Future of Computing: Moore's Law & Its Implications

15-213: Introduction to Computer Systems
27th Lecture, Dec. 3, 2015

Instructors:

Randal E. Bryant and David R. O'Hallaron

Moore's Law Origins

April 19, 1965

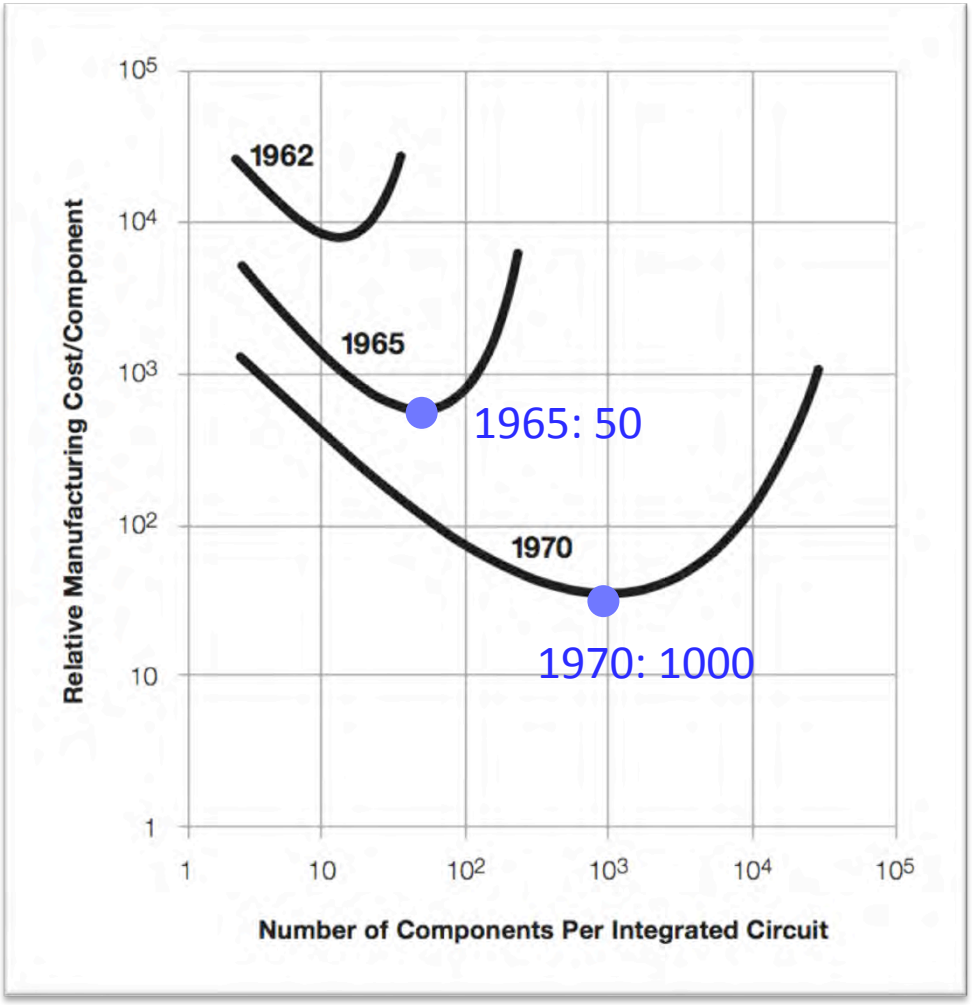
Cramming more components onto integrated circuits

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

By Gordon E. Moore

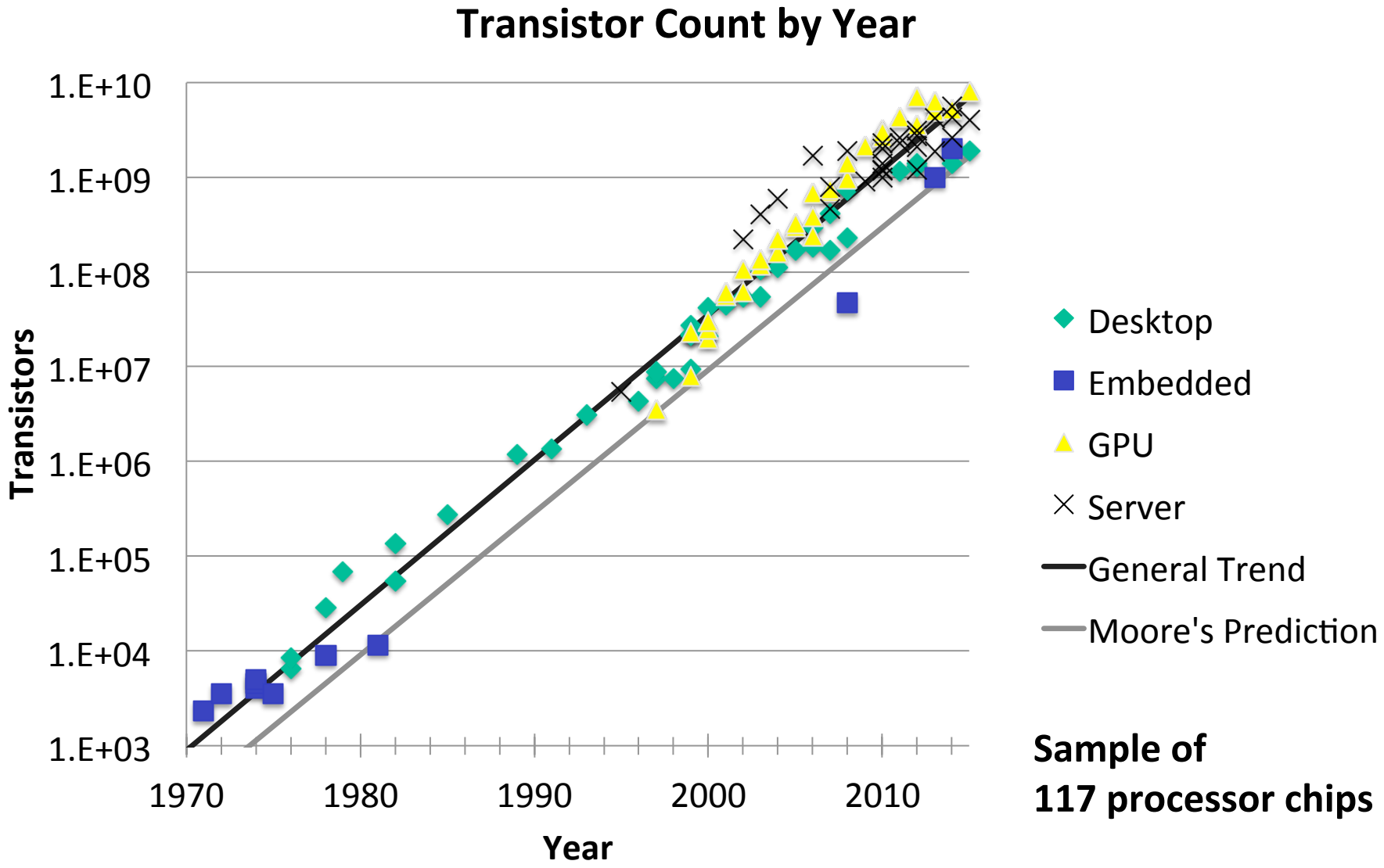
Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

Moore's Law Origins

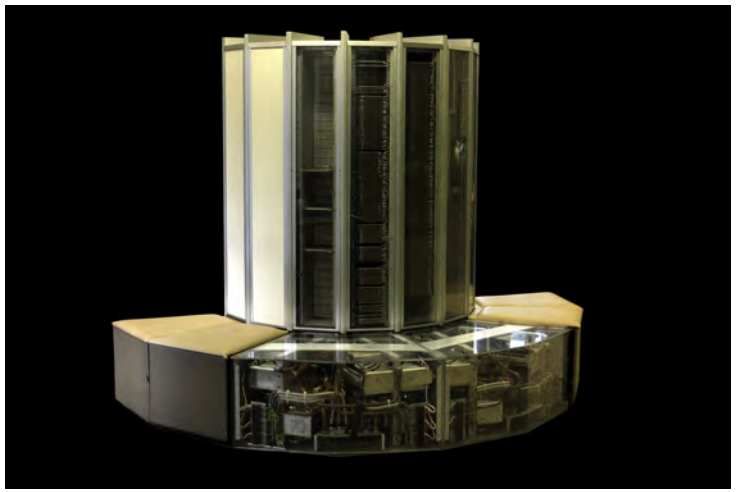


- **Moore's Thesis**
  - Minimize price per device
  - Optimum number of devices / chip increasing 2x / year
- **Later**
  - 2x / 2 years
  - "Moore's Prediction"

Moore's Law: 50 Years



What Moore's Law Has Meant

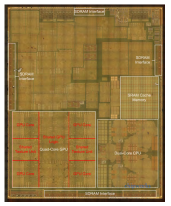


- **1976 Cray 1**
  - 250 M Ops/second
  - ~170,000 chips
  - 0.5B transistors
  - 5,000 kg, 115 KW
  - $9M
  - 80 manufactured
- **2014 iPhone 6**
  - > 4 B Ops/second
  - ~10 chips
  - > 3B transistors
  - 120 g, < 5 W
  - $649
  - 10 million sold in first 3 days

What Moore's Law Has Meant

- 1965 Consumer Product

- 2015 Consumer Product

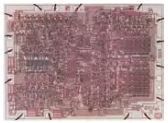


Apple A8 Processor
2 B transistors

Visualizing Moore's Law to Date

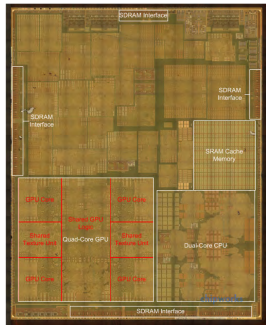
If transistors were the size of a grain of sand

