

Piotr Luszczek

Energy footprint

for

LINPACK benchmark

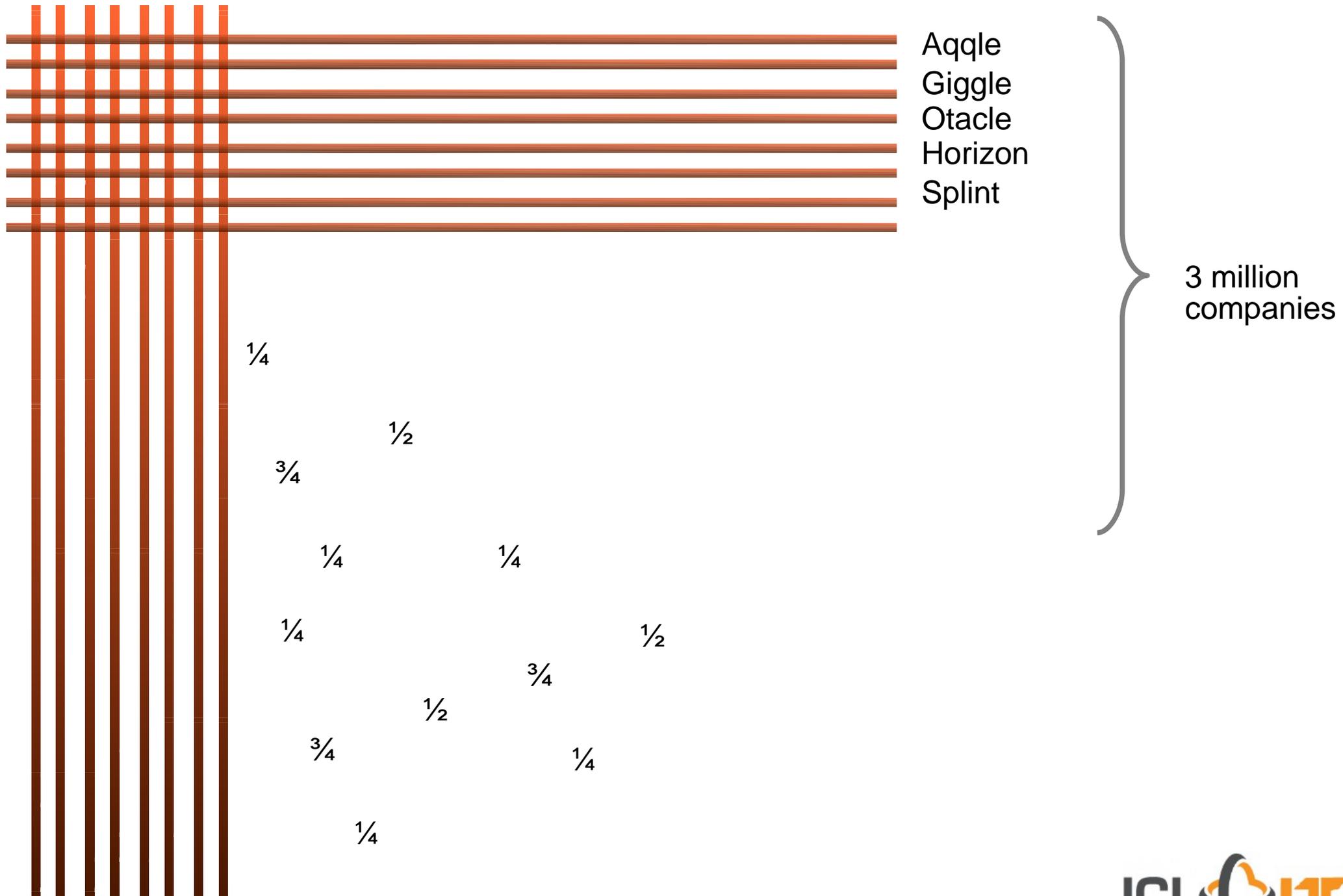
from

supercomputers

to

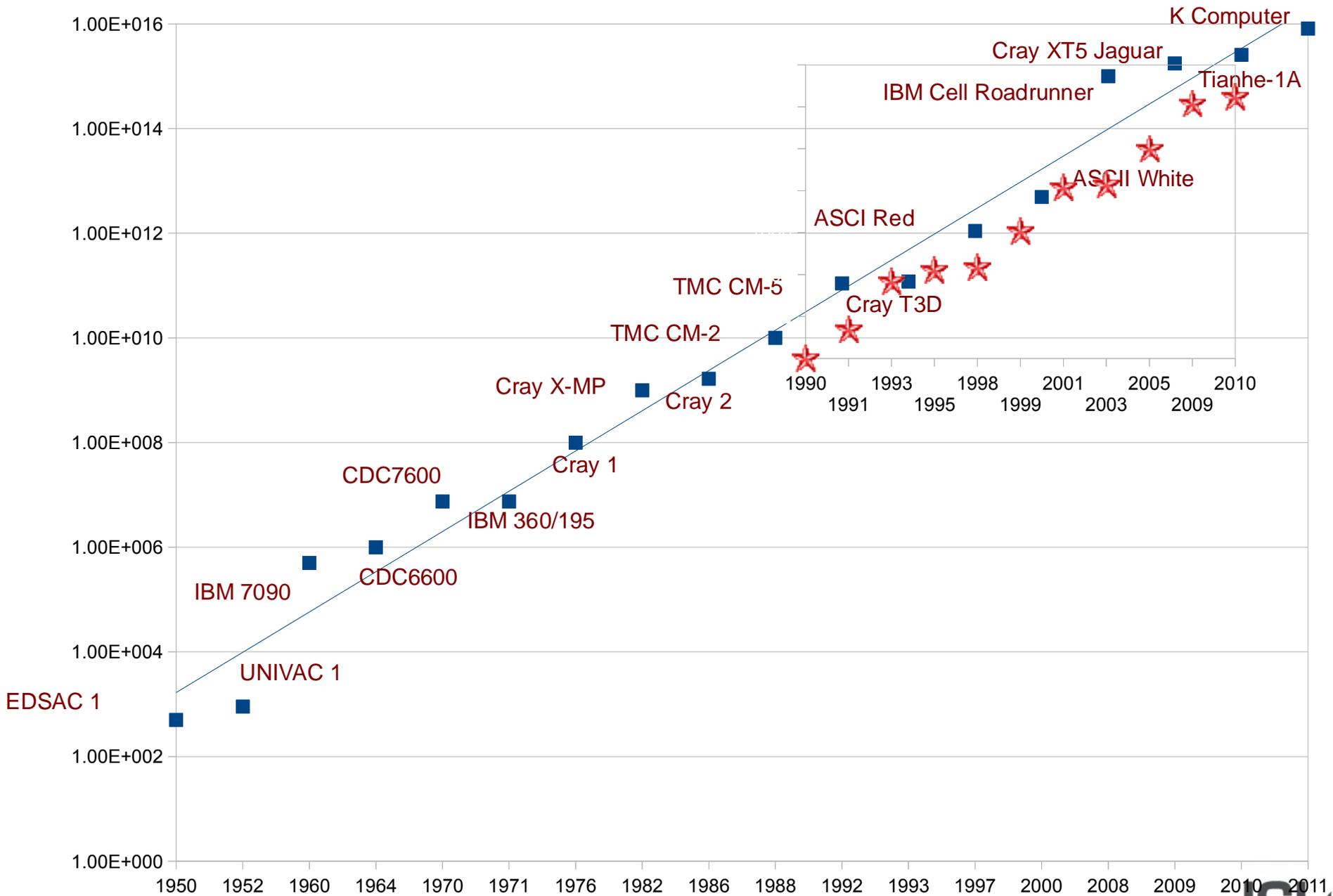
tablet devices

Business Intelligence at IBM



Why LINPACK Benchmark?

5 Decades of the LINPACK: the LINPACK Law



Why talk about power consumption?

x86 at 4GHz cancelled
Multicore era begins:
chip surface hotter than
rocket nozzle.

Power leakage recognized as
the long term challenge for
the CMOS technology.

