1959



Gordon Moore's 1965 article…. In 1968 he founded the company that became Intel

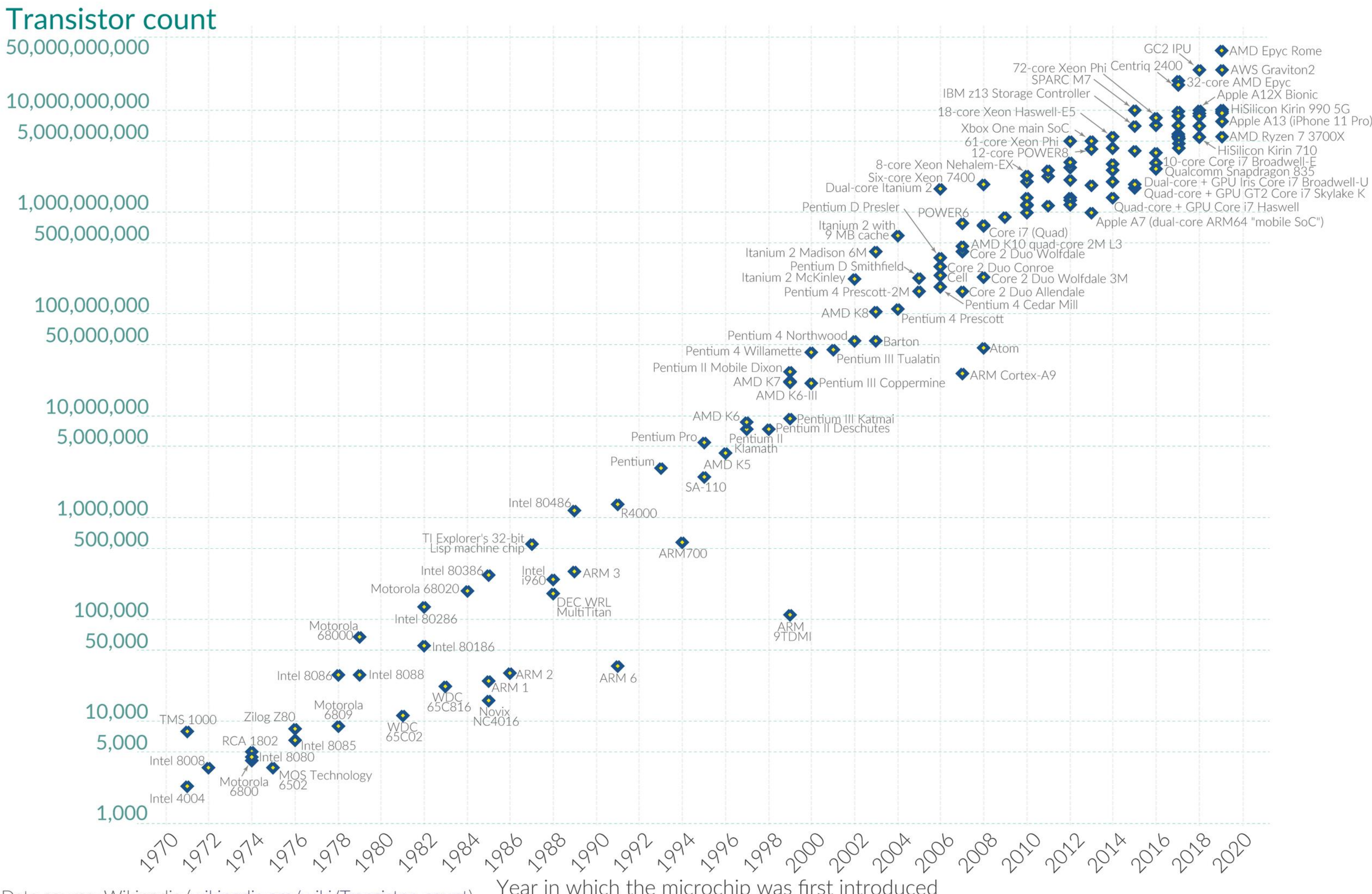
"With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip" Gordon Moore, 1965

Cramming more components onto integrated circuits, Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.

