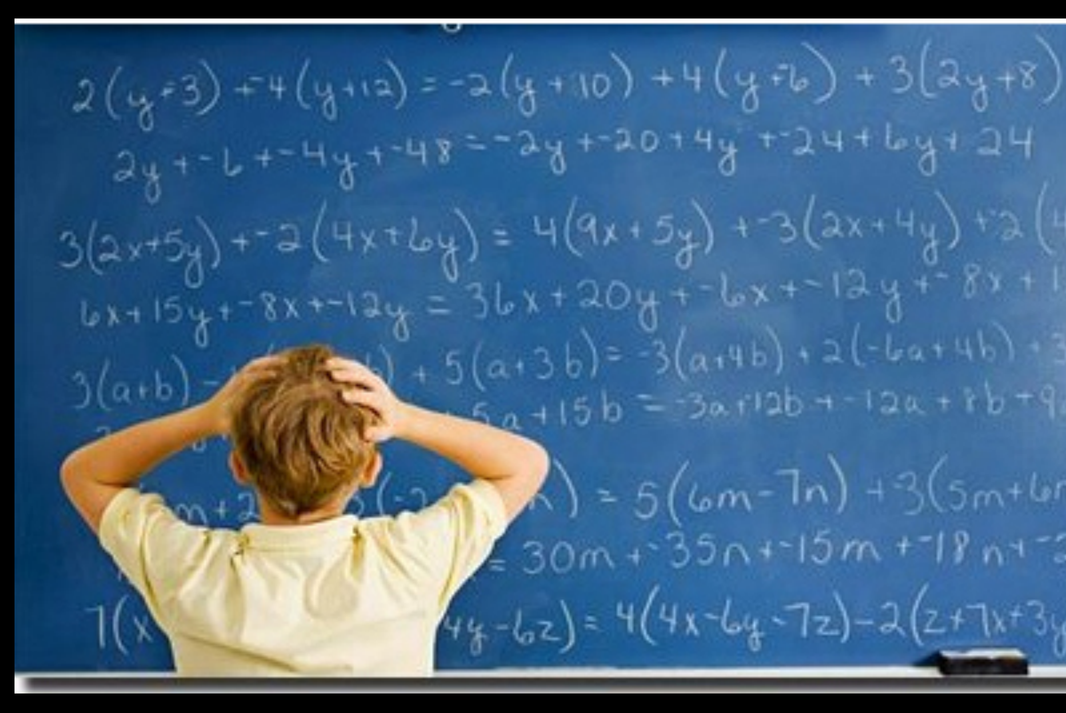
Scaling up CPU/GPU Computing

Dr. Andrew Anderson
Software Tools Group
SCSS @ TCD

How do we measure computing power?



FLOPS

(Floating Point operations/
Second)

What is a FLOP anyway?

- Nobody really agrees
- Hardware manufacturers like big numbers
- Commonly accepted usage:
  - 1 multiply-accumulate (MAC)
  - 1 fused multiply-add (FMA)

$$1.2 \times 3.0 + 1.0$$

Climate Change is Real

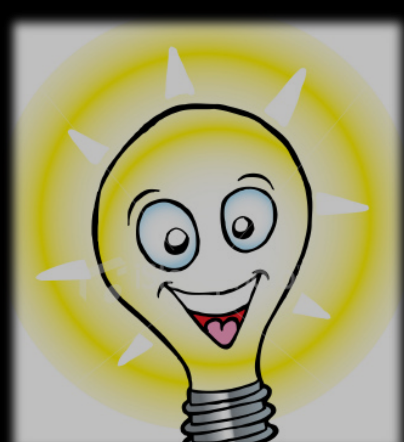
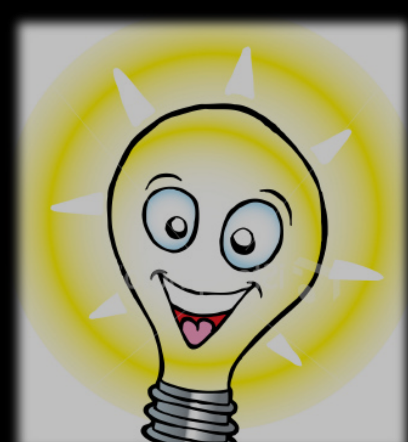
- **TOP500 is the list of the 500 most powerful computers**
- **GREEN500 is the list of the 500 most energy-efficient**
- **Inefficient compute resources require lots of cooling**
  - **Uses lots of additional power**
  - **Have to build data centres differently**

ENERGY STAR

FLOPS/Watt

How efficient is a human?