NVIDIA Fermi
AMD Radeon
2 Gf op/s/W

2008

2004

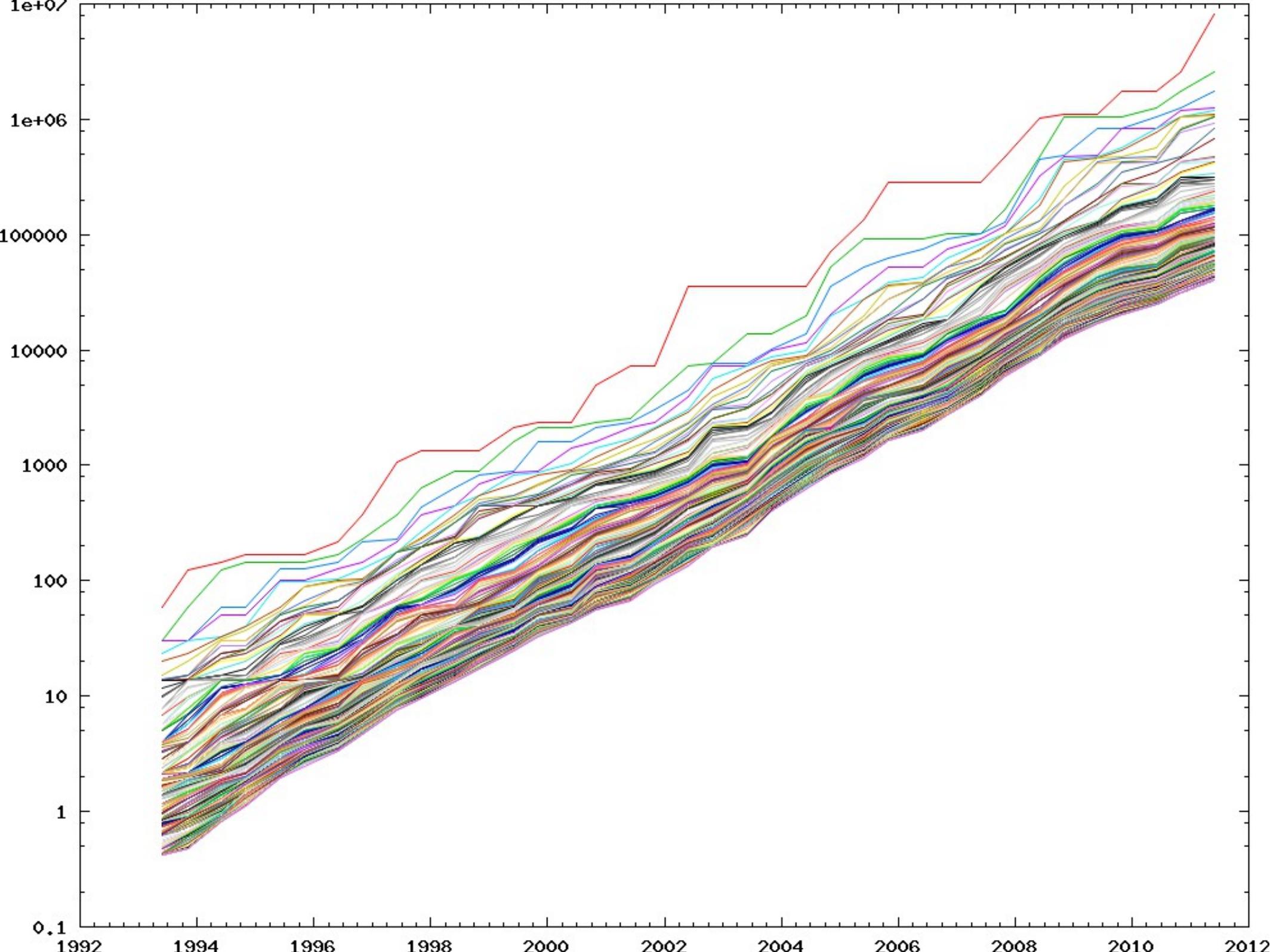
2009

2011

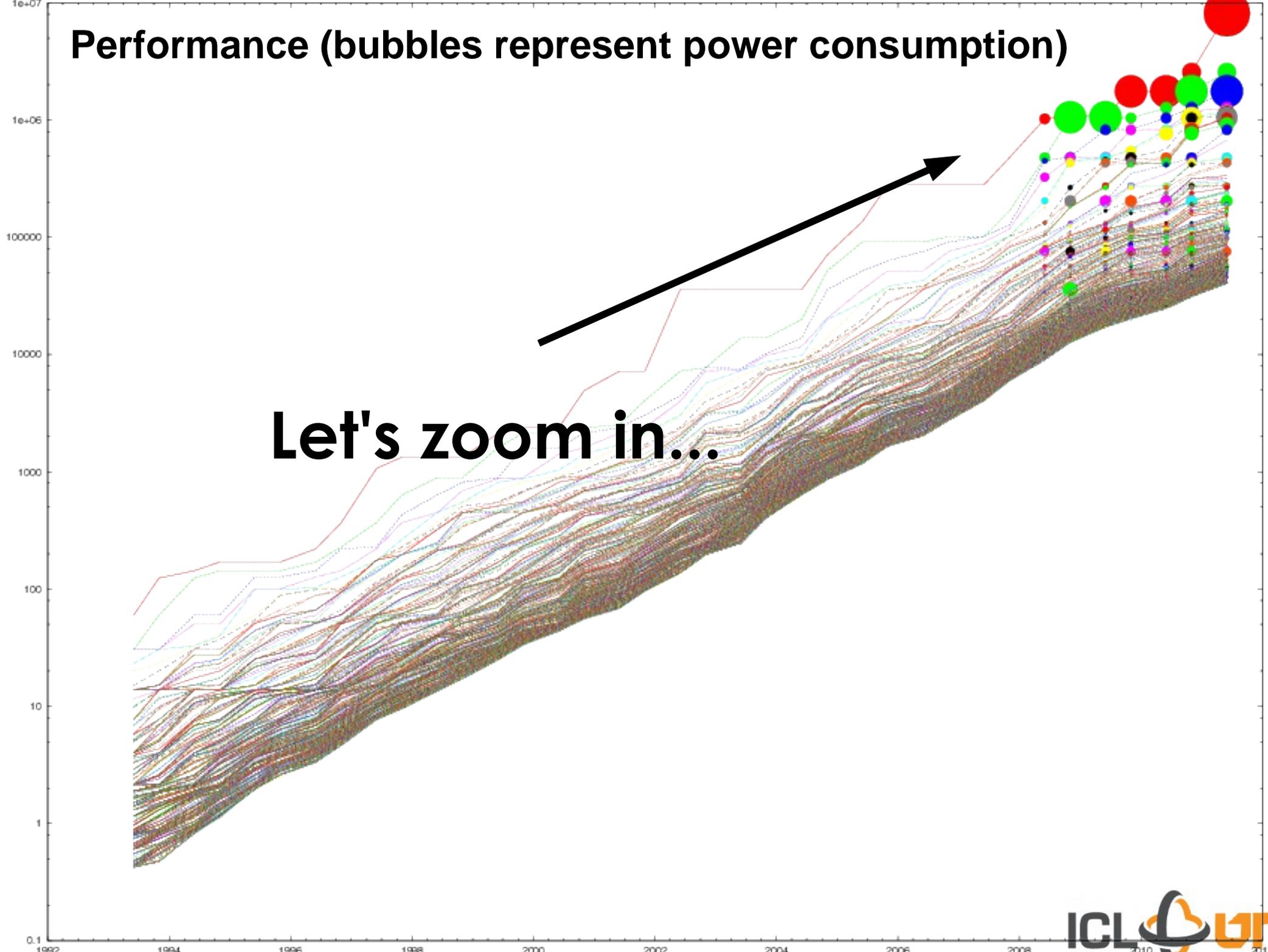
DARPA
Exascale
Study

ARM Cortex A9
iPad 2, Xoom, ...
4 Gf op/s/W

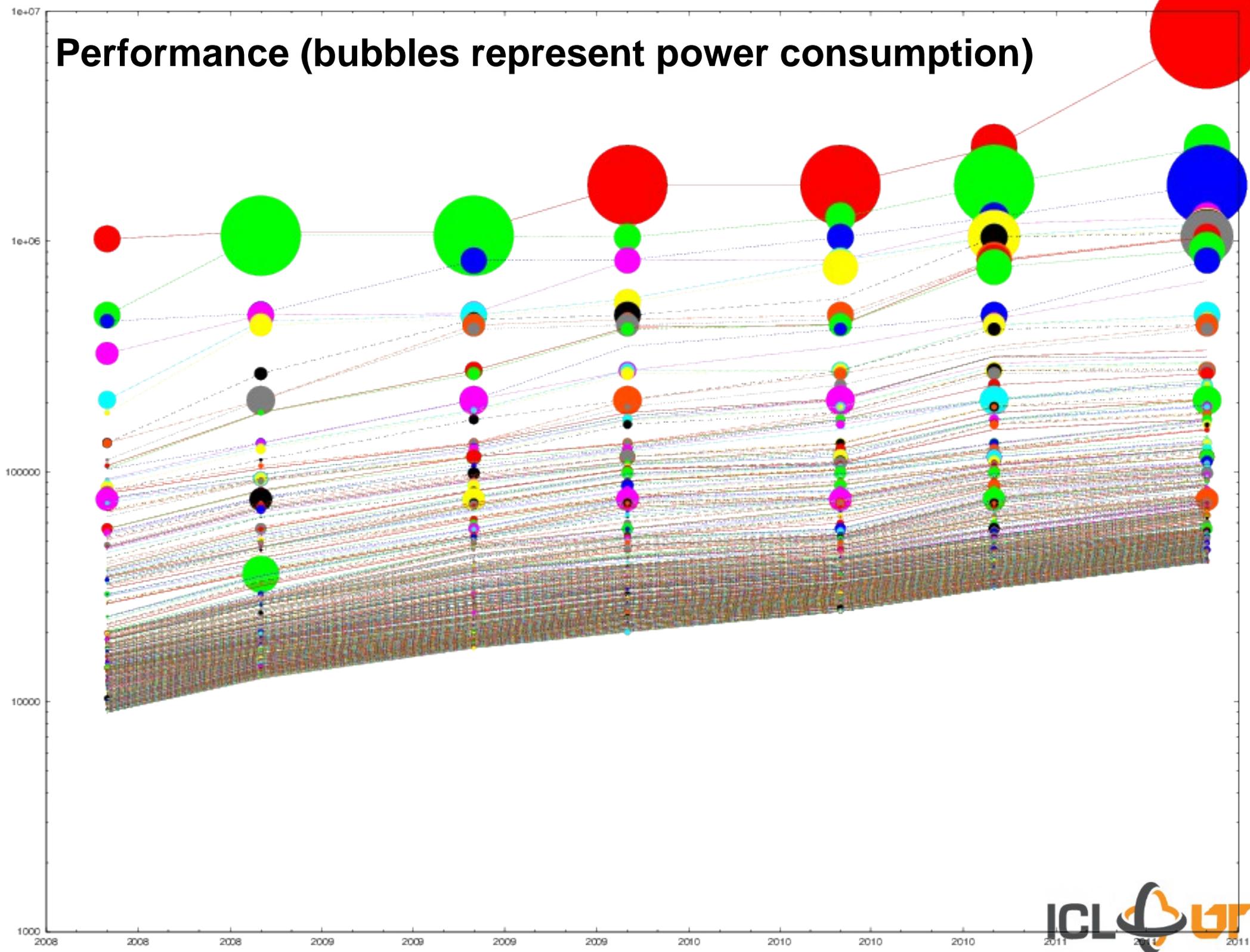
Power in TOP500



