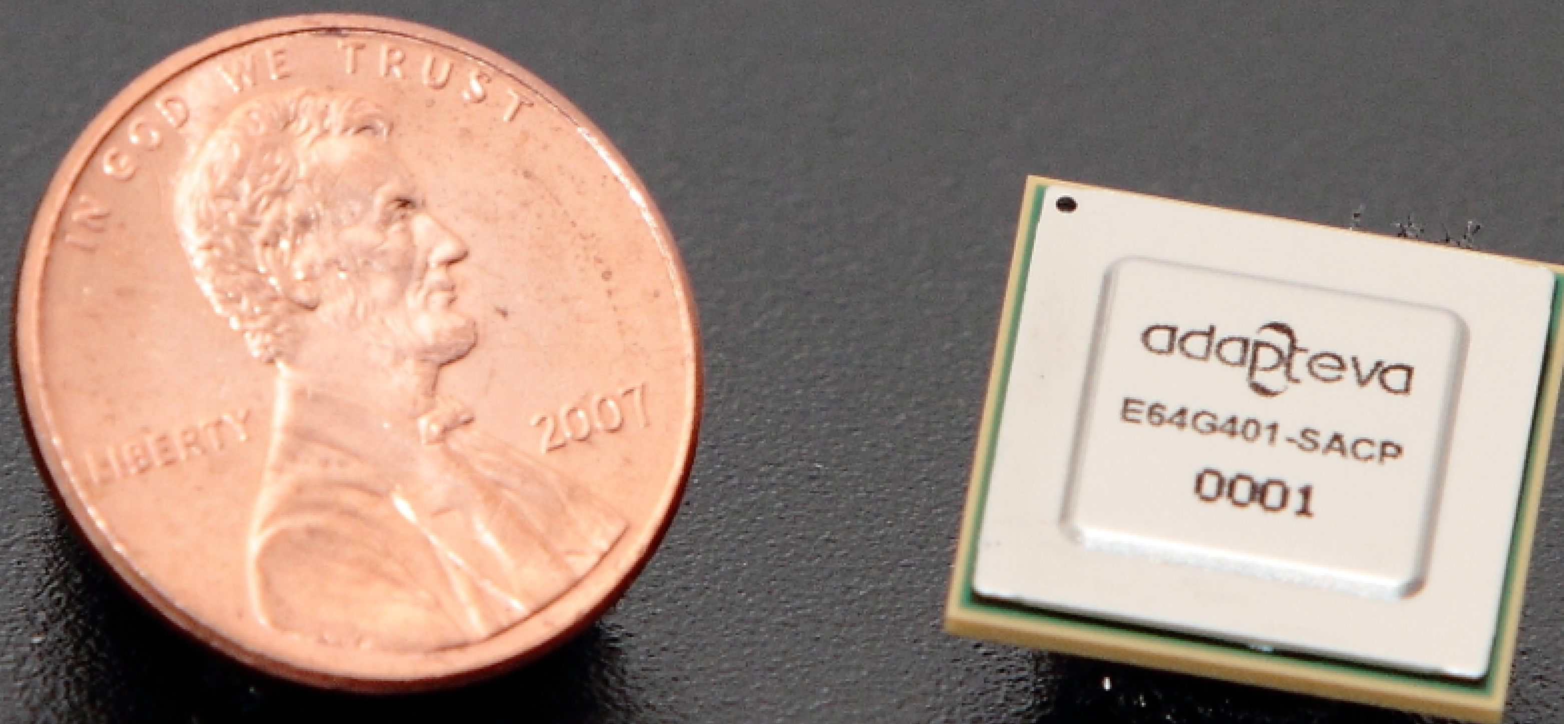
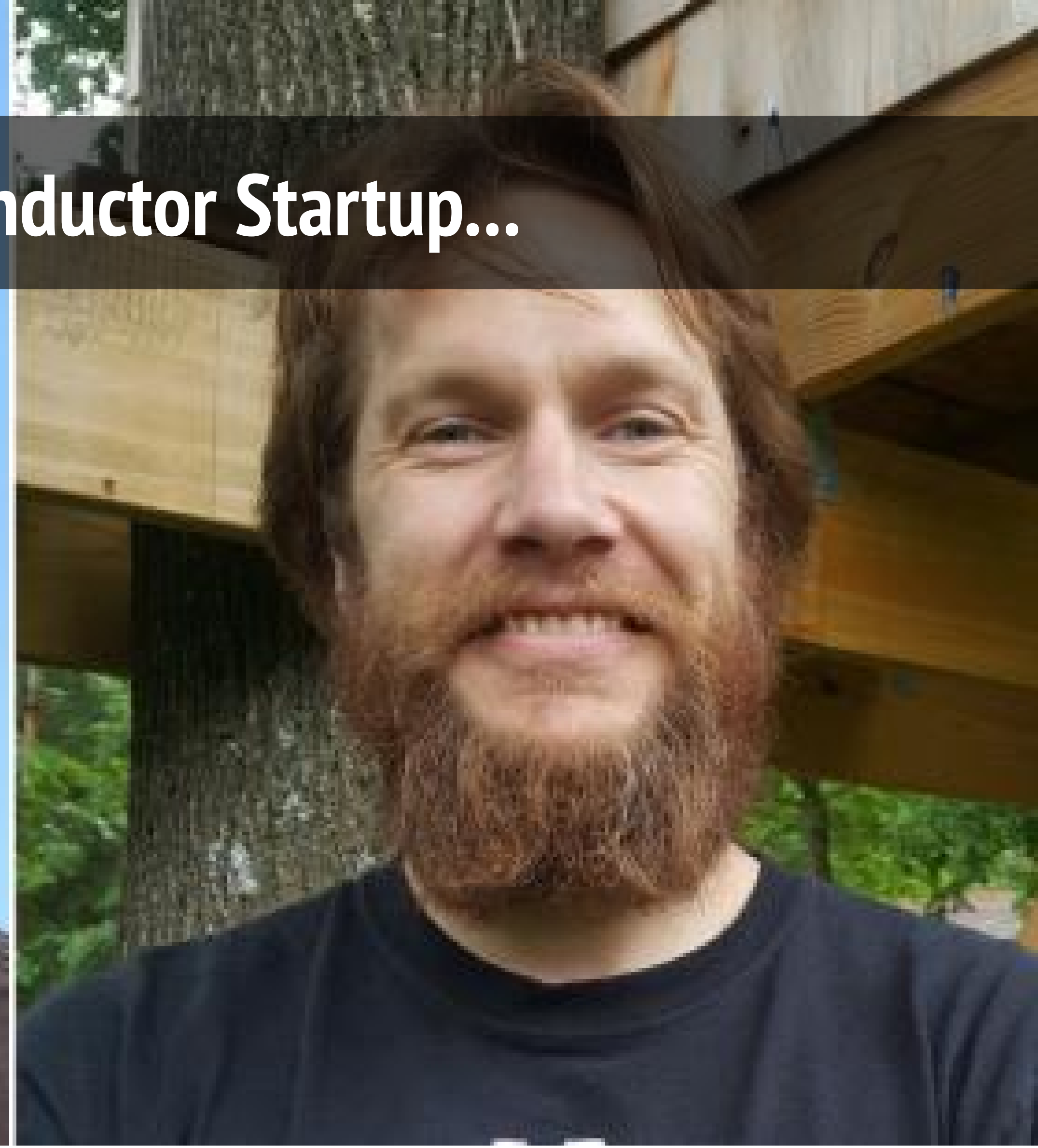
# Computing in the Post-Moore Era