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

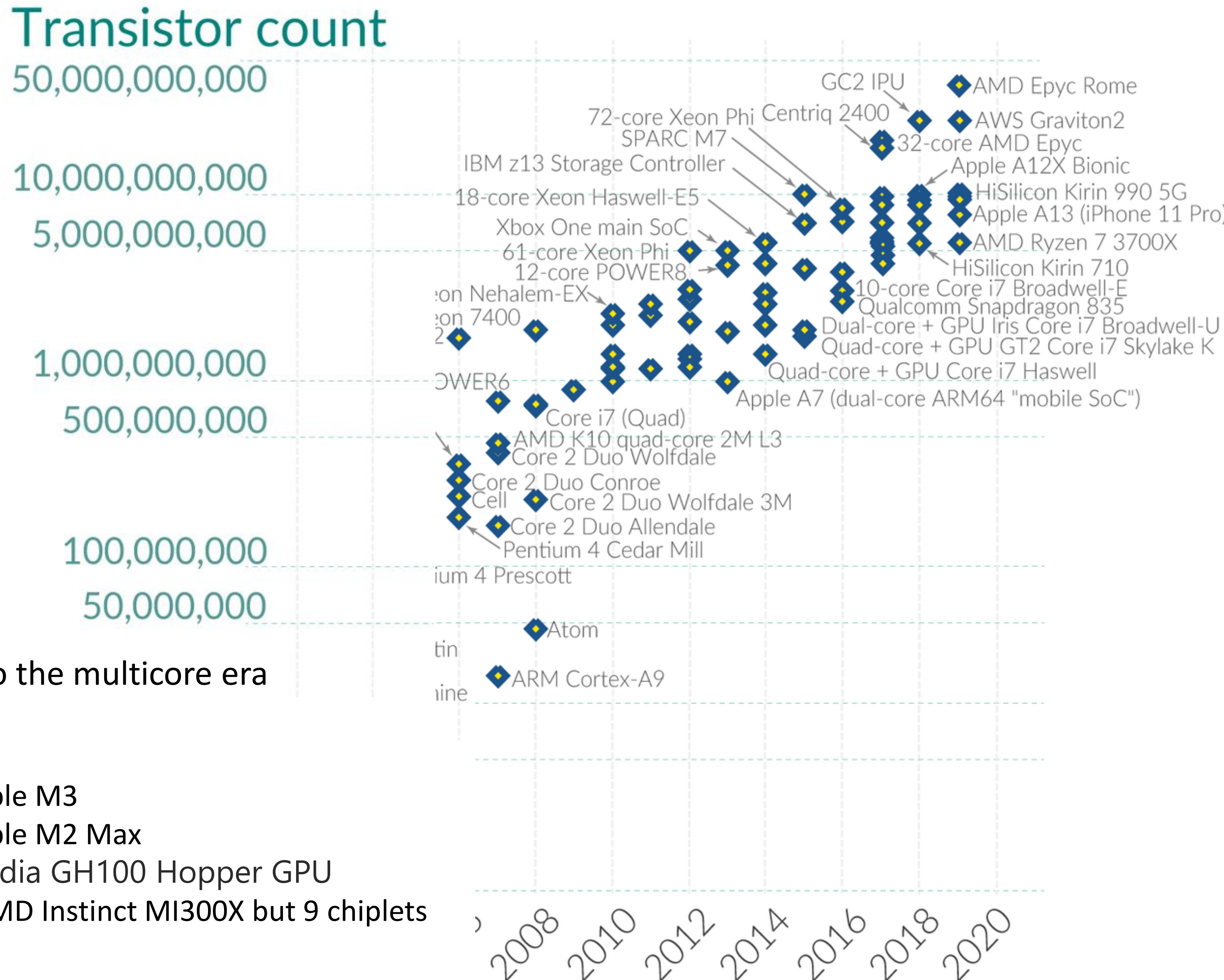
Our World in Data

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.