Performance (bubbles represent power consumption)

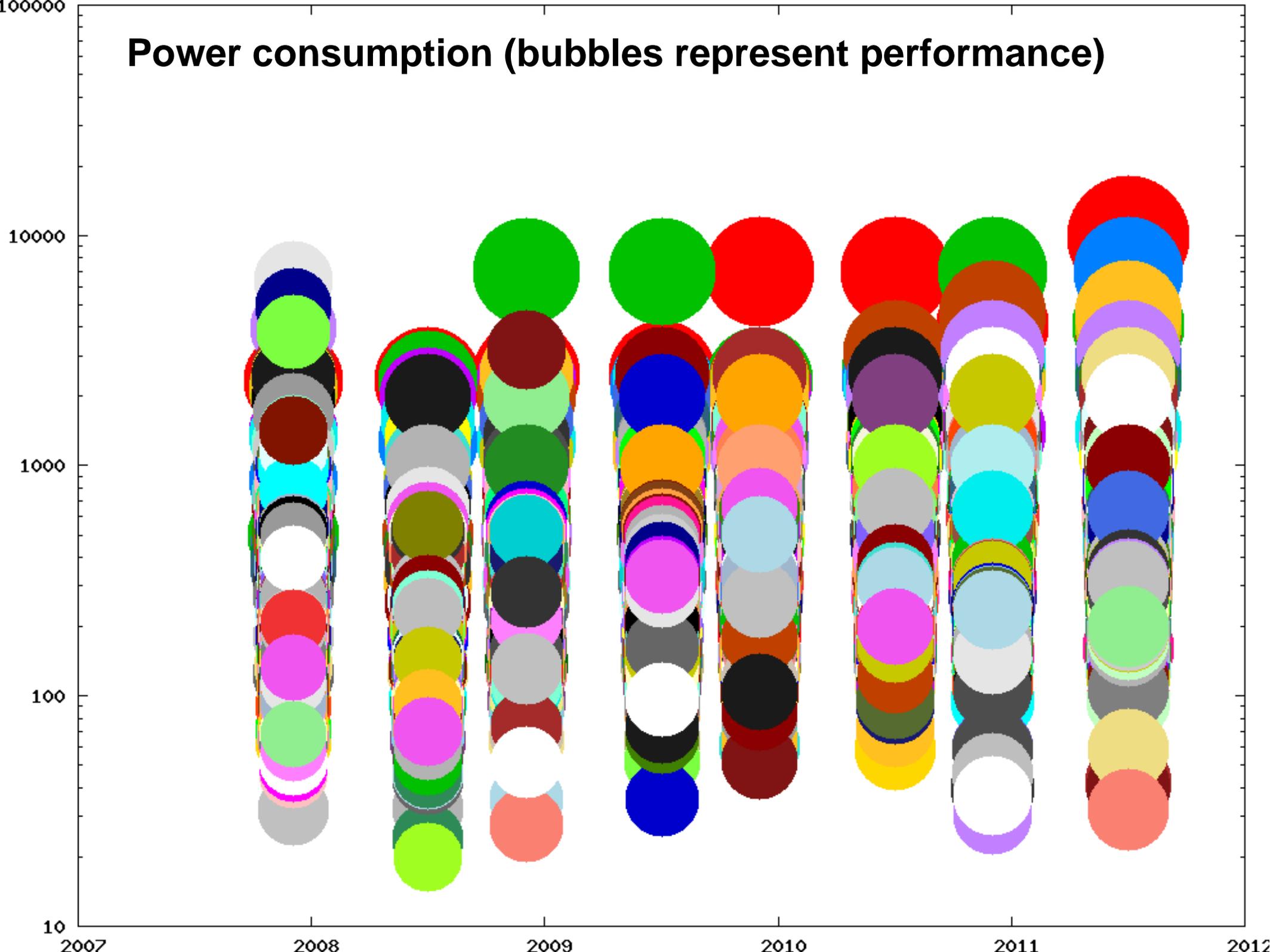


Performance (bubbles represent power consumption)



Can we just talk about power?

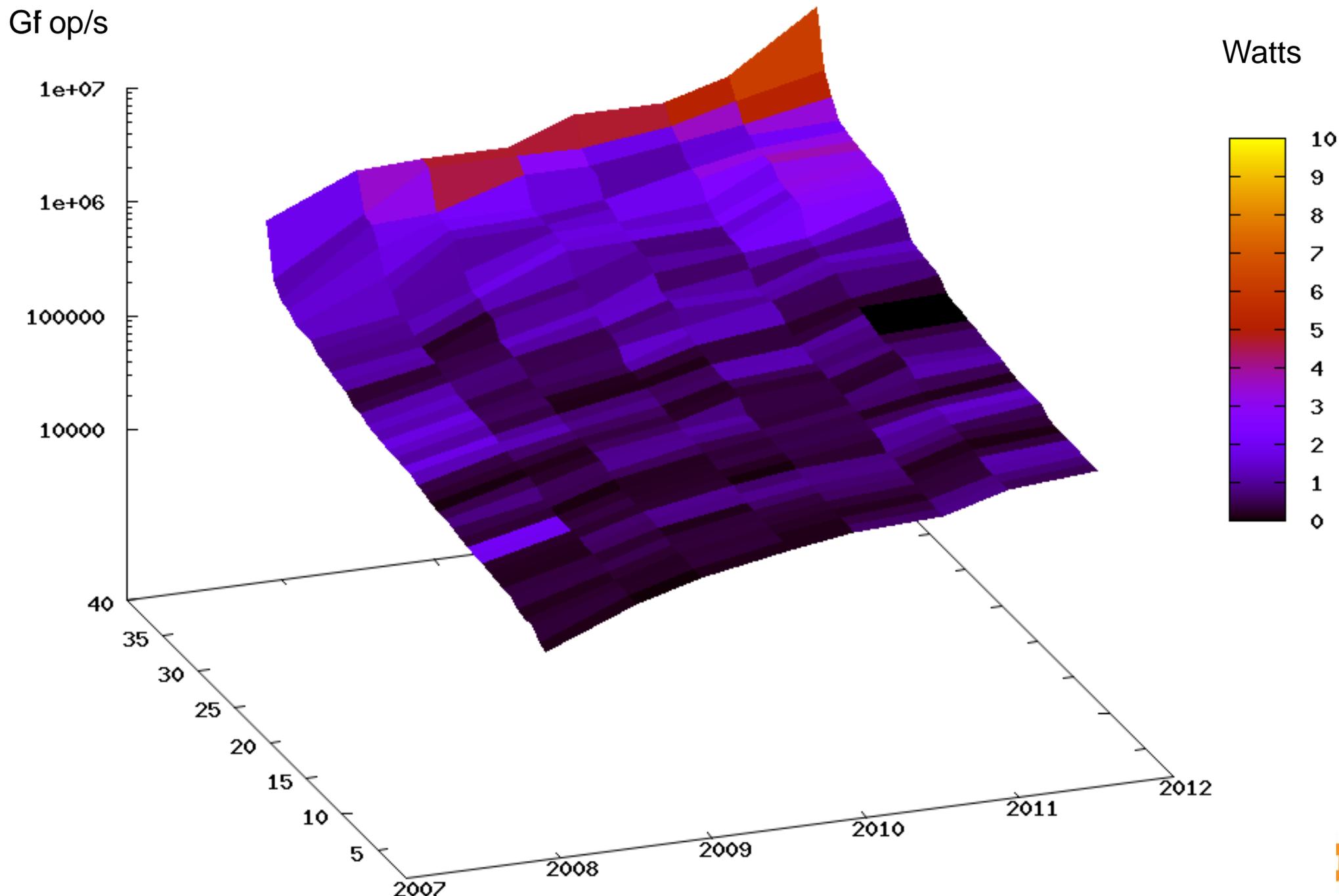
Power consumption (bubbles represent performance)



Let's add another dimension...