*by Andreas Olofsson*



# Before & After Semiconductor Startup...



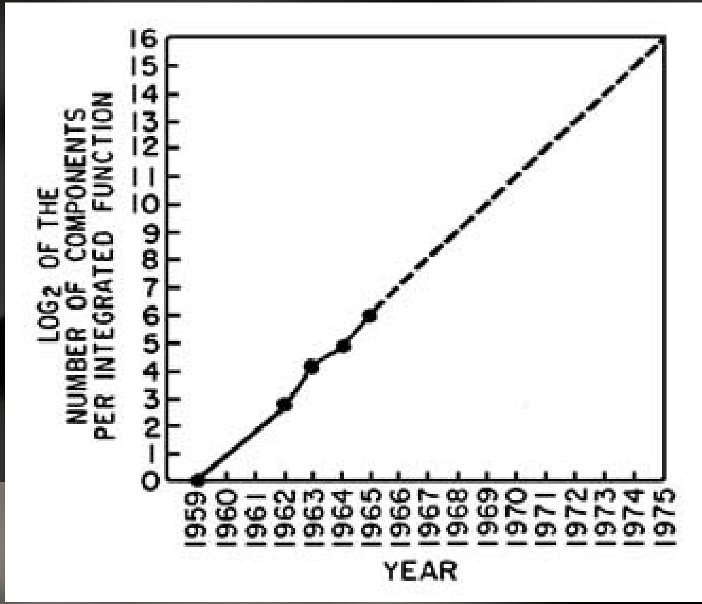
# Talk Outline

- Moore's Law Primer
- Moore's Law Impact
- Predicting the Future (provided ASIS)

# Moore's Law Definition

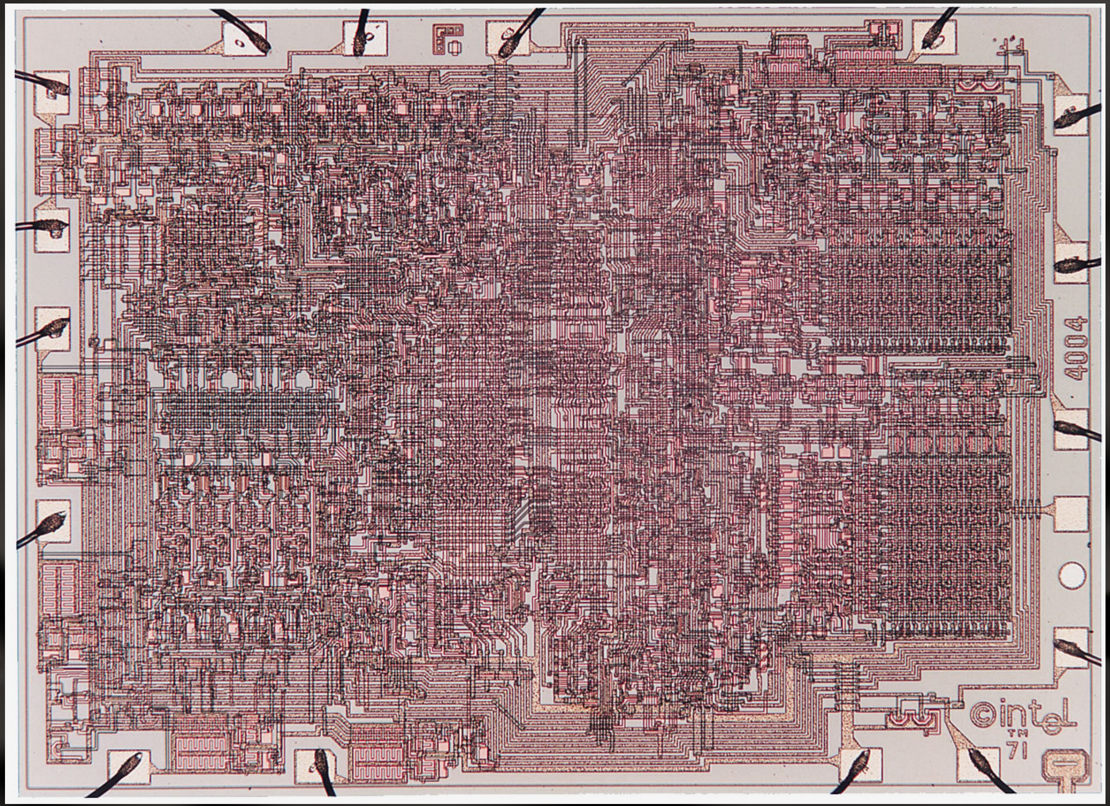
> *The observation that the number of transistors in a dense integrated circuit doubles approximately every two years.*