Intel 4004
1970
2,300 transistors



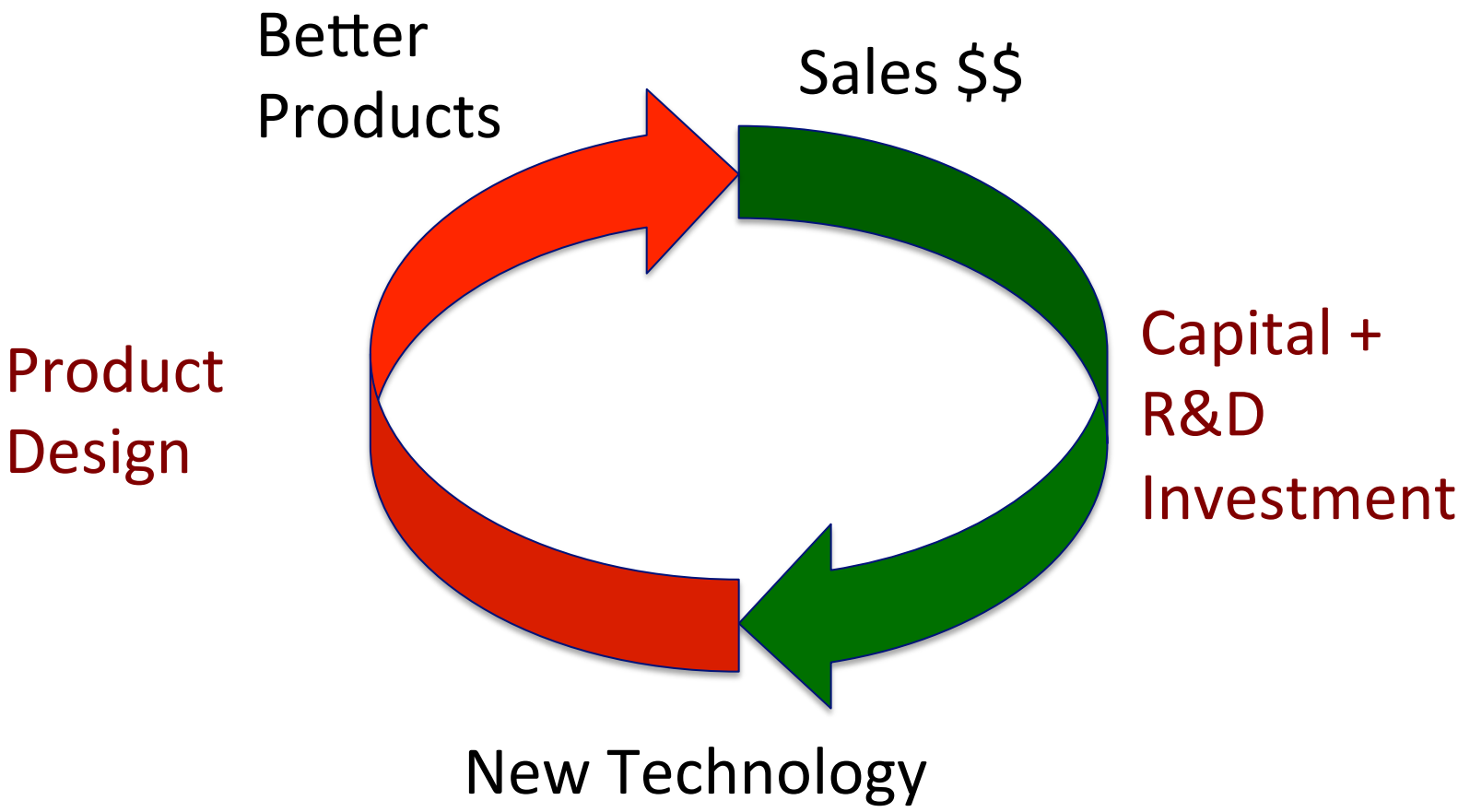
0.1 g

Apple A8
2014
2 B transistors



88 kg

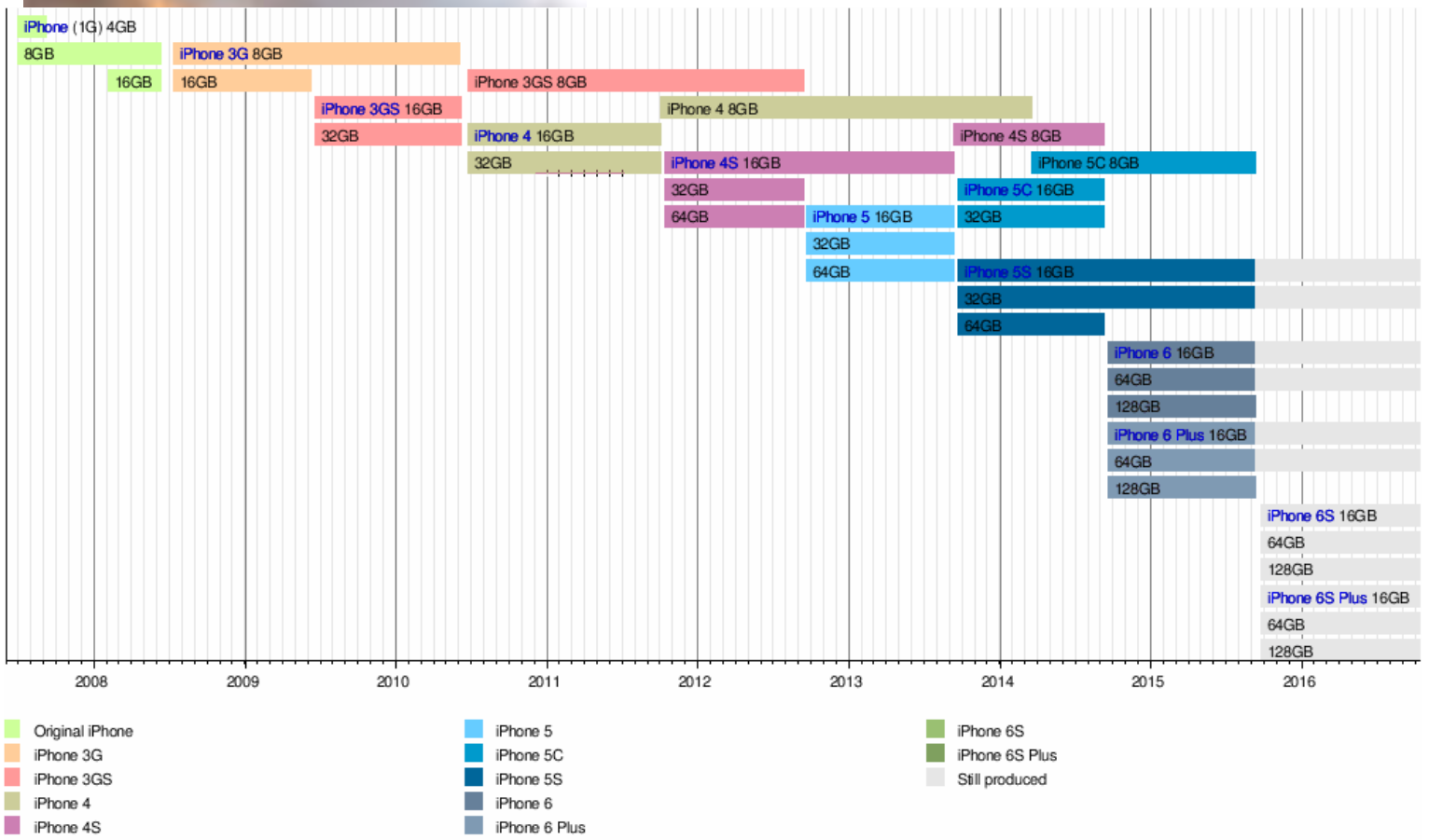
Moore's Law Economics



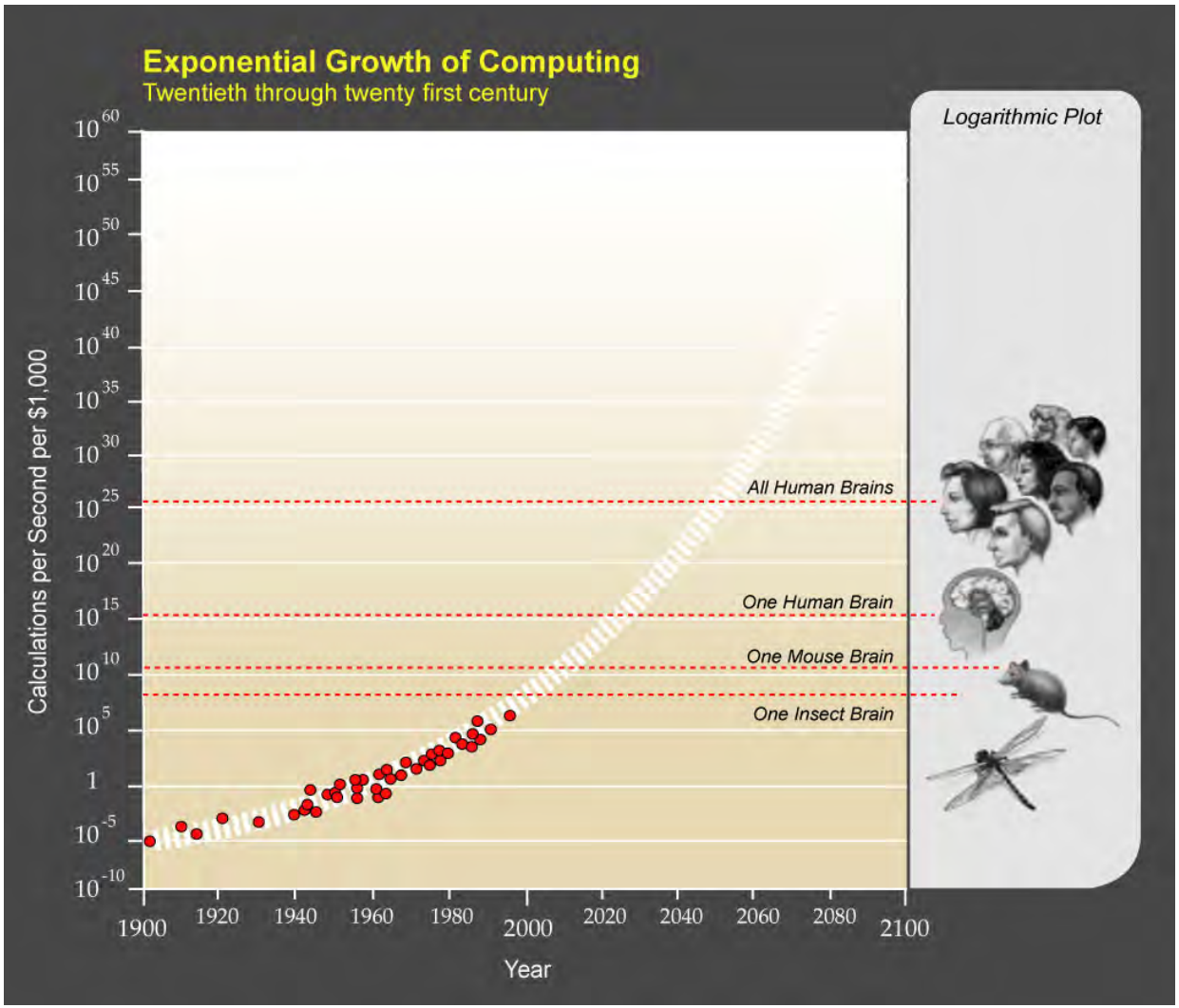
Consumer products sustain the $300B semiconductor industry

What Moore's Law Has Meant

9 generations of iPhone since 2007



What Moore's Law Could Mean



Bryant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

Kurzweil, *The Singularity is Near*, 2005

What Moore's Law Could Mean

- **2015 Consumer Product**

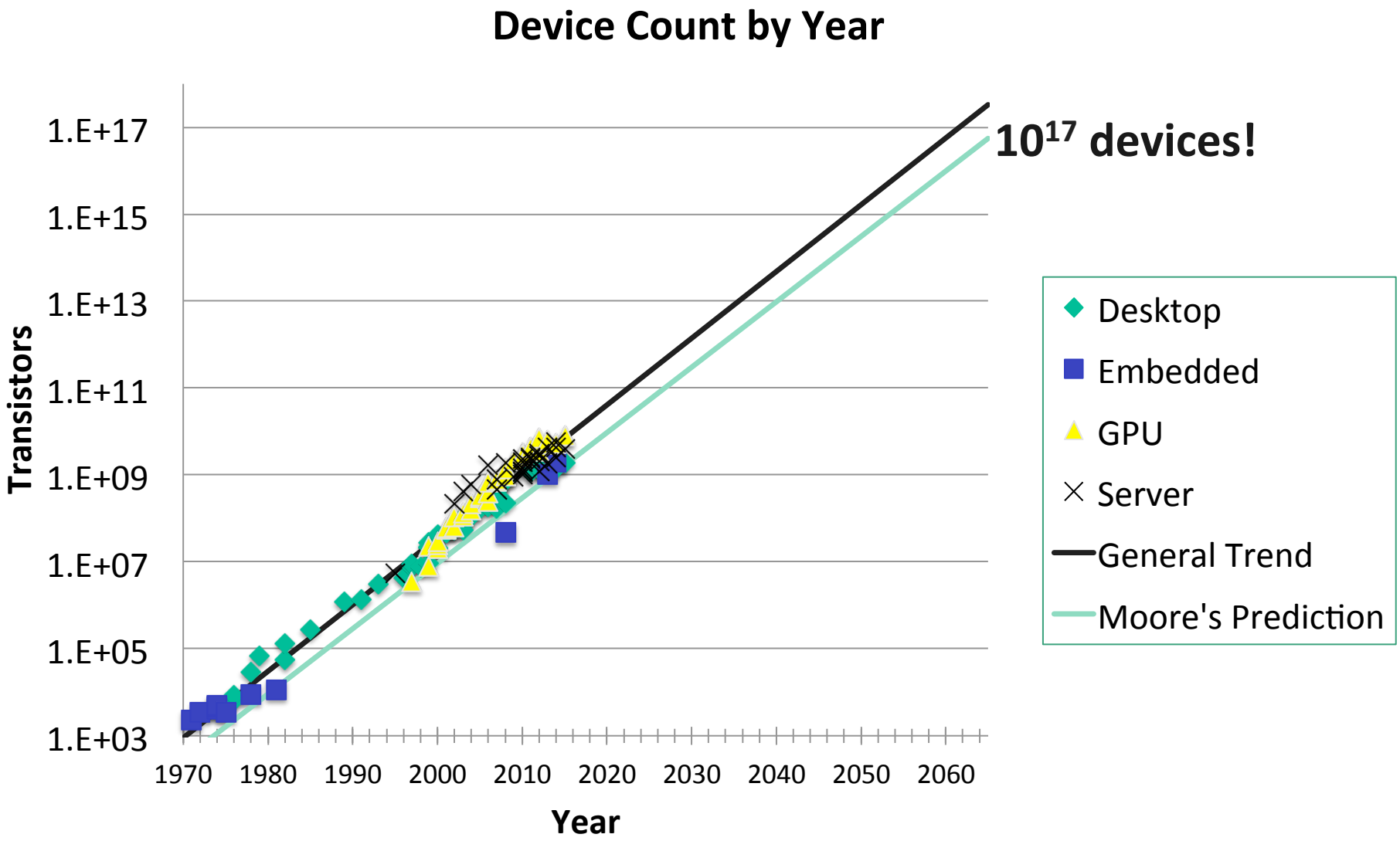
- **2065 Consumer Product**

  - Portable
  - Low power
  - Will drive markets & innovation

Requirements for Future Technology

- **Must be suitable for portable, low-power operation**
  - Consumer products
  - Internet of Things components
  - Not cryogenic, not quantum