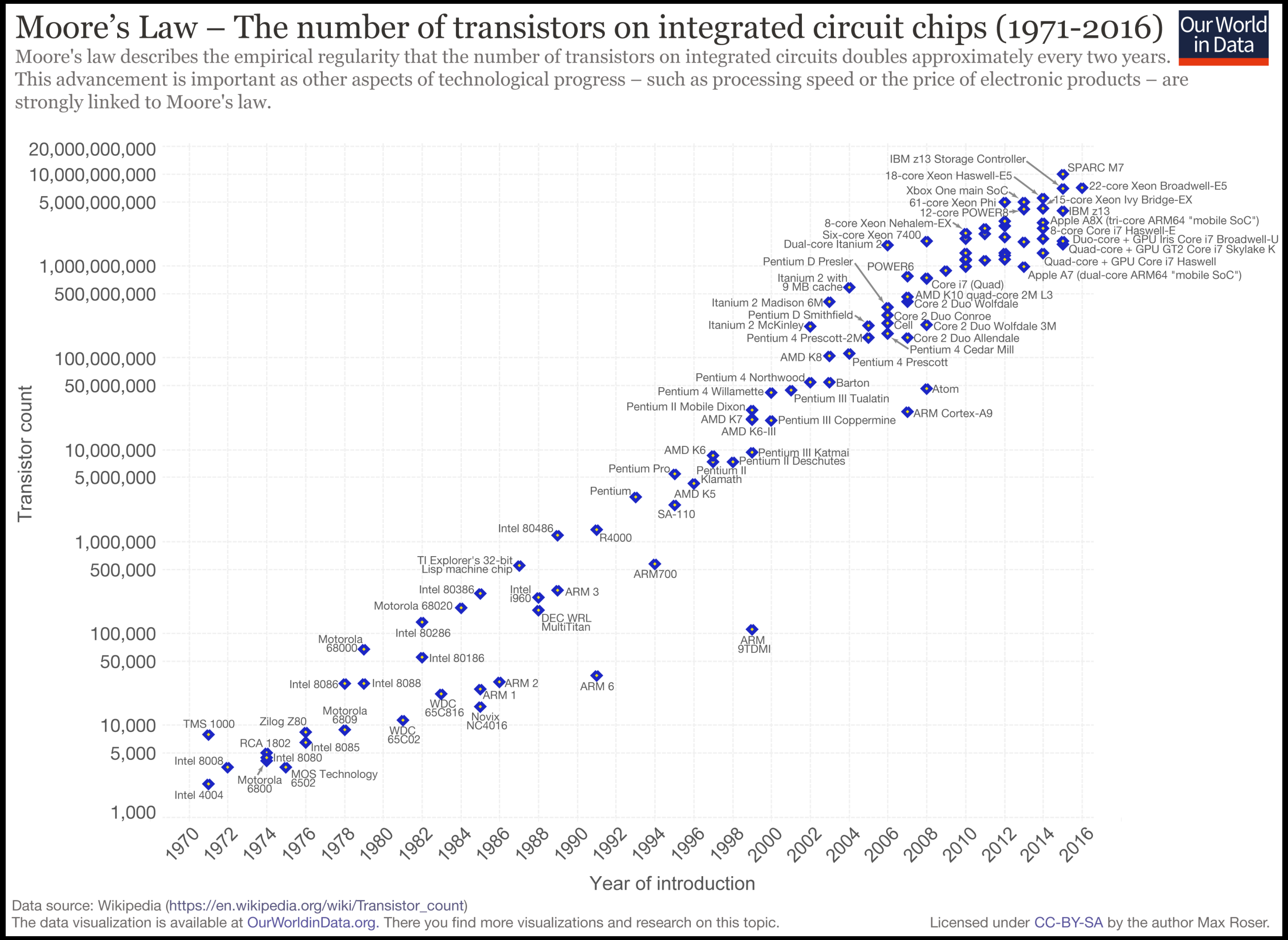
- ~4.2kj = 1kcal
- ~2200kcal = 1 day = ~9.24 x 10^6 j
- ~8.64 x 10^4 s = 1 day
- 9.24 x 10^6 / 8.64 x 10^4 = ~107W
- About two light bulbs or a tenth of a kettle