- All about the $$$
- Since 1955 cost/xtor reduced by 10B
- Don't confuse with performance!
- Profound societal impact



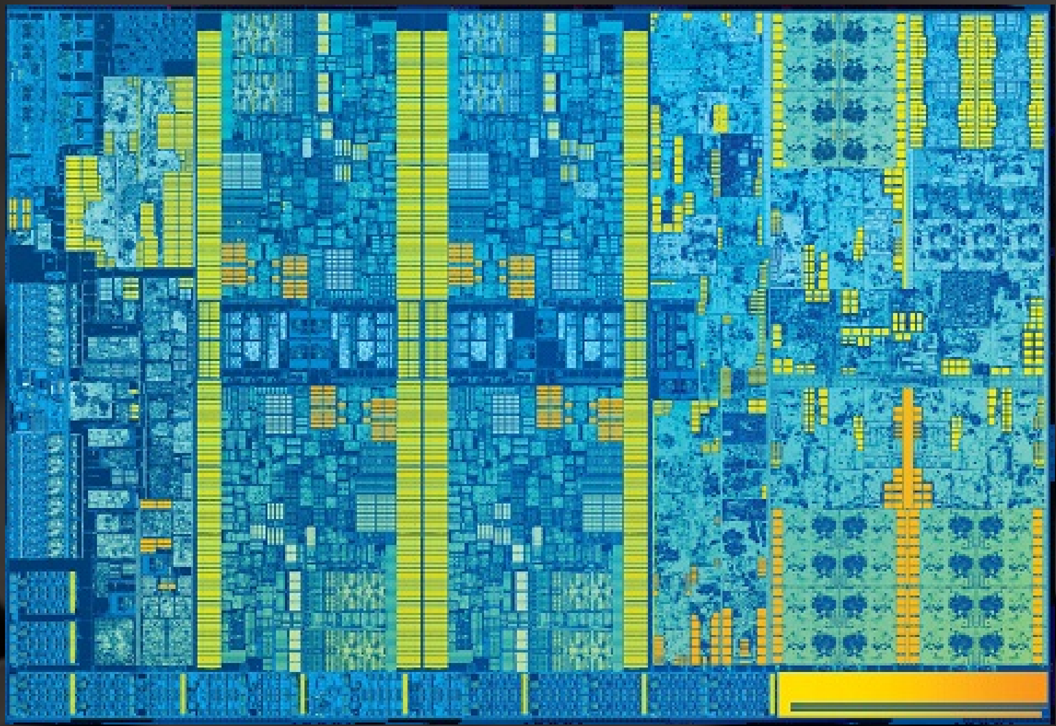
# Moore's Law in 1971

- Intel 4004
- 1 core
- 2,300 xtors
- 12mm^2
- 740Khz
- 4-bit processor
- 10um process