- **Must be inexpensive to manufacture**
  - Comparable to current semiconductor technology
    - O(1) cost to make chip with O(*N*) devices
- **Need not be based on transistors**
  - Memristors, carbon nanotubes, DNA transcription, ...
  - Possibly new models of computation
  - But, still want lots of devices in an integrated system

Moore's Law: 100 Years



Visualizing 10^{17} Devices

If devices were the size of a grain of sand

0.1 m^3

3.5 X 10^9 grains

1 million m^3

0.35 X 10^{17} grains

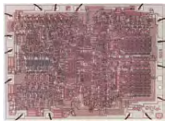
Increasing Transistor Counts

1. **Chips have gotten bigger**
   - 1 area doubling / 10 years
2. **Transistors have gotten smaller**
   - 4 density doublings / 10 years

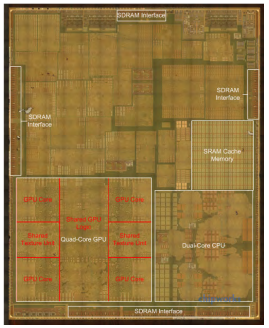
Will these trends continue?

Chips Have Gotten Bigger

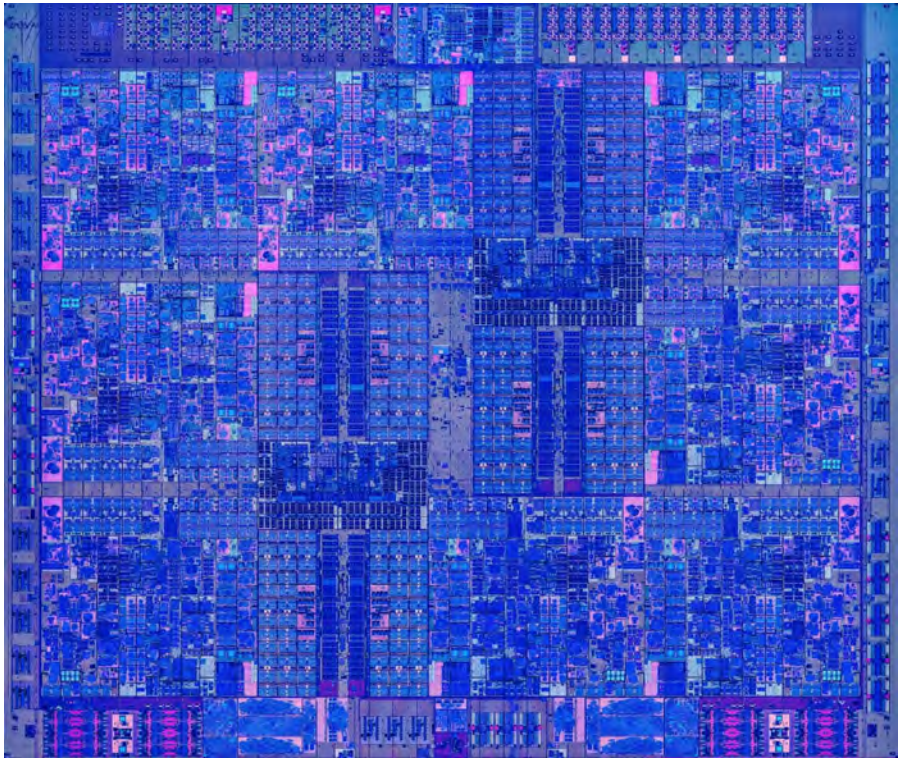
Intel 4004
1970
2,300 transistors
12 mm²