Class Activity

1.2 x 3.0 + 1.0	1.2 x 3.0 + 2.0	1.2 x 3.0 + 3.0	1.2 x 3.0 + 4.0
1.2 x 4.0 + 2.0	1.2 x 4.0 + 3.0	1.2 x 4.0 + 4.0	1.2 x 4.0 + 5.0
1.2 x 5.0 + 3.0	1.2 x 5.0 + 4.0	1.2 x 5.0 + 5.0	1.2 x 5.0 + 6.0
1.2 x 6.0 + 4.0	1.2 x 6.0 + 5.0	1.2 x 6.0 + 6.0	1.2 x 6.0 + 7.0
1.2 x 7.0 + 5.0	1.2 x 7.0 + 6.0	1.2 x 7.0 + 7.0	1.2 x 7.0 + 8.0

Moore's Law:
Transistor count doubles every 2 years



But Moore's Law runs into the Quantum Wall at 5nm!

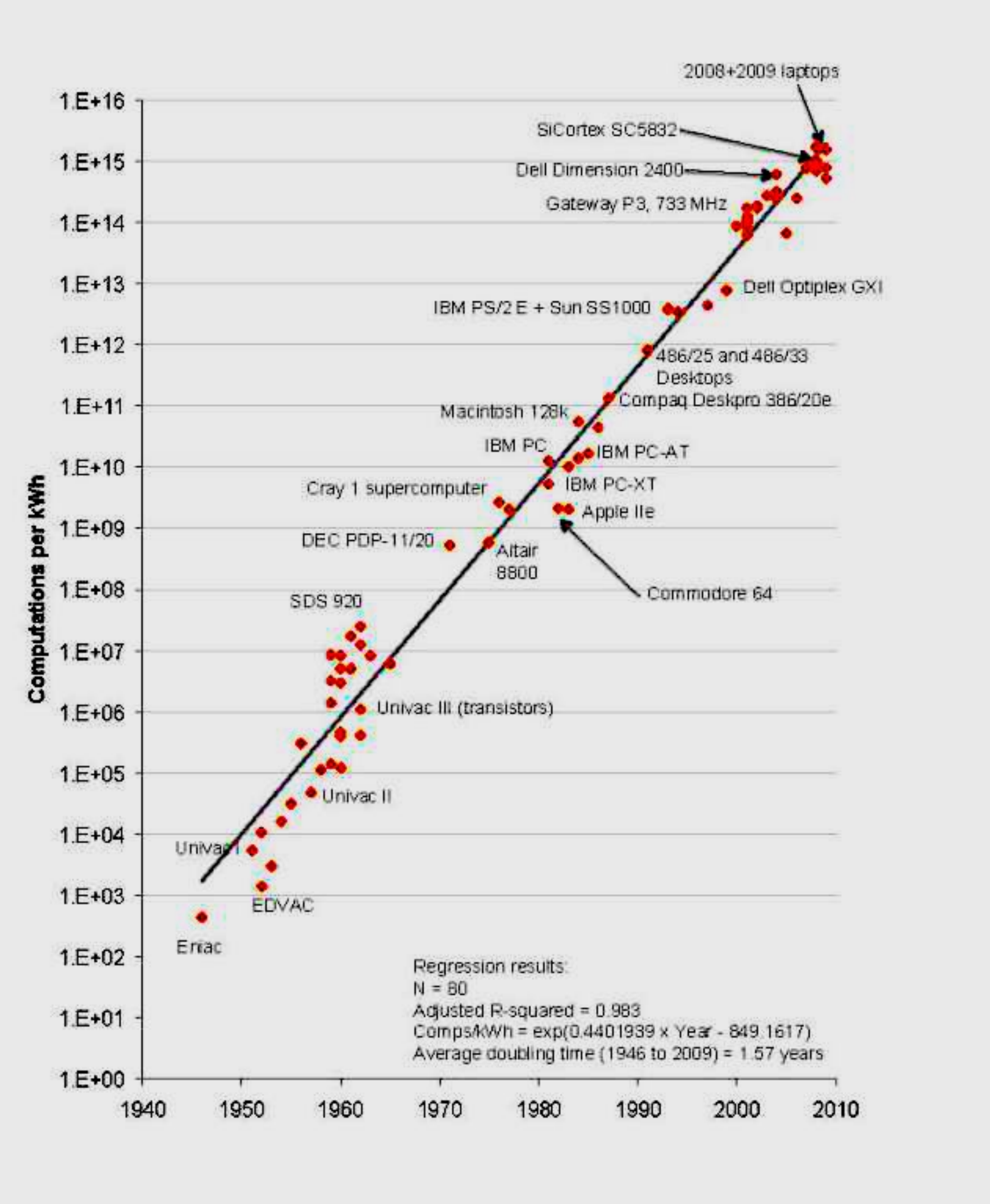
Dennard Scaling:
Transistor power is proportional to area