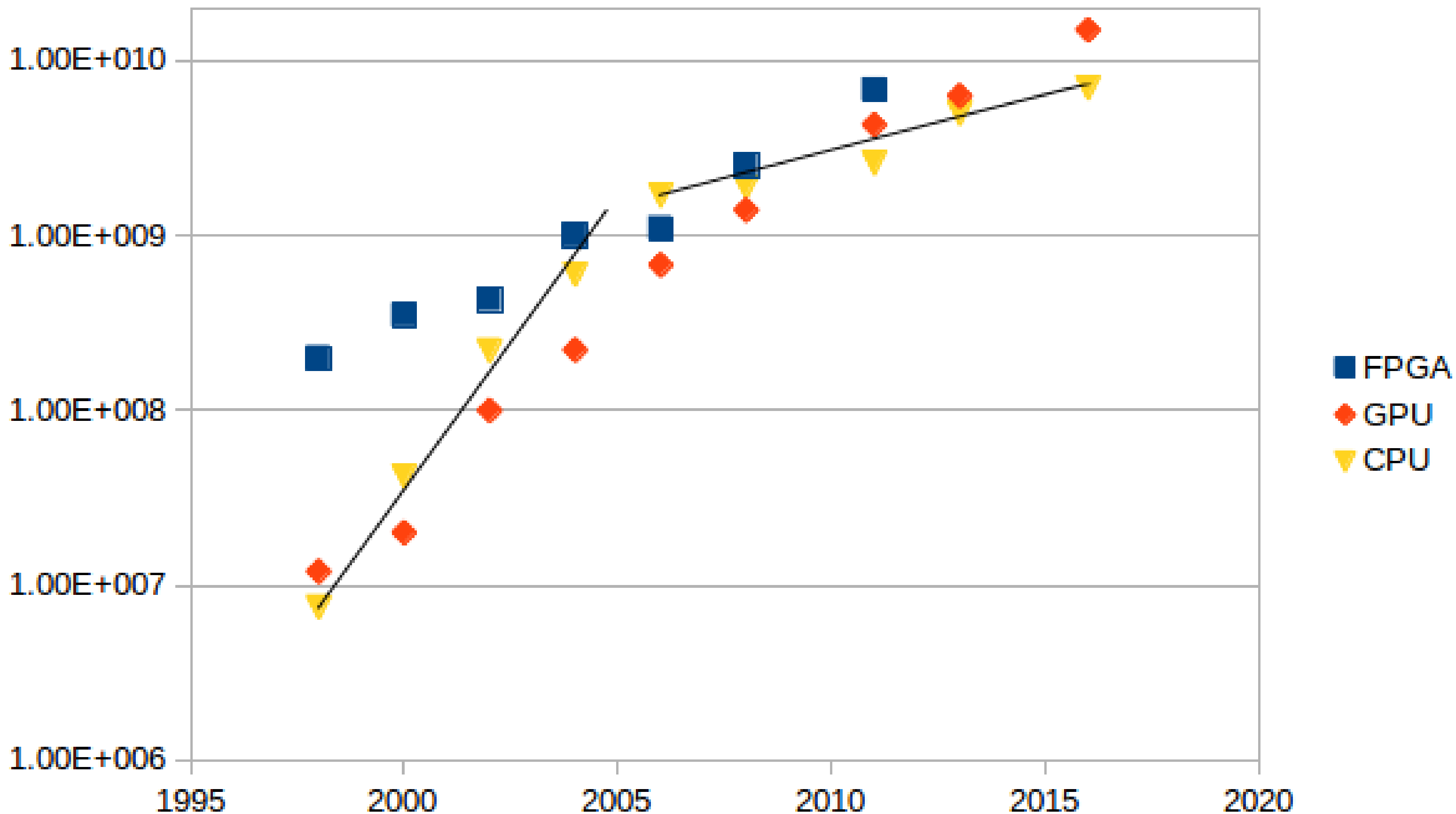


# Moore's Law in 2016

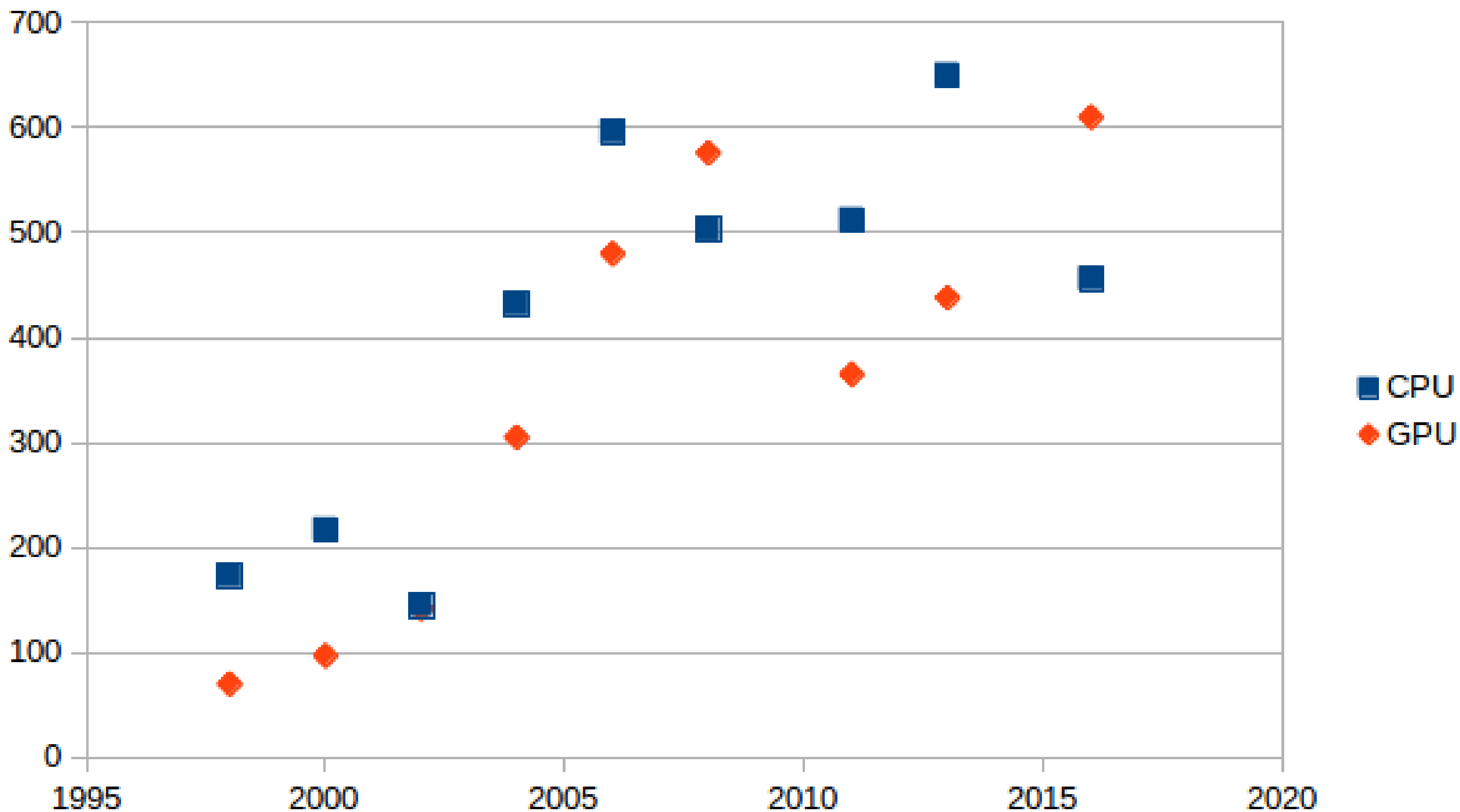
- Intel Broadwell
- 22 cores
- 7B xtors
- 456mm^2
- 4 GHz
- 64-bit processor
- 14nm process (Picture shows Skylake)