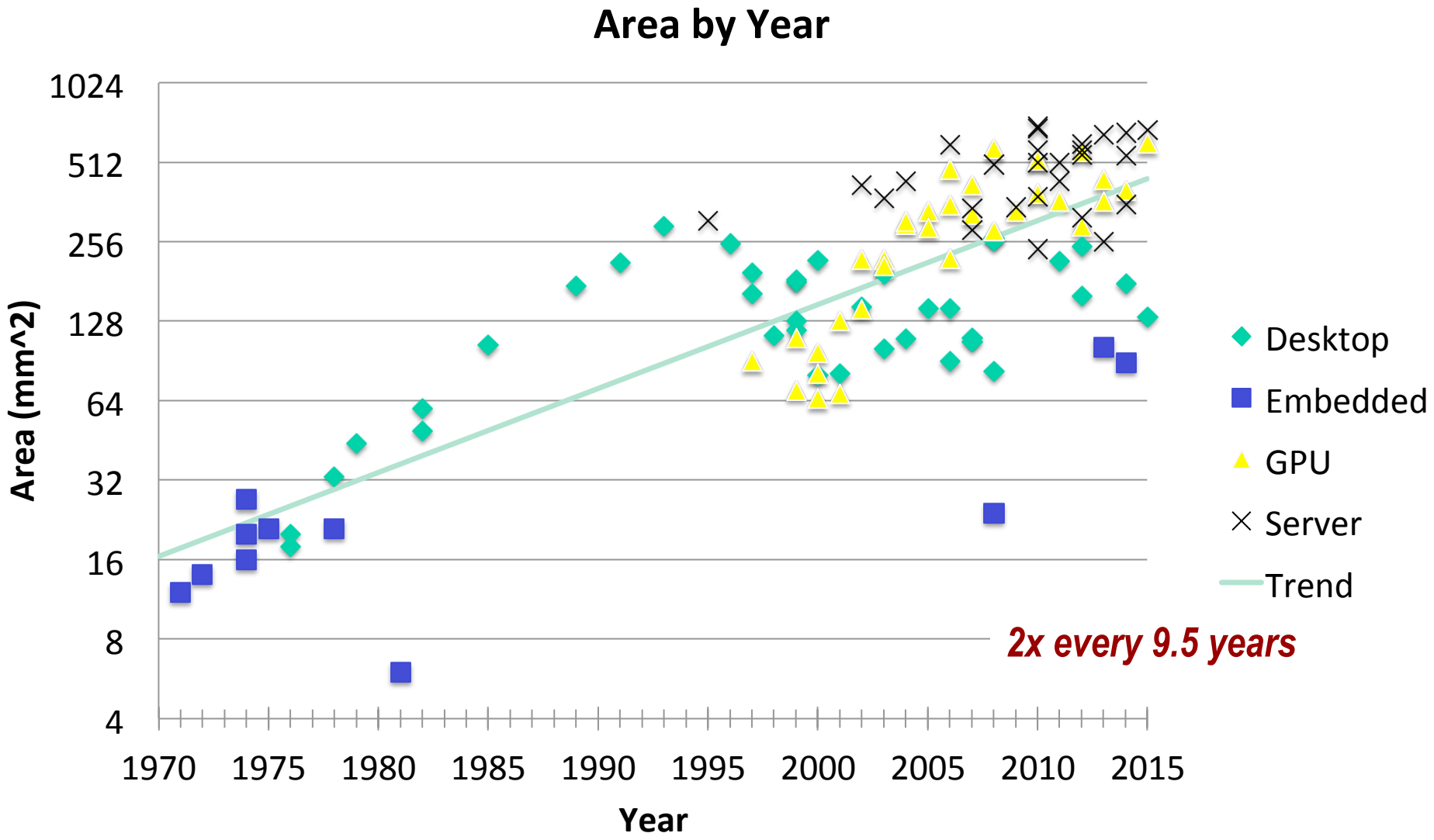
Apple A8
2014
2 B transistors
89 mm²



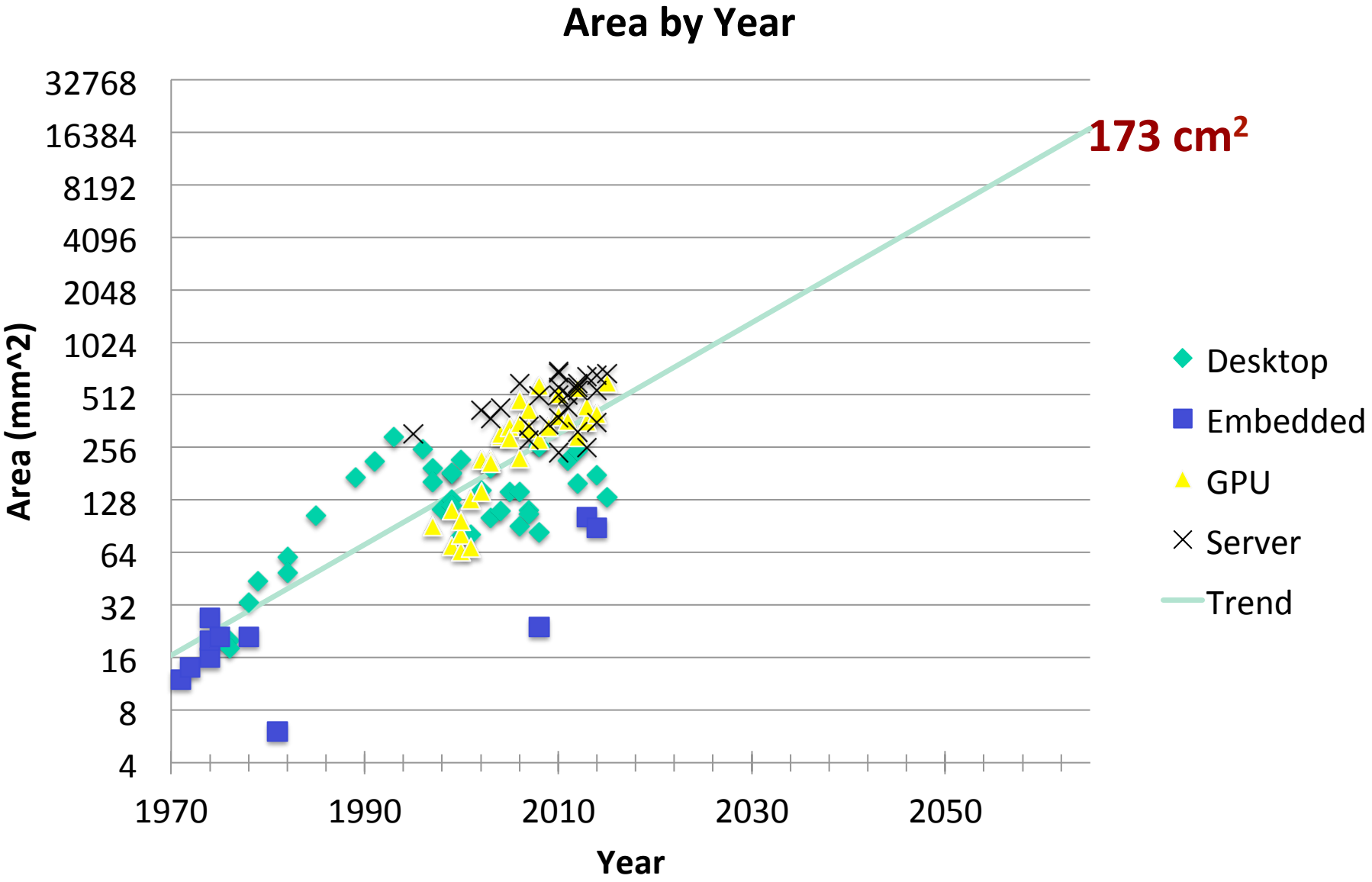
IBM z13
205
4 B transistors
678 mm²



Chip Size Trend



Chip Size Extrapolation



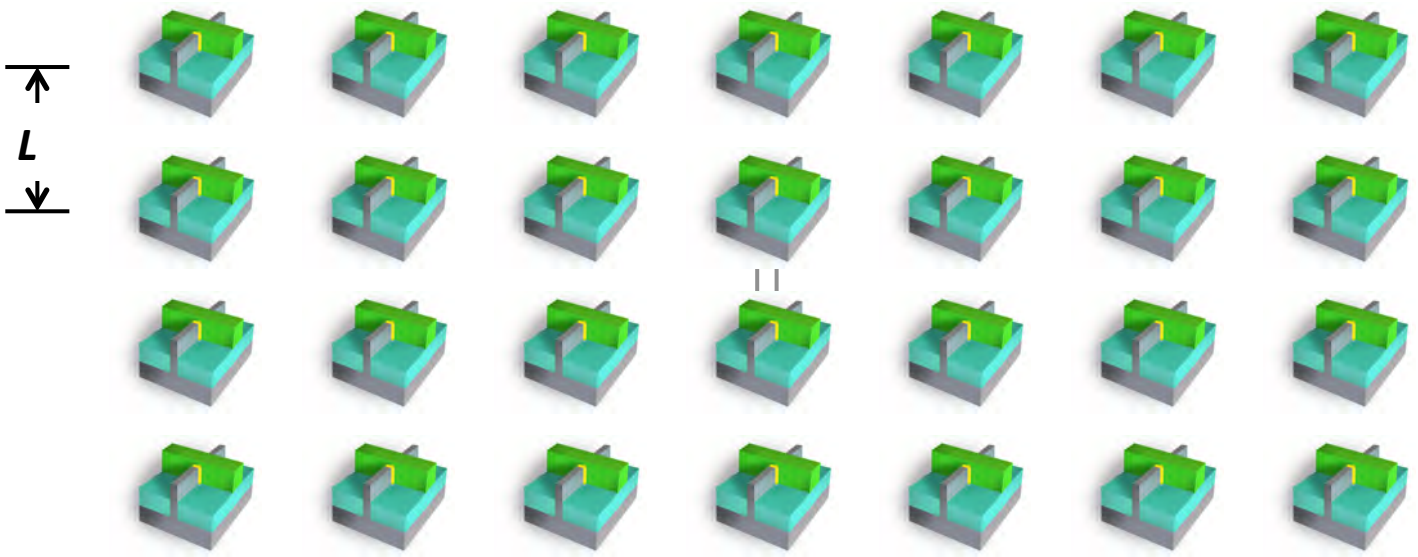
Extrapolation: The iPhone 31s

Apple A59
2065
10^{17} transistors
173 cm^2

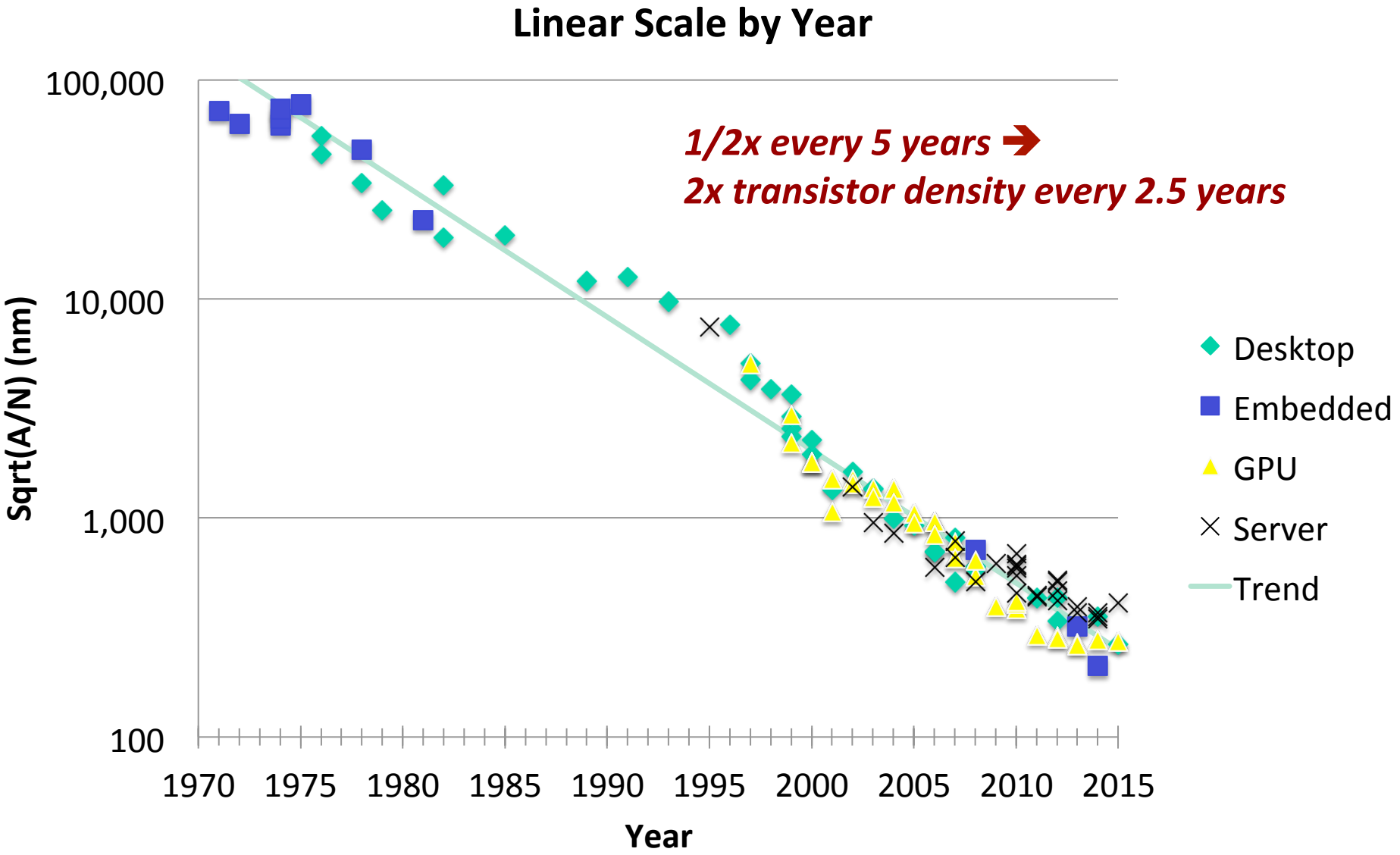