TOP 40: Performance and Power

Performance and power consumption of top 40 systems



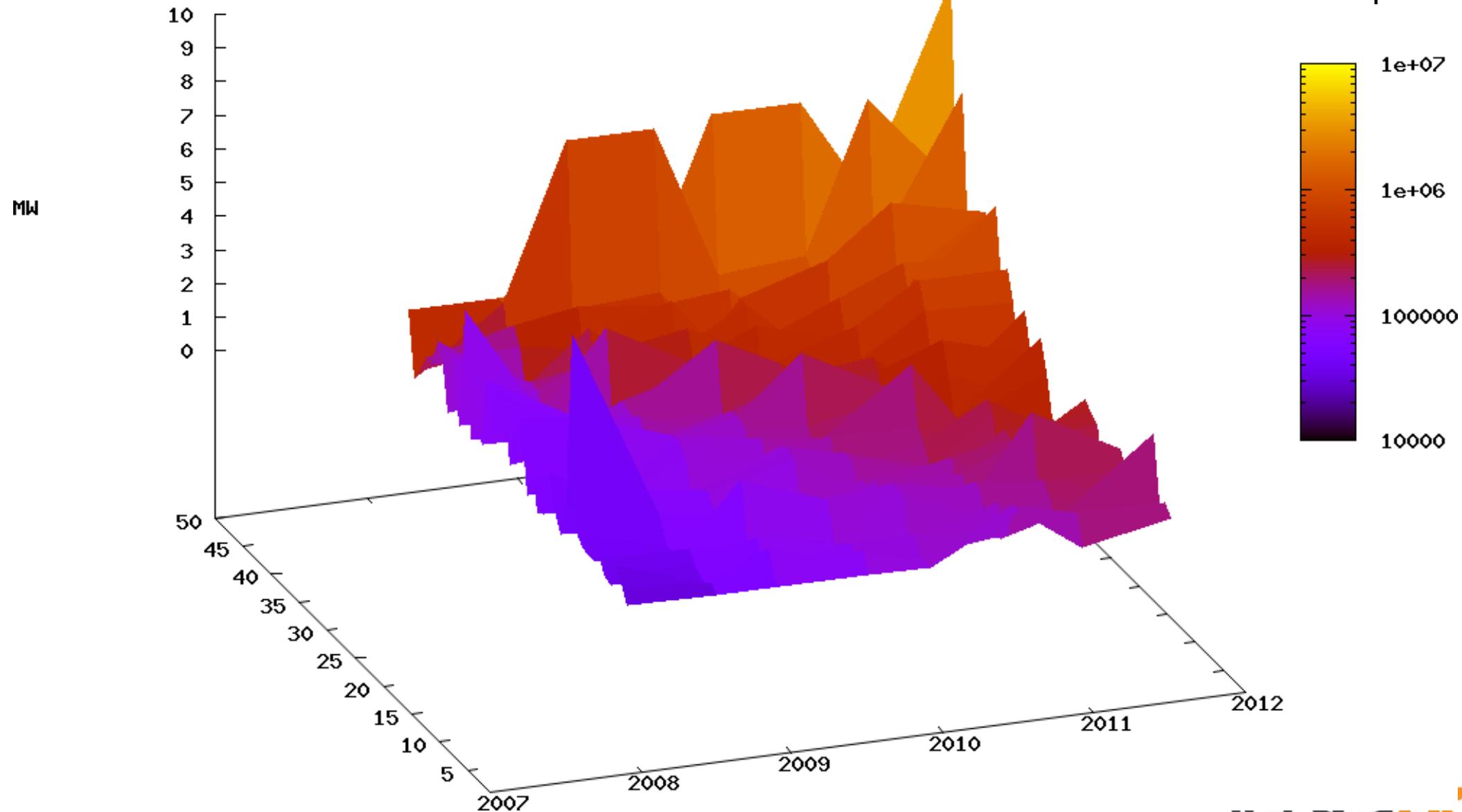
And switch the power and performance

TOP 50: Power and Performance

Power consumption of top 50 systems

Mega Watts

Gf op/s

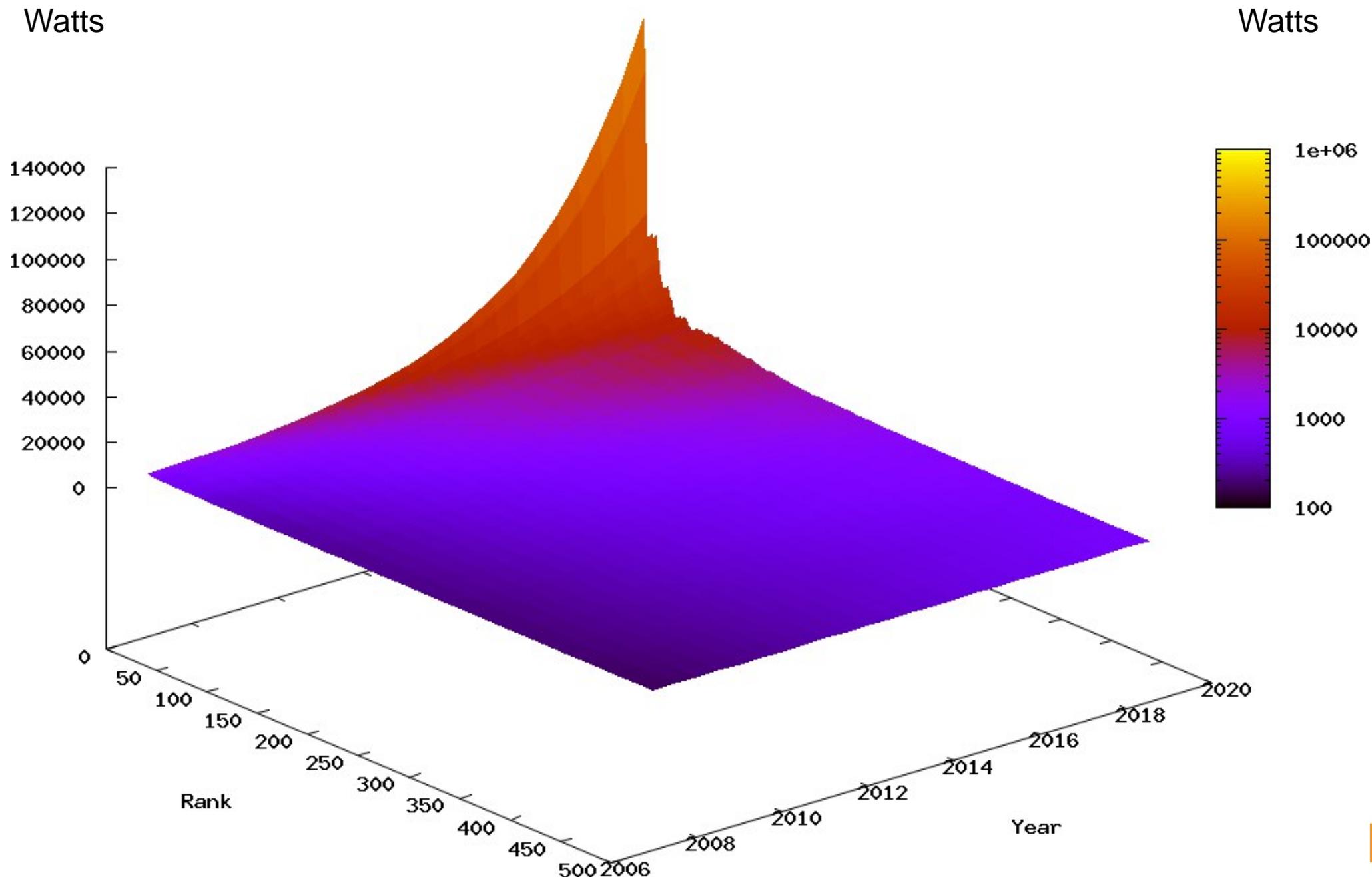


Power Trends

- Power consumption increases dramatically for the top systems
- The best fit for power seems to be exponential function
- What if this trend continues?