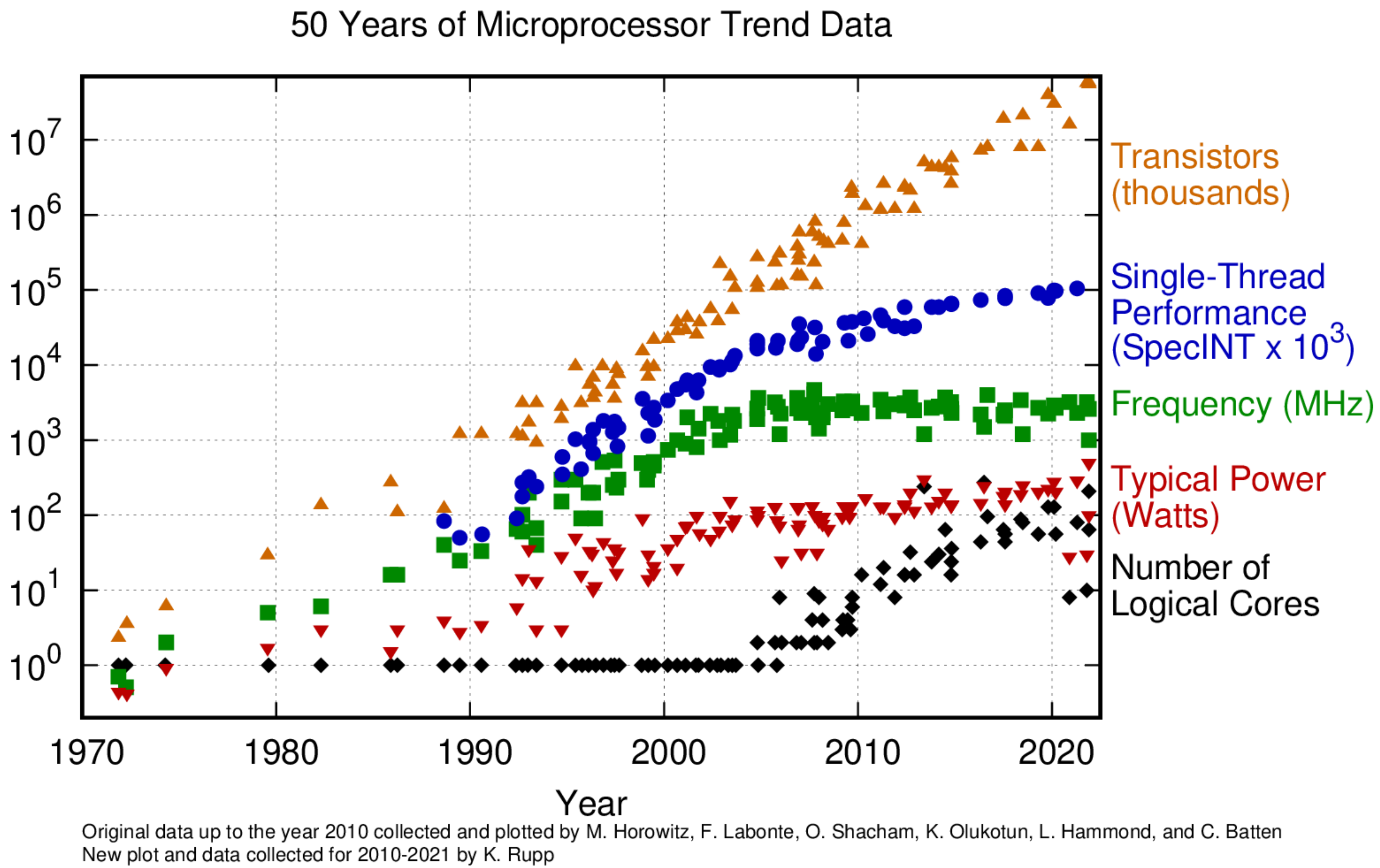


Zooming into the multicore era

2023 chips:

- 25B: Apple M3
- 67B: Apple M2 Max
- 80B: Nvidia GH100 Hopper GPU
- 146B: AMD Instinct MI300X but 9 chiplets