Transistors Have Gotten Smaller

- Area *A*
- *N* devices
- Linear Scale *L*

$$L = \sqrt{A / N}$$

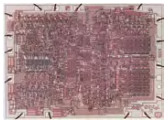


Linear Scaling Trend

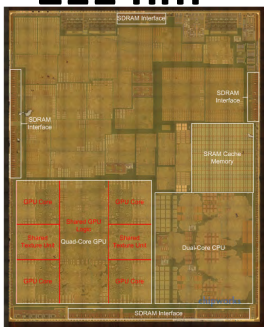


Decreasing Feature Sizes

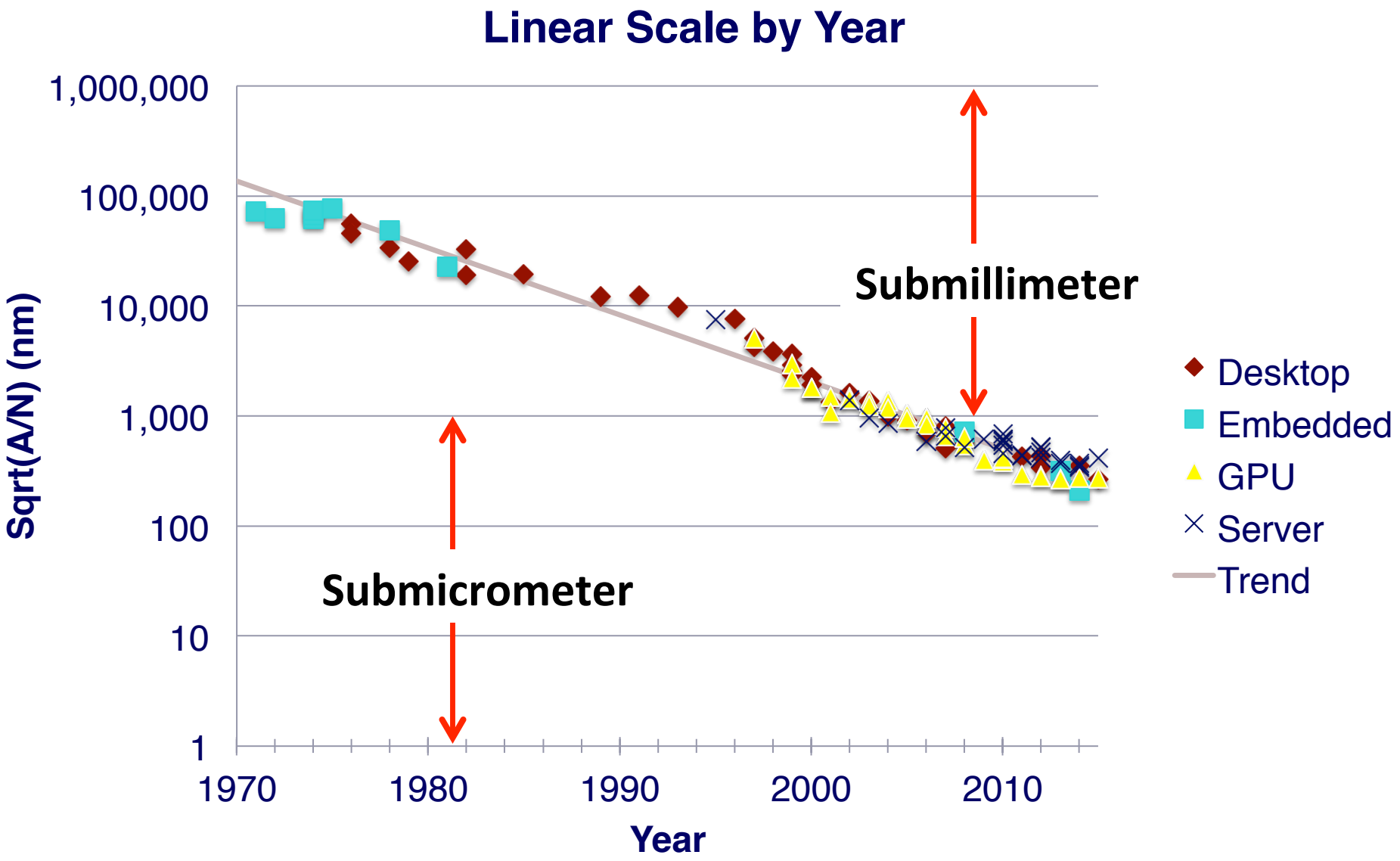
Intel 4004
1970
2,300 transistors
***L* = 72,000 nm**