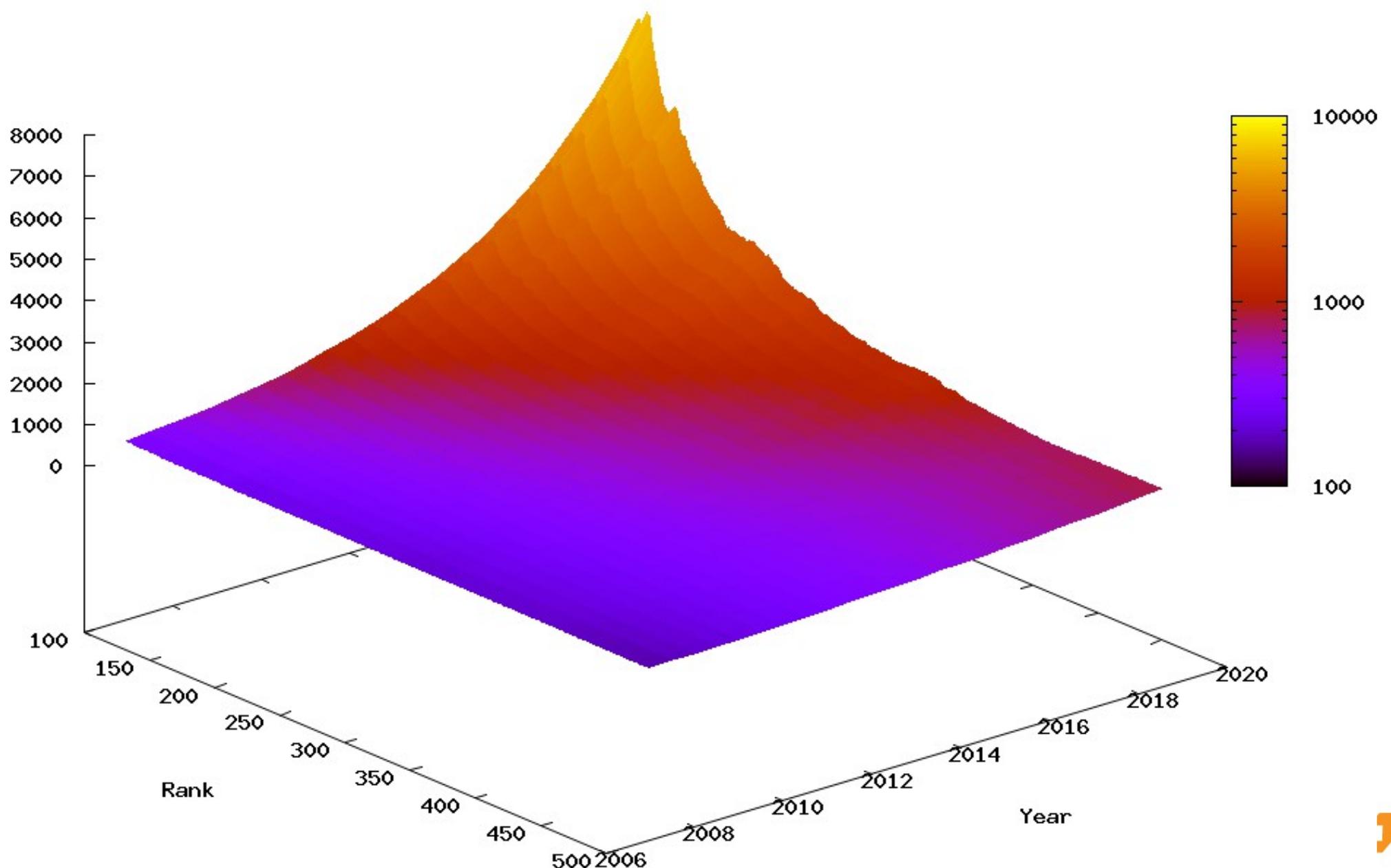
Projected Power Consumption

Projected power consumption of all 500 systems



Projected Power Consumption: Bottom 400

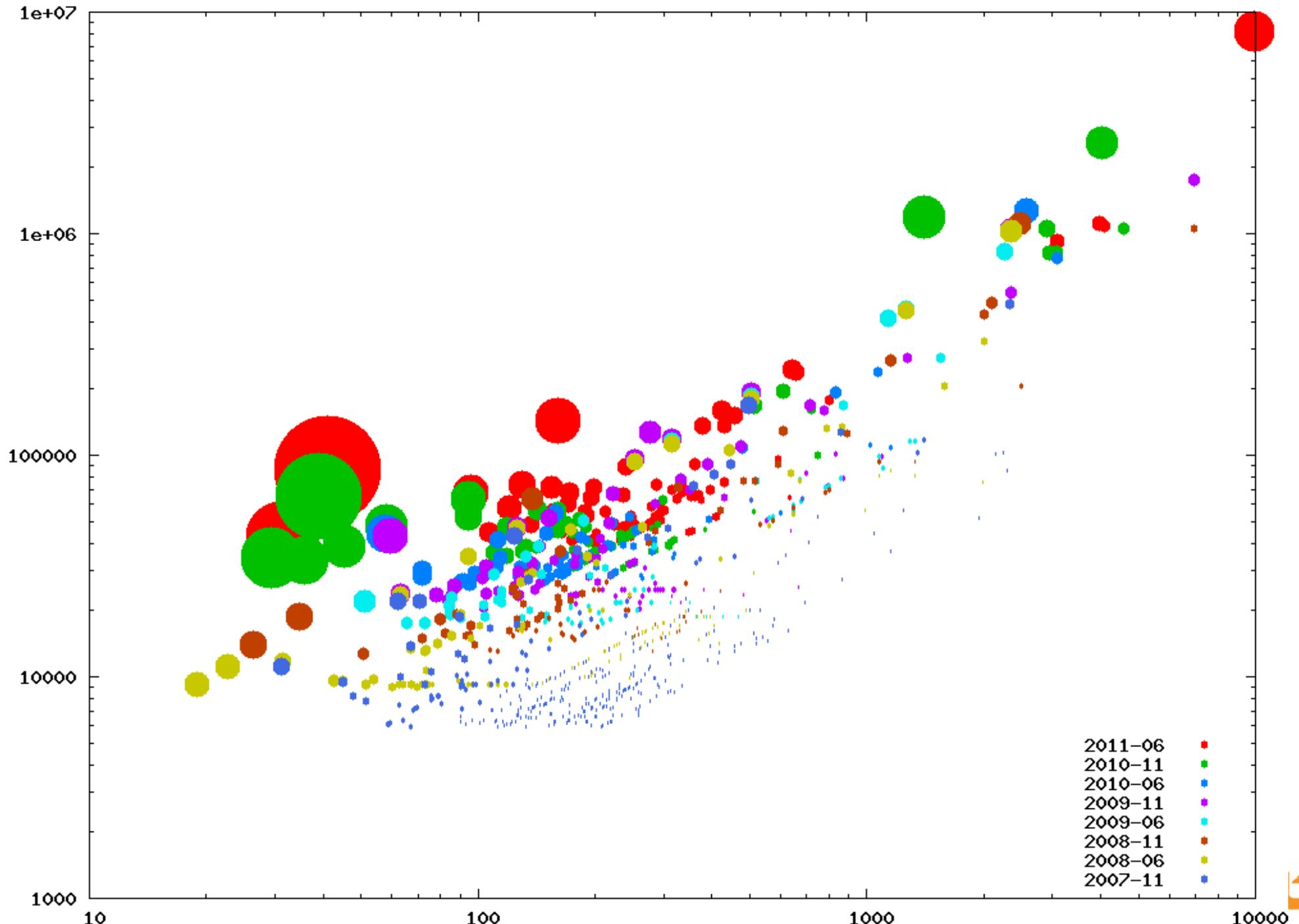
Projected power consumption of bottom 400 systems