See also https://handwiki.org/wiki/Engineering:Transistor_count

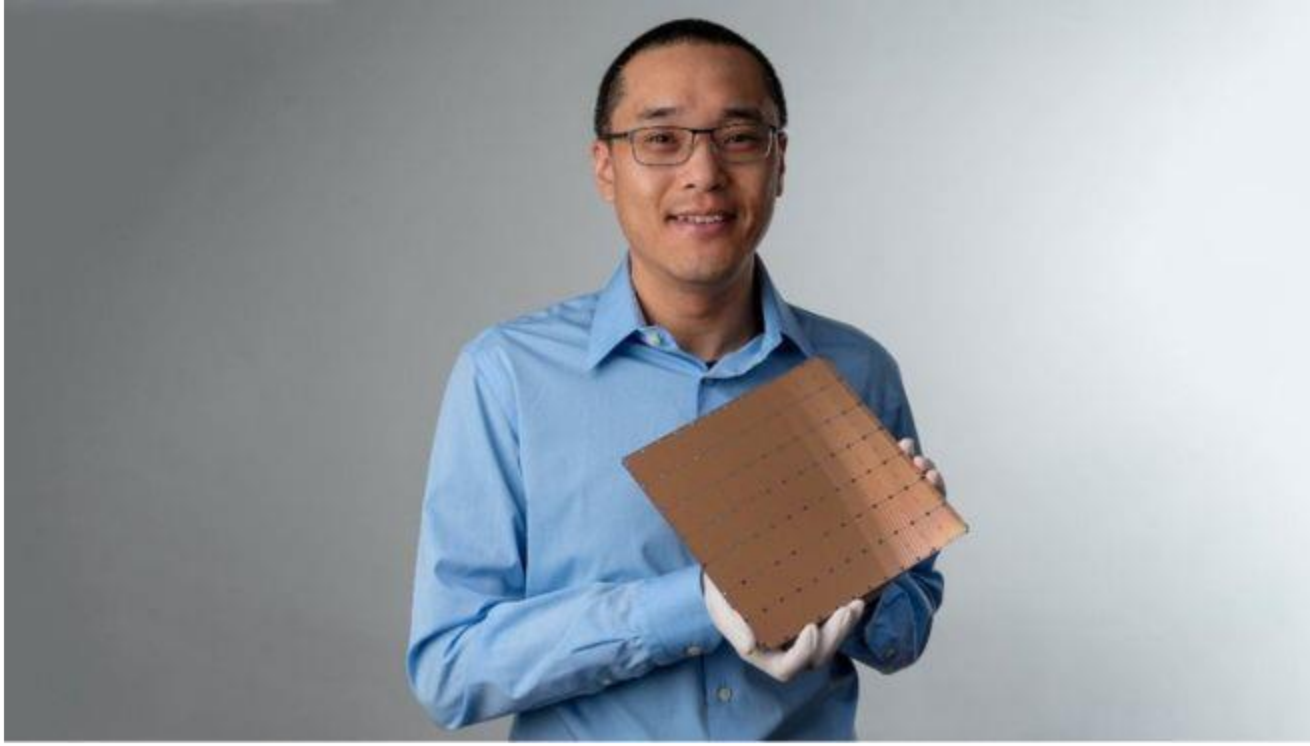


Transistor count continues to grow, but clock frequency increases stopped ca.2005

Single-core performance grows very slowly, but cores/chip has been increasing rapidly since 2005

Power per chip hit >100W around 2005 and has increased only slightly since then

[GitHub - karlrupp/microprocessor-trend-data: Data repository for my blog series on microprocessor trend data.]()



Cerebras CS2

https://www.bbc.co.uk/news/technology-49395577

- 2.6 trillion transistors
- 46,225 mm^2 of silicon
- 850,000 AI programmable cores
- 40 GB of on-die memory (SRAM)
- 20 Petabytes/s memory bandwidth
- 220 Petabits/s fabric bandwidth
- 1.2Tb/s I/O via 12 100Gb ethernet links

Limits

- Historically, computer architecture has been limited severely by what we could build

- We have been on a technology “escalator” driven by increasing VLSI fabrication capability