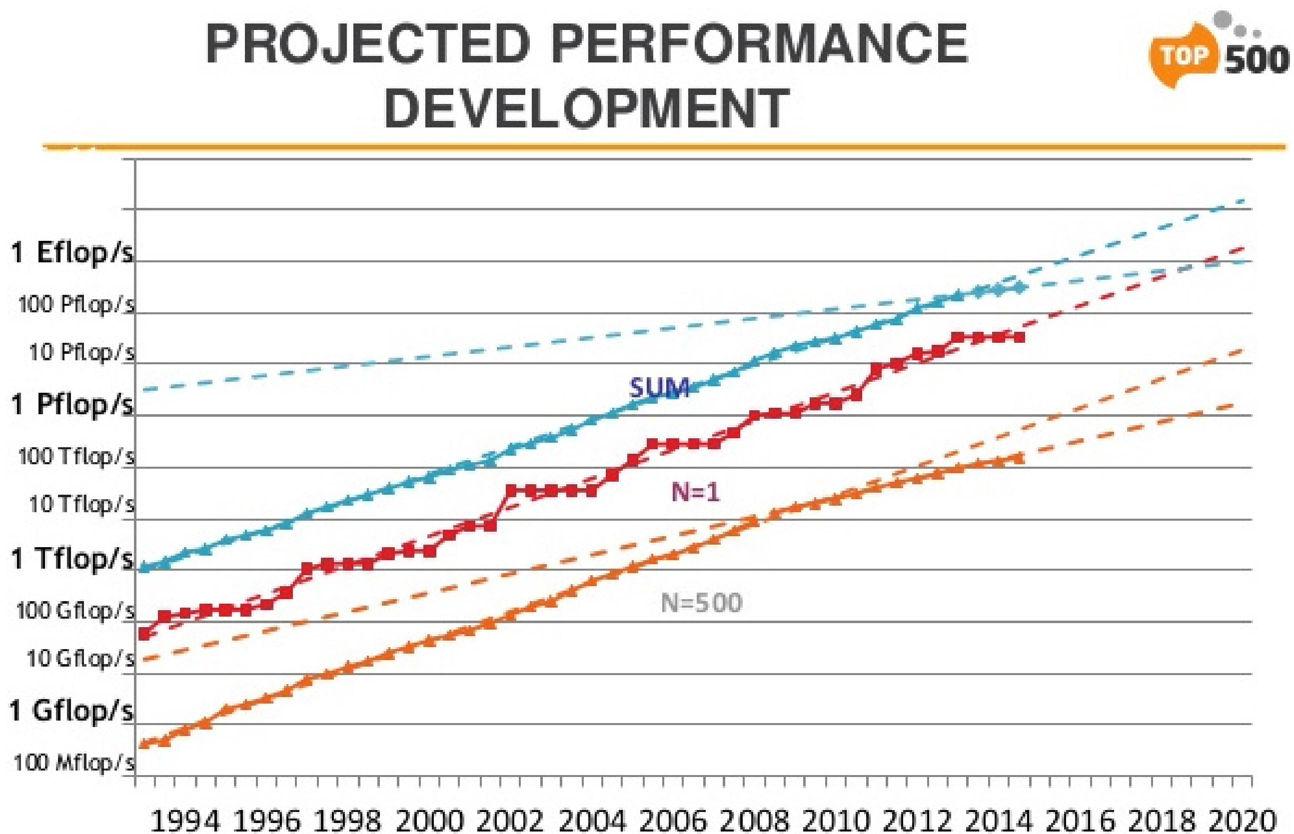
# Moore's Law Transistor Trend



# Moore's Law Area Trend



# Moore Effect: Performance



# Moore Effect: Cost

> *2 Billion people with Thinking machines in their pockets*

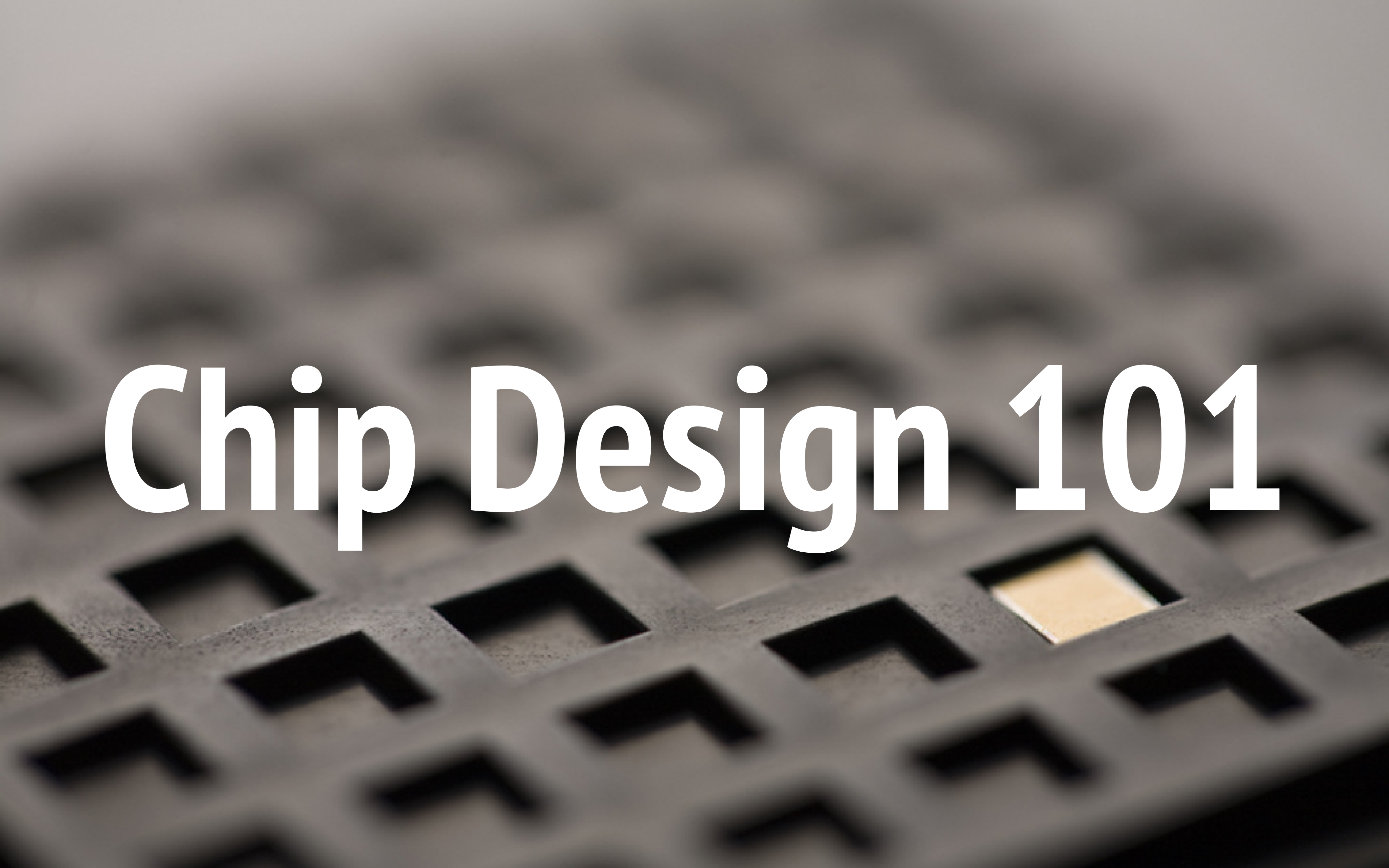
- 100 GFLOPS smart-phones possible
- 5-50B transistors per phone
- Exaflop level connected clouds

# Moore Effect: Democratization

- Parallella: "An open $99 supercomputer"
- "Raspberry Pi for parallel computing"
- 18 CPU cores on a credit card and @ 5W
- Democratized access to parallel computing
- $898K raised on Kickstarter in Oct 2012
- First ever crowd funded chip!
- Almost 100 publications
- Over 10,000 shipped, available at Amazon & Digikey