A different look at power consumption

Performance vs. Power

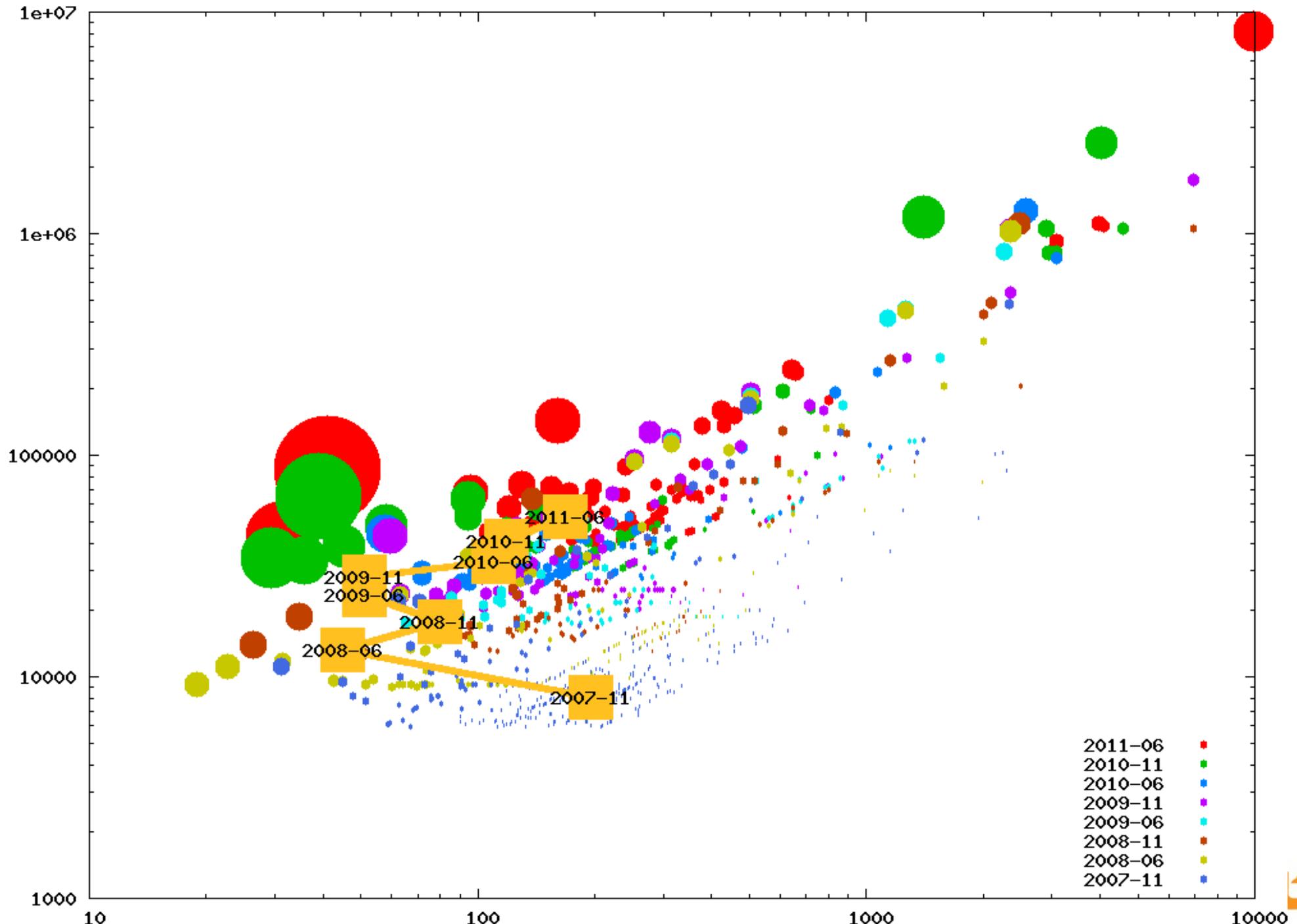
The bigger the bubble the better Gflop/s to Watts ratio



And the trend is...

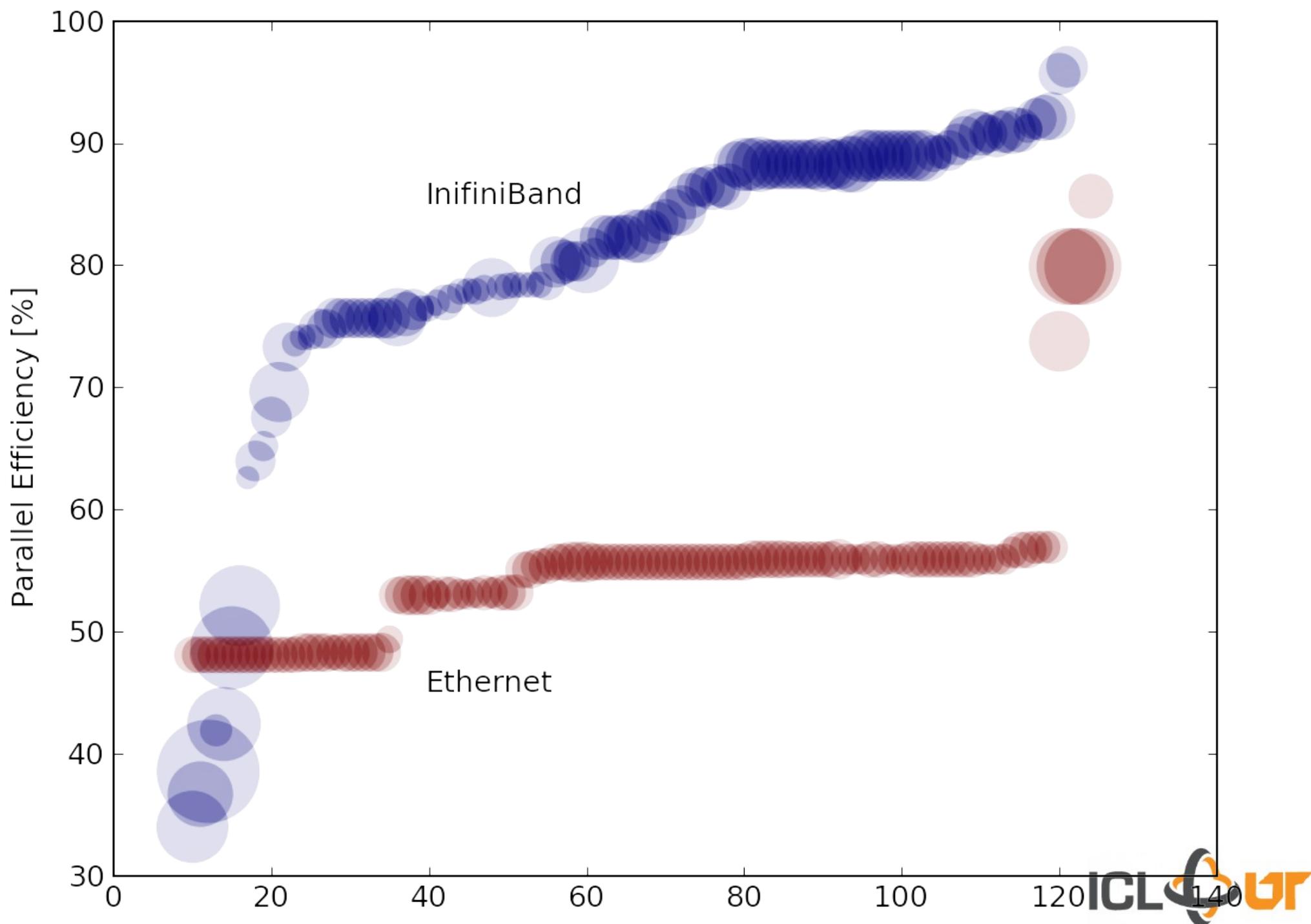
Performance vs. Power and the Median Trend