- **Smaller transistors = less power**
- **But we can ramp up the clock speed**
- **So we get more transistors for equivalent power**

WIKIPEDIA:

Since around 2005–2007 Dennard scaling appears to have broken down. As of 2016, transistor counts in integrated circuits are still growing, but the resulting improvements in performance are more gradual than the speed-ups resulting from significant frequency increases.[2][5] The primary reason cited for the breakdown is that at small sizes, **current leakage** poses greater challenges, and also **causes the chip to heat up**, which creates a threat of thermal runaway and therefore further increases energy costs.

Koomey's Law:
Efficiency doubles every 1.57 years



WIKIPEDIA:

"By the second law of thermodynamics and Landauer's principle, irreversible computing cannot continue to be made more energy efficient forever. As of 2011, computers have a computing efficiency of about 0.00001%.[8] Assuming that the energy efficiency of computing will continue to double every 1.57 years, the Landauer bound will be reached in 2048. Thus, after about 2048, Koomey's law can no longer hold."

Physics is in the way!

If we can't have better computers, let's just have

MORE computers

A Really Big Computer: ASPIRE @ NSCC

NSCC
National Supercomputing Centre
SINGAPORE

1 PFLOPS System (**10^15 FLOPS**)

- 1,288 nodes (dual socket, 12 cores/CPU E5-2690v3)
- 128GB DDR4 RAM/node
- 10 large memory nodes (1x6TB, 4x2TB, 5x1TB)

13PB Storage

- GPFS & Lustre File Systems
- I/O bandwidth up to 500GB/s

Accelerator Nodes

- 128 nodes with Tesla K40 GPUs

Infiniband Interconnection

- EDR (100Gbps) full bandwidth within cluster

How many humans-equivalent?

Current GREEN500 Top 5

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	359	**Shoubu system B** - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 , PEZY Computing / Exascaler Inc. Advanced Center for Computing and Communication, RIKEN Japan	794,400	857.6	47	18.404
2	419	**Suiren2** - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 , PEZY Computing / Exascaler Inc. High Energy Accelerator Research Organization /KEK Japan	762,624	798.0	47	16.835
3	385	**Sakura** - ZettaScaler-2.2, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband EDR, PEZY-SC2 , PEZY Computing / Exascaler Inc. PEZY Computing K.K. Japan	794,400	824.7	50	16.657
4	227	**DGX SaturnV Volta** - NVIDIA DGX-1 Volta36, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla V100 , Nvidia NVIDIA Corporation United States	22,440	1,070.0	97	15.113
5	1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,282,544	122,300.0	8,806	13.889

How do we program a computer with >1000 cores?

- Short answer: with great difficulty
- Long answer: we're working on it!

OpenCL

OpenVX

Halide

Cray

IBM

Questions?