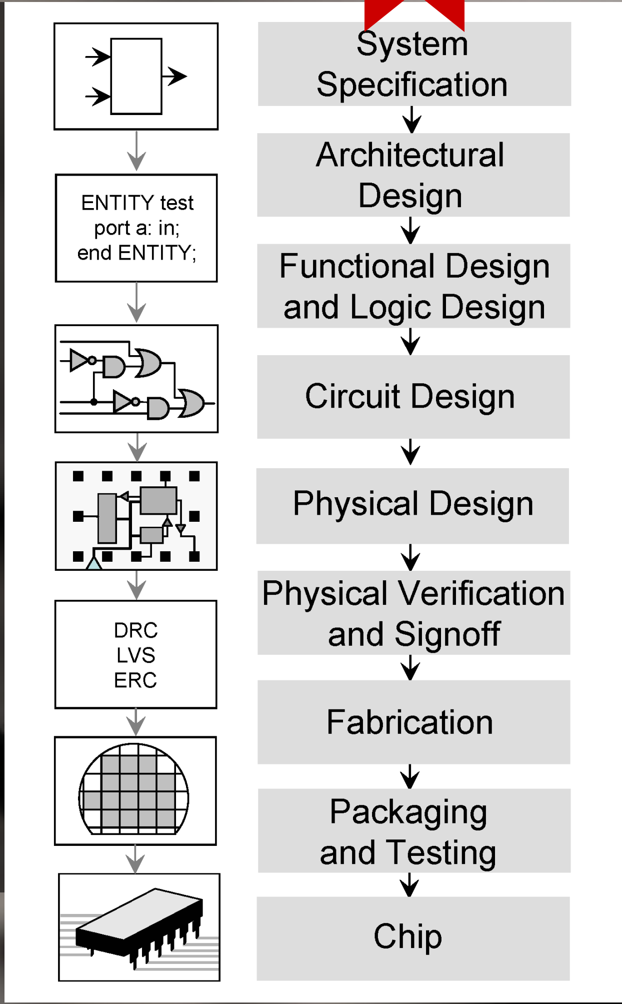
# Moore Effect: Casualties

| | | | | |
|---|---|---|---|---|
| Achronix | Brightscale | Cradle | Mathstar | Sandbridge |
| Adapteva | Calxeda | C-Switch | Mobileye | Silicon Sp. |
| Ambric | Chameleon | ElementCXI | Monarch | Stream Proc |
| Asocs | Clearspeed | Greenarrays | Octasic | Stretch |
| Aspex | Cognivue | Icera | Picochip | Tilera |
| Axis Semi | Coherent L. | Intellasys | Plurality | Transputer |
| BOPS | IBM-Cell | IP-flex | PACT | XMOS |
| Boston C. | CPUTech | Larrabee | Quicksilver | Zilabs |


Chip Design 101

# Chip Design Flow

- ~1 cent / million logic gates
- Arcane languages (Verilog / VHDL)
- 1 year compilation cycle
- $1M / compiler seat
- $1M / hardware bug
- Completely opaque and proprietary flow



# Don't believe the hype!

Adapteva's Story...

- 4 chips in 2 years
- 1-3 engineers
- <$2M spent
- Complexity==$$$

# Real Chip Design Costs

| | |