The bigger the bubble the better Gflop/s to Watts ratio

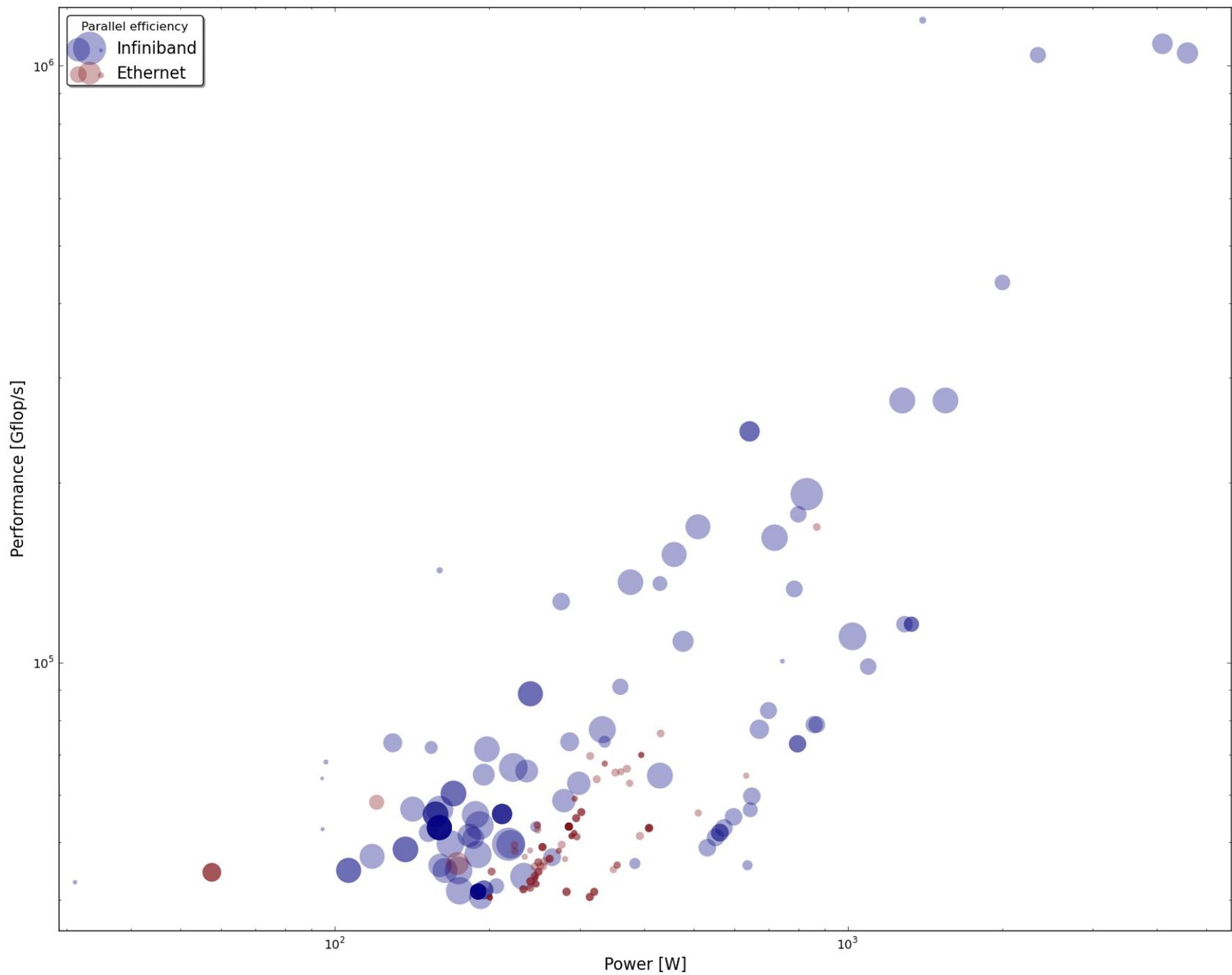


Can power differentiate interconnects?

Interconnect Comparison: Efficiency and Power



Interconnect Comparison: Power and Performance



Reducing effects of increasing power consumption

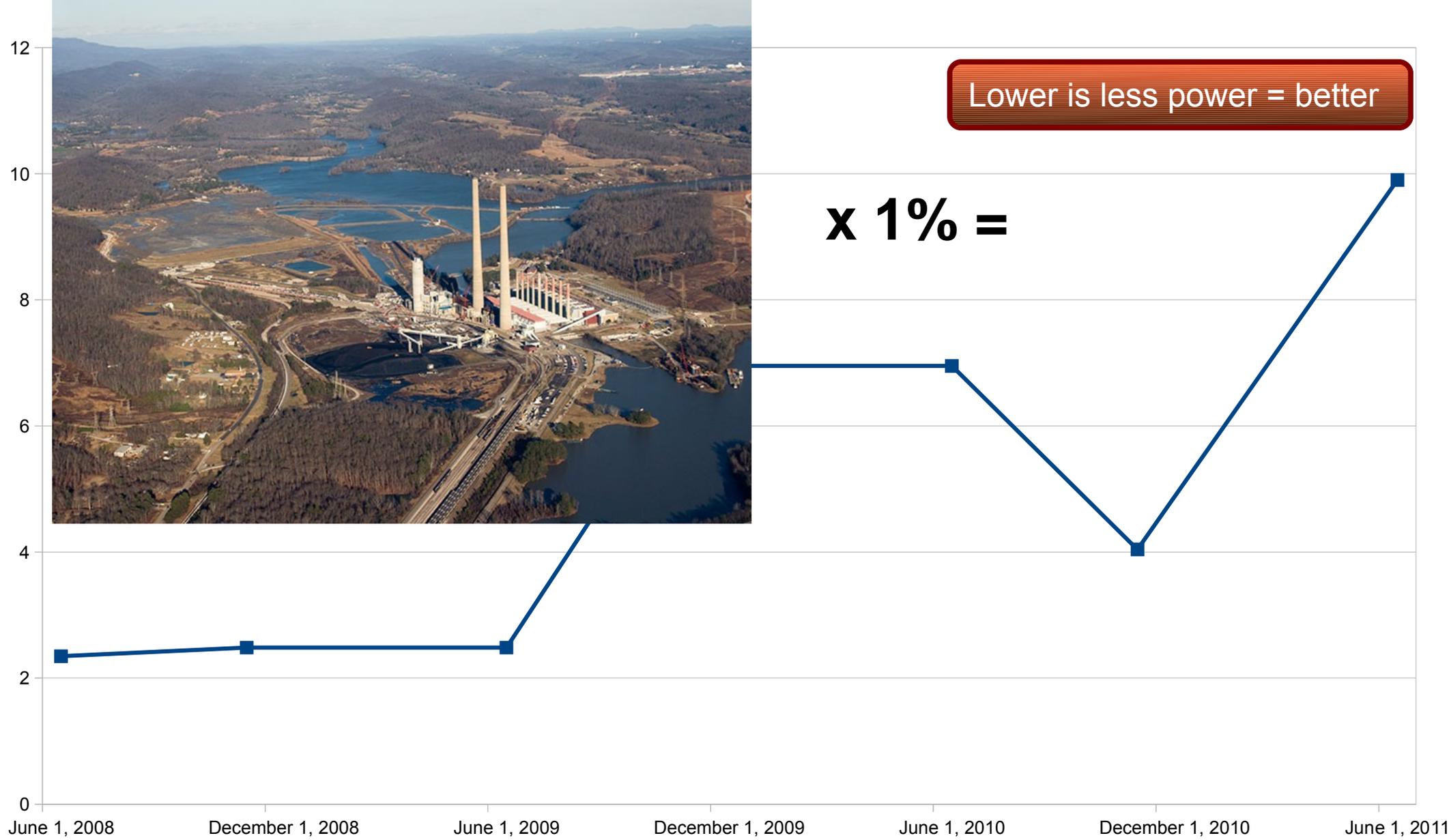
Power consumed

by #1 entry on TOP500 (in MW)