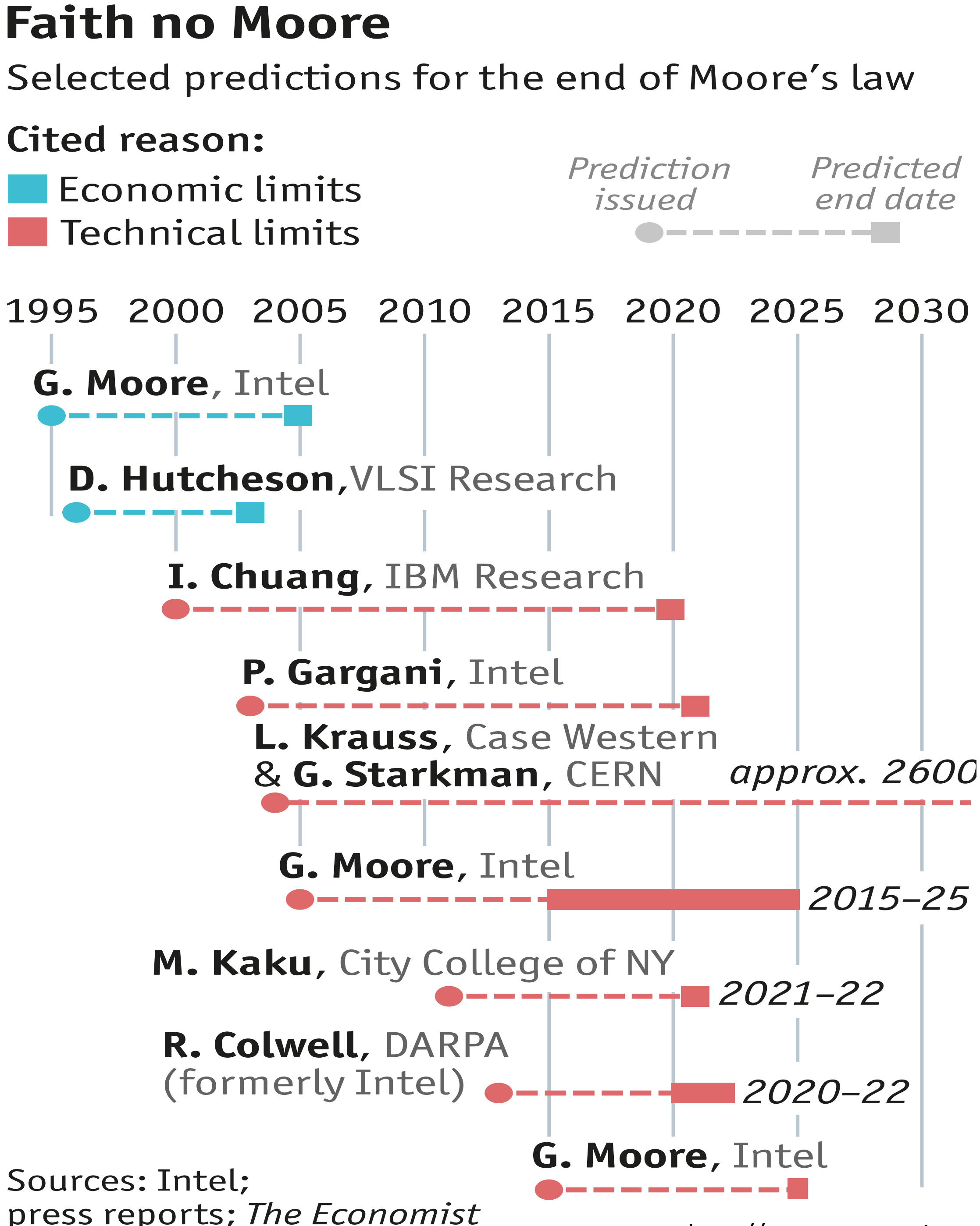
- All exponential growth processes come to end

- At the limit of Moore’s Law, computer architecture will be limited by the fundamental physics of computation
  - Current high-end fabs operate at 3nm (ca.200M transistors/mm^2
  - 1nm is “in development”
  - Key idea is to build in 3D – FinFET, GAA, VTFET

  - Wire length, signal propagation delays
  - Wiring density

  - Power dissipation

- What will change for computer architects as we are more and more limited by fundamental physics?



http://www.economist.com/technology-quarterly/2016-03-12/after-moores-law

How much more room at the bottom is there?

How will our priorities change when year-on-year technology improvements in fabrication capability stop?