|---|---|
| Engineering | N*($150K/eng) |
| IP Licensing | $1-10M |
| EDA Tools (Compilers) | $1-10M |
| Tapeout (Tooling) | $5M |
| Chip packaging | $50K |
| Qualification | $1M |
| ***TOTAL*** | ***$1-$1,000M*** |

# Moore's Law Economic Challenges

| Challenge | Industry | Hurdle | Current | Future |
|---|---|---|---|---|
| Open source chip IP | $5B | NIH | $1M+ | $0 |
| Open source EDA | $6B | Complexity | $1M+ | $0 |
| Engineering | -- | Time | 9 months | 24hrs |
| Packaging | $13B | Logistics | $50K | $0 |
| Manufacturing | $40B | Logistics | $2M+ | $1,000* |

# Post-Moore Predictions

- Laws of physics prevail (again)
- Semiconductor goes 3D (again)
- Silicon efficiency becomes important (again)
- Optimization engineering becomes important (again)
- Programming gets hard (again)
- ASICs will make a comeback (again)
- Parallel architectures win!

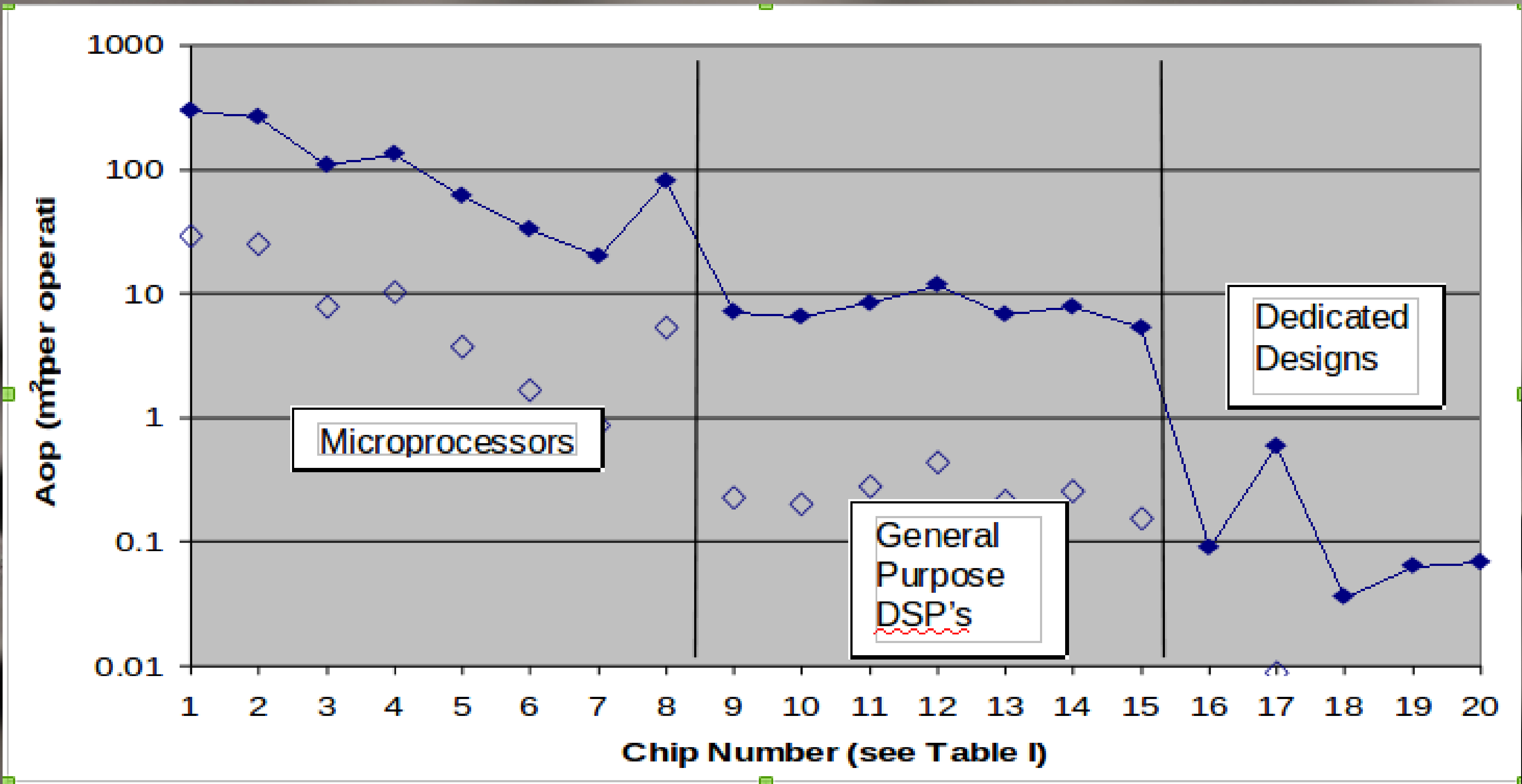
# Physics: Getting Harder!