Apple A8
2014
2 B transistors
***L* = 211 nm**



Linear Scaling Trend

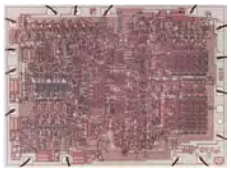


Submillimeter Dimensions

10^{-3} 1 millimeter (mm)

500µm: Length of amoeba

10^{-4}



72µm: Intel 4004 linear scale

50µm: Average size of cell in human body

10^{-5}

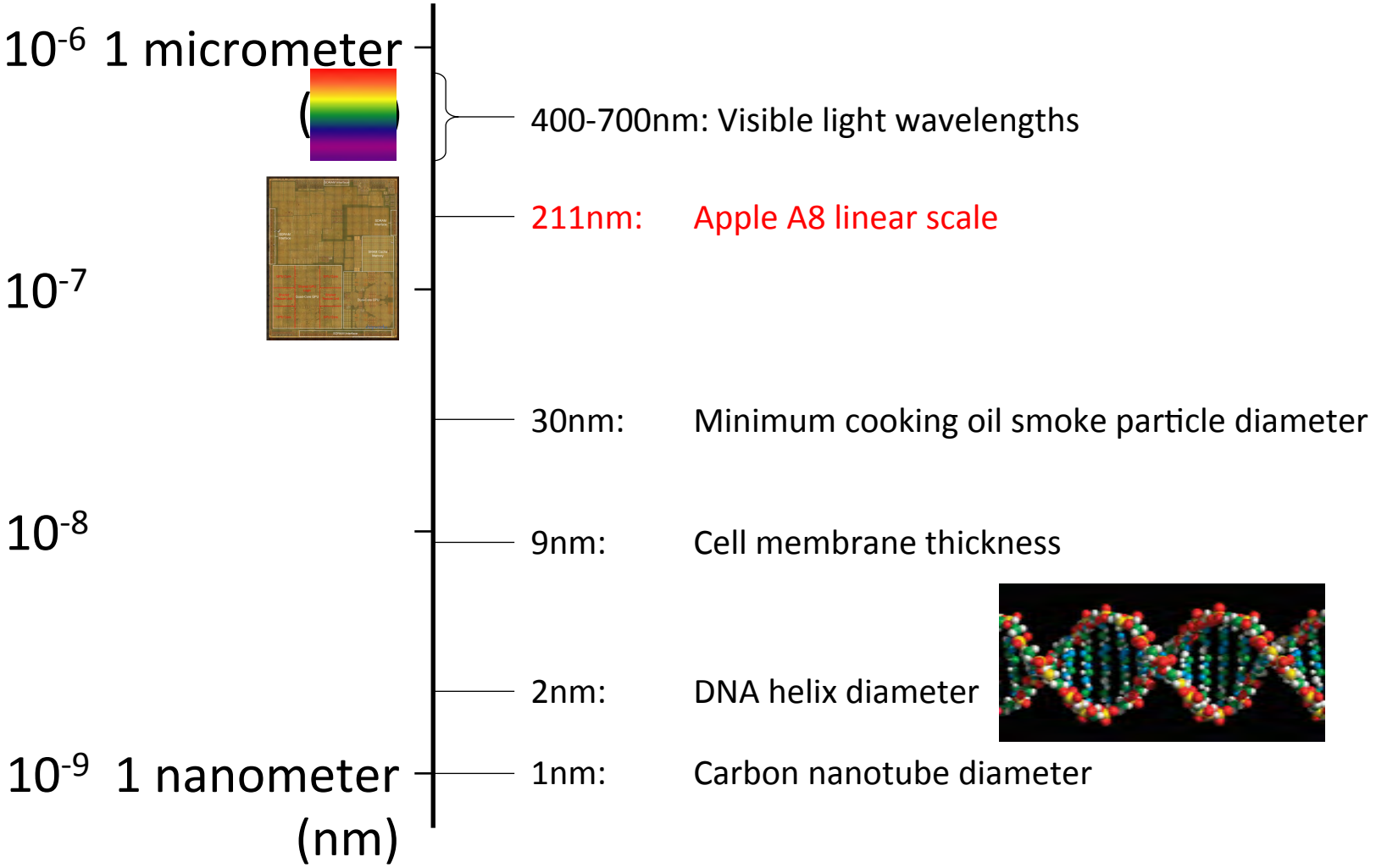
10µm: Thickness of sheet of plastic food wrap

5µm: Spider silk thickness

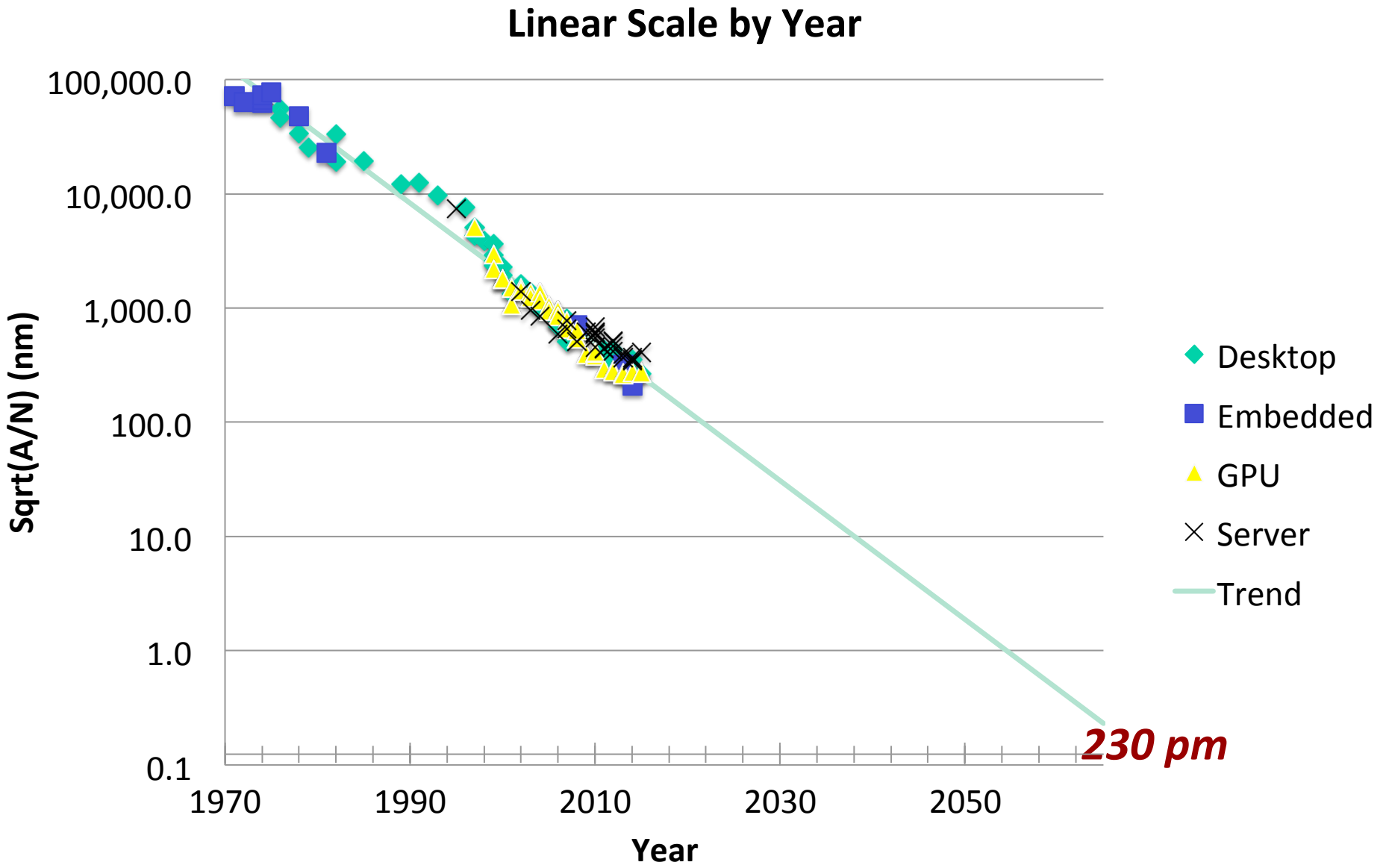
2µm: E coli bacterium length

10^{-6} 1 micrometer (µm)

Submicrometer Dimensions



Linear Scaling Extrapolation

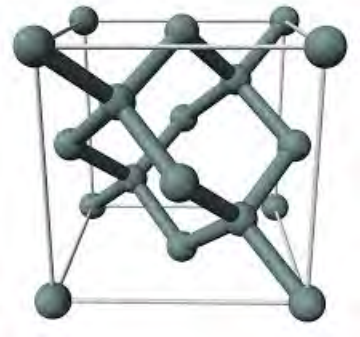


Subnanometer Dimensions

10^{-9} 1 nanometer (nm)

1nm: Carbon nanotube diameter

543pm: Silicon crystal lattice spacing

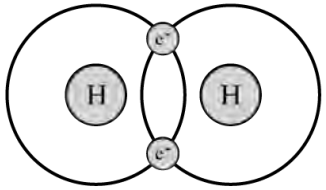


230pm: 2065 linear scale projection

10^{-10}

74pm: Spacing between atoms in hydrogen molecule

53pm: Electron-proton spacing in hydrogen (Bohr radius)



10^{-11}

2.4pm: Electron wavelength (Compton wavelength)

10^{-12} 1 picometer (pm)