- Digital Power ~= cap x voltage^2 x freq (derived)
- Switching Delay ~= resistance x capacitance
- Speed Limit = 3 x 10^8 m/s (how far is one nanosecond?)
- Atomic Size Limit ~= 0.1nm
- Cooling ~= Area x dT x HC(v)
- Thermal Noise ~= FUNC(RES,temp, V)

# 3D: Easy! Plenty of Room at the Bottom

| Rule | Value |
|---|---|
| Chip wire pitch | ~0.1um |
| 2.5D wire pitch | 4um |
| Wirebond pitch | 30um |
| 2.5D Bump pitch | 45um |
| Flip-chip pitch | 170um |
| BGA pitch (advanced) | 400um |
| Ethernet connector | ~10,000um |

# Silicon Efficiency (REF: Brodersen)



# Optimization Engineering

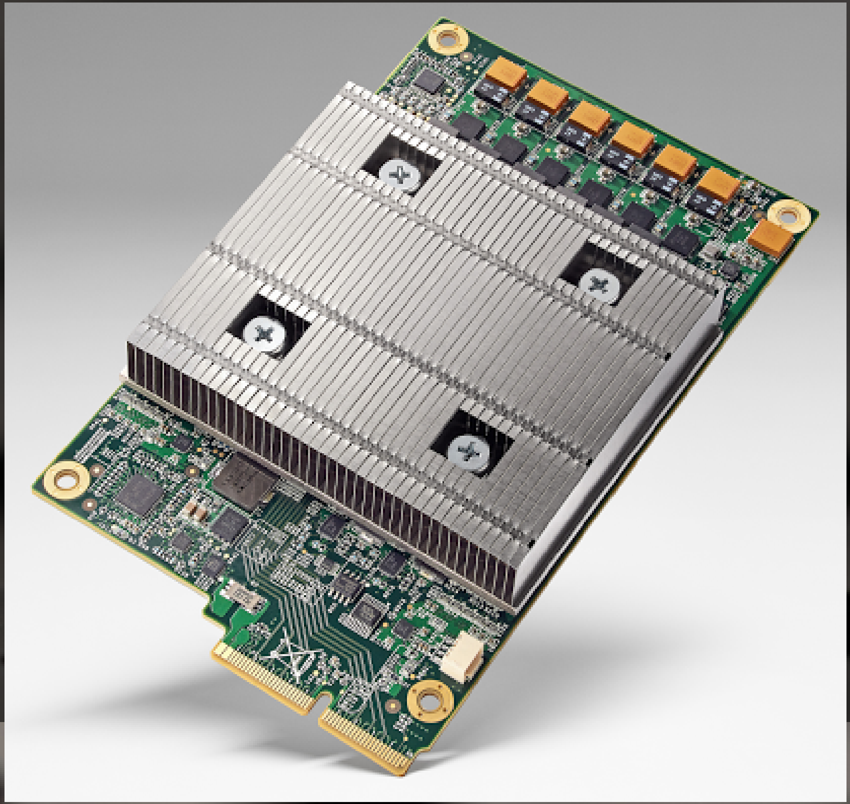
- 200,000 difference between unoptimized Java and assembly
- As things slow down, there is more time for optimization
- Engineers innovate when they have to (free lunch is over)
- Architecture convergence makes optimization effective
- Open source trend making a big difference

# Programming Challenges

| Metric | Chip Designer | Programmer |
|---|---|---|
| Correctness | Always | Always |
| Performance | Always | Sometimes* |
| Parallelism | Always | Sometimes* |
| Timing | Always | Sometimes |
| Size, Power | Always | Sometimes* |
| Fault-tolerance | Often | Rarely* |

# ASICs Making a Comeback

- Tail that wags the dog
- Can't leave 100X on the table
- Design cheaper than ever
- Cisco, Ericsson, Huawei
- Apple (A9x)
- Google (TPU)



# The long tail of electronics

> *"Axiom: Big semiconductor companies only cares about big $$"*

...but what about low volume designs (1-100K units)?

- Health (diagnostics, embedded)
- Robotics (smarter, smaller)
- Communication (free and pervasive)
- Special supercomputers (to answer really tough questions)

# Parallel Computing Will Win Eventually!

(Computing normalized for silicon area at 14/16FF)

| Metric | GPU | CPU | Epiphany Arch |
| --- | --- | --- | --- |
| Performance(FLOPS) | 5,300 | 500 | 10,000 |
| Area (mm^2) | 610 | 456 | 600 |
| Power(W) | 300+150 | 150 | 120 |