Reaching 2065 Goal

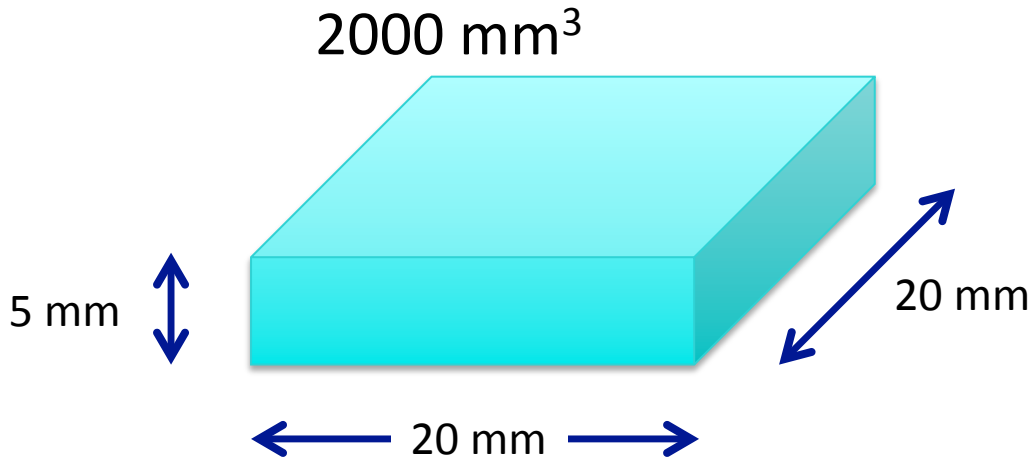
- **Target**
  - 10^{17} devices
  - 400 mm^2
  - L = 63 pm

- **Is this possible?**

No!

Not with 2-d fabrication

Fabricating in 3 Dimensions



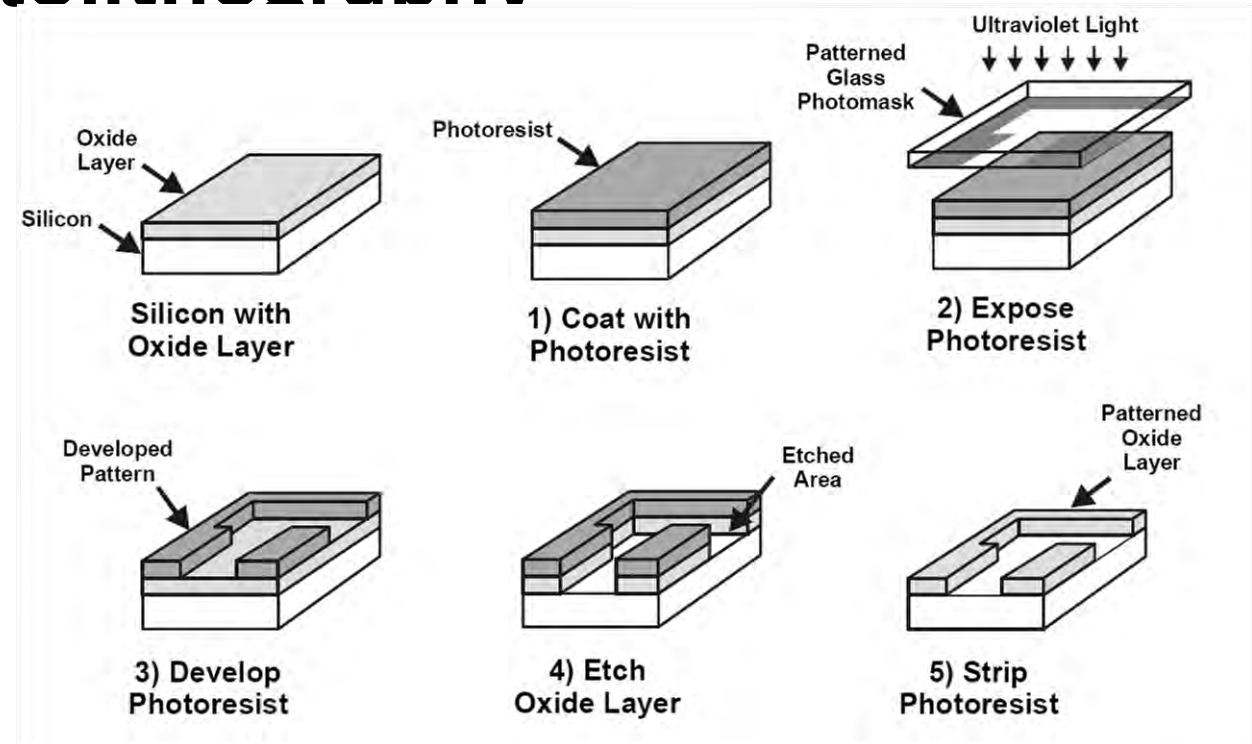
- **Parameters**
  - 10^{17} devices
  - 100,000 logical layers
    - Each 50 nm thick
    - ~1,000,000 physical layers
      - To provide wiring and isolation
  - *L* = 20 nm
    - 10x smaller than today

2065 mm³

3D Fabrication Challenges

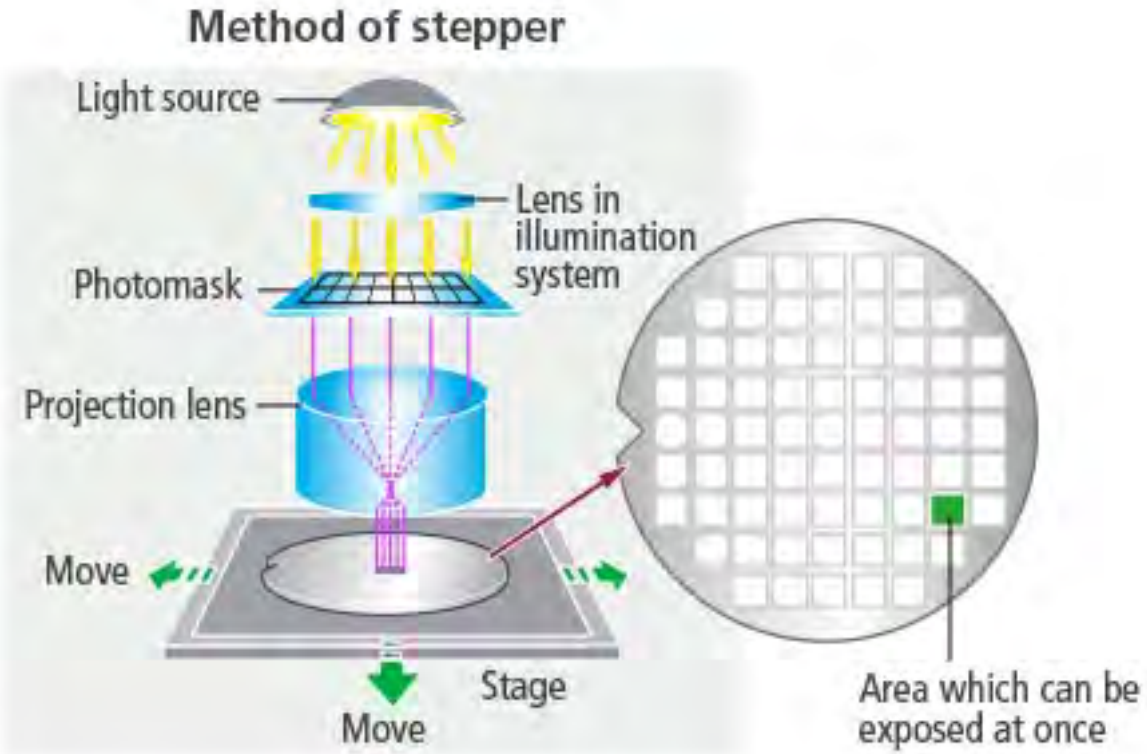
- **Yield**
  - How to avoid or tolerate flaws
- **Cost**
  - High cost of lithography
- **Power**
  - Keep power consumption within acceptable limits
  - Limited energy available
  - Limited ability to dissipate heat

Photolithography



- Pattern entire chip in one step
- Modern chips require ~60 lithography steps
- Fabricate *N* transistor system with O(1) steps

Fabrication Costs

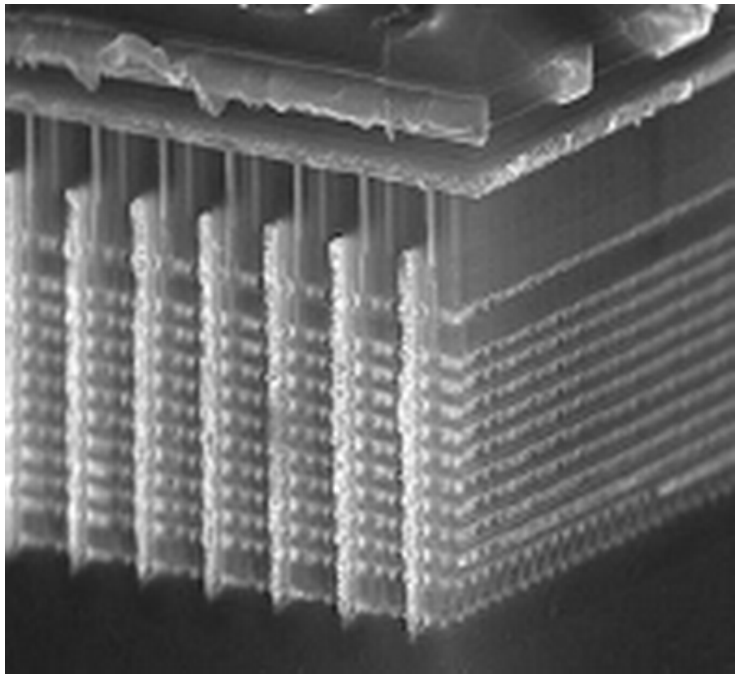


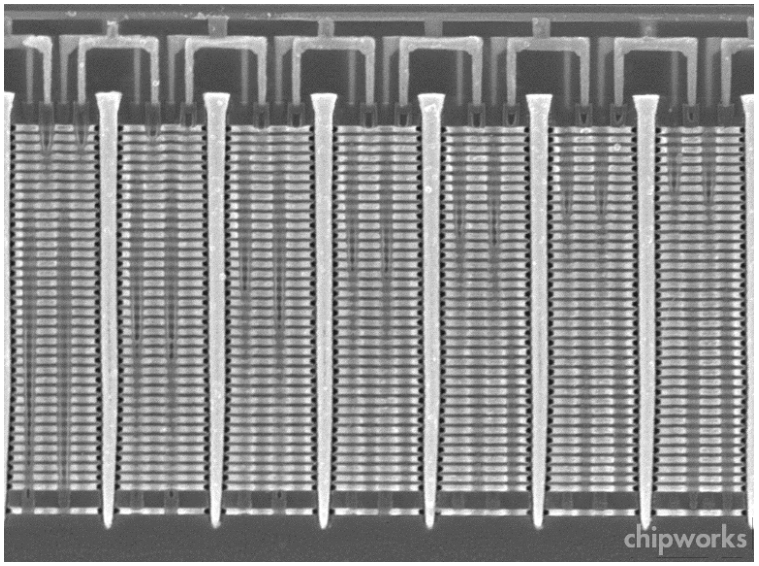
- **Stepper**
  - Most expensive equipment in fabrication facility
  - Rate limiting process step
    - 18s / wafer
  - Expose 858 mm^2 per step
    - 1.2% of chip area

Fabrication Economics

- **Currently**
  - Fixed number of lithography steps
  - Manufacturing cost $10–$20 / chip
    - Including amortization of facility
- **Fabricating 1,000,000 physical layers**
  - Cannot do lithography on every step
- **Options**
  - Chemical self assembly
    - Devices generate themselves via chemical processes
  - Pattern multiple layers at once

Samsung V-Nand Flash Example





- Build up layers of unpatterned material
- Then use lithography to slice, drill, etch, and deposit material across all layers
- ~30 total masking steps
- Up to 48 layers of memory cells
- Exploits particular structure of flash memory circuits

Meeting Power Constraints

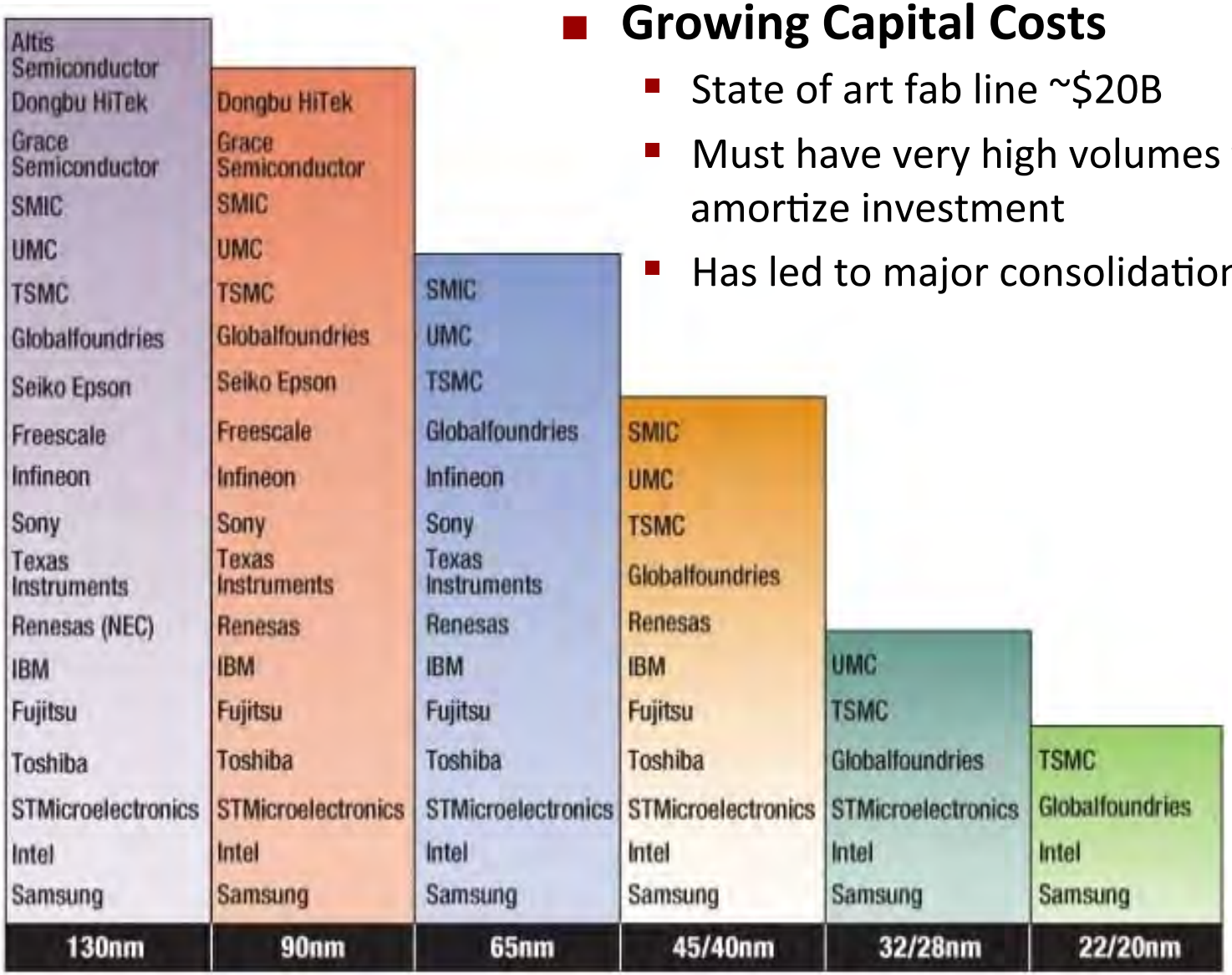
- 2 B transistors
- 2 GHz operation
- 1—5 W

Can we increase number of devices by 500,000x without increasing power requirement?

- 64 B neurons
- 100 Hz operation
- 15—25 W
  - Liquid cooling
  - Up to 25% body's total energy consumption

Challenges to Moore's Law: Economic

- **Growing Capital Costs**
  - State of art fab line ~$20B
  - Must have very high volumes to amortize investment
  - Has led to major consolidations

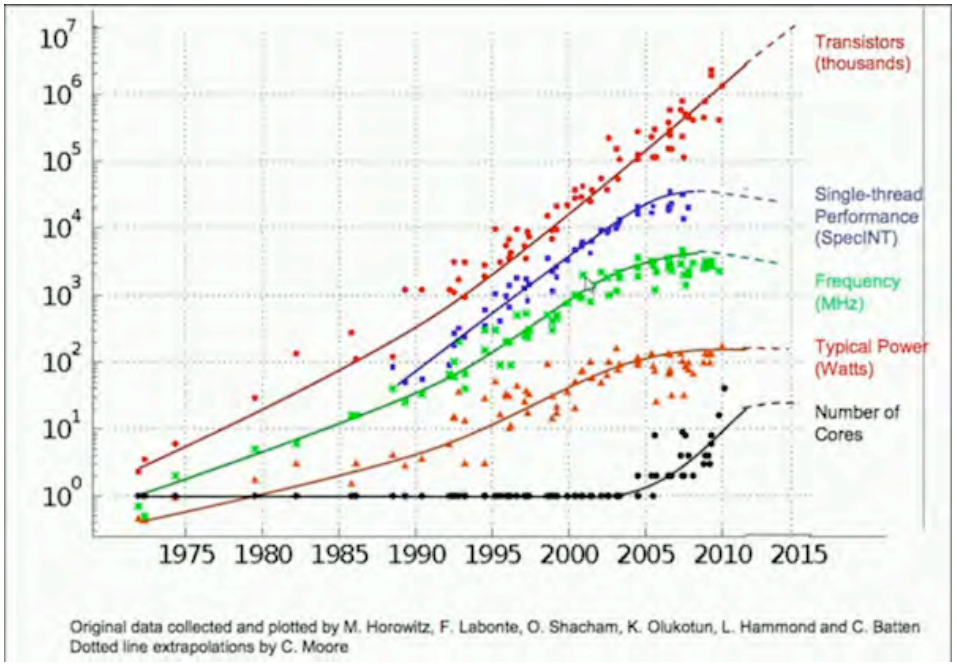


Dennard Scaling

- Due to Robert Dennard, IBM, 1974
- Quantifies benefits of Moore's Law

- **How to shrink an IC Process**
  - Reduce horizontal and vertical dimensions by *k*
  - Reduce voltage by *k*
- **Outcomes**
  - Devices / chip increase by k^2
  - Clock frequency increases by *k*
  - Power / chip constant
- **Significance**
  - Increased capacity and performance
  - No increase in power

End of Dennard Scaling



- **What Happened?**
  - Can't drop voltage below ~1V
  - Reached limit of power / chip in 2004
  - More logic on chip (Moore's Law), but can't make them run faster
    - Response has been to increase cores / chip

Final Thoughts

- **Compared to future, past 50 years will seem fairly straightforward**
  - 50 years of using photolithography to pattern transistors on two-dimensional surface
- **Questions about future integrated systems**
  - Can we build them?
  - What will be the technology?
  - Are they commercially viable?
  - Can we keep power consumption low?
  - What will we do with them?
  - How will we program / customize them?