Lower is less power = better

x 1% =



Number of US houses without power for a year

to keep #1 of TOP500 running



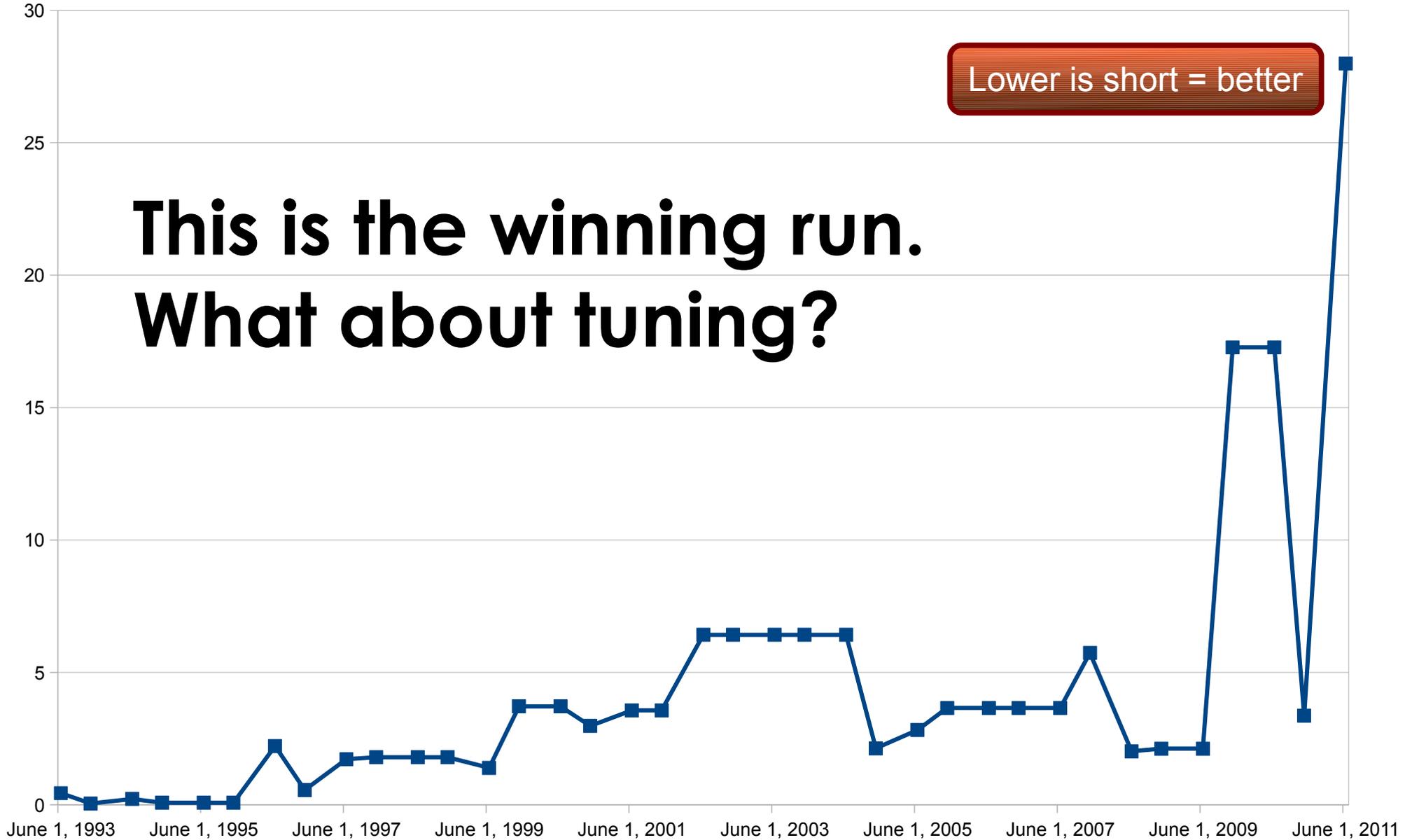
x 55 years =



Time to completion in hours

for #1 entry on TOP500

Hours

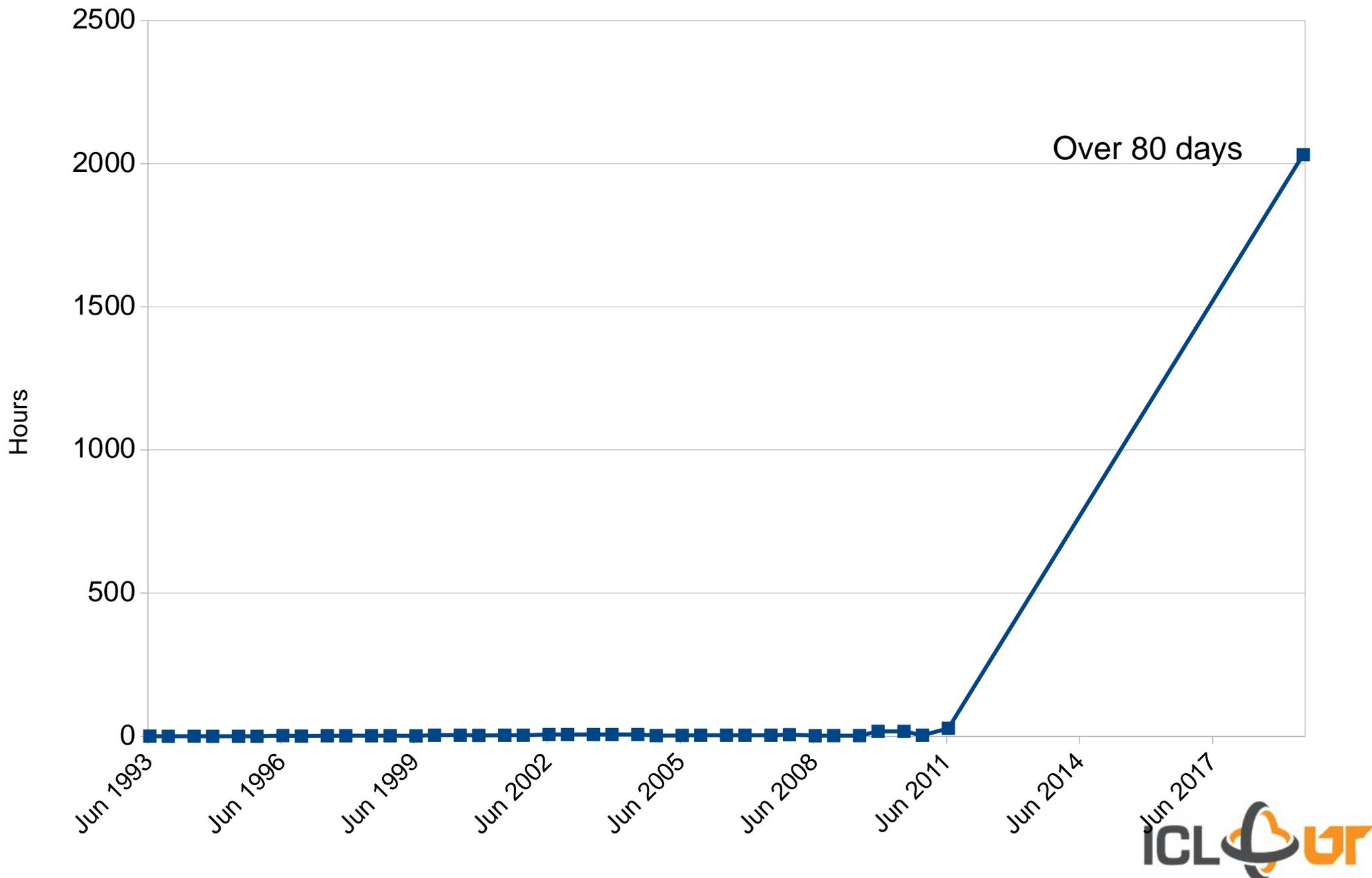


**This is the winning run.
What about tuning?**

Lower is short = better

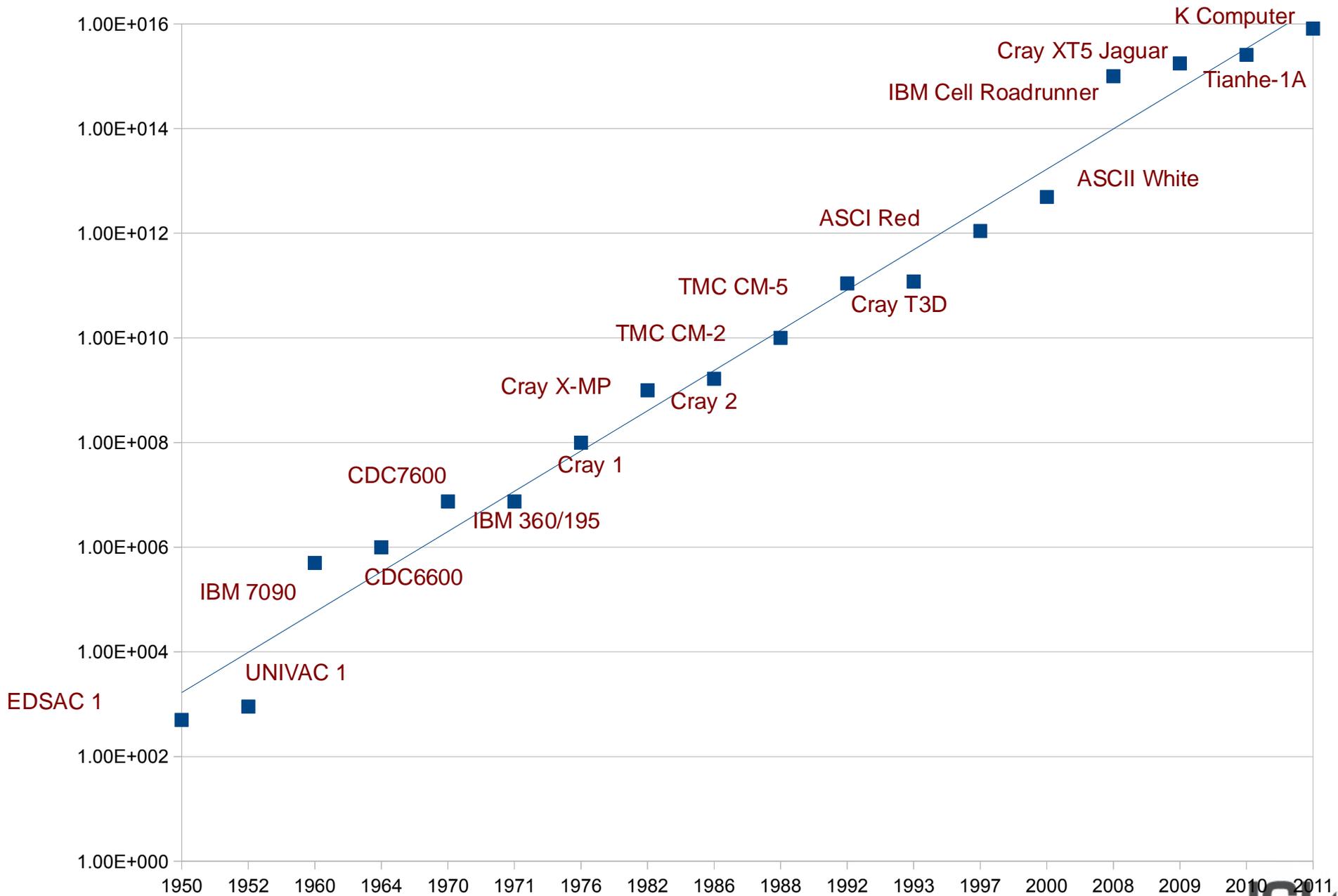
And at Exa-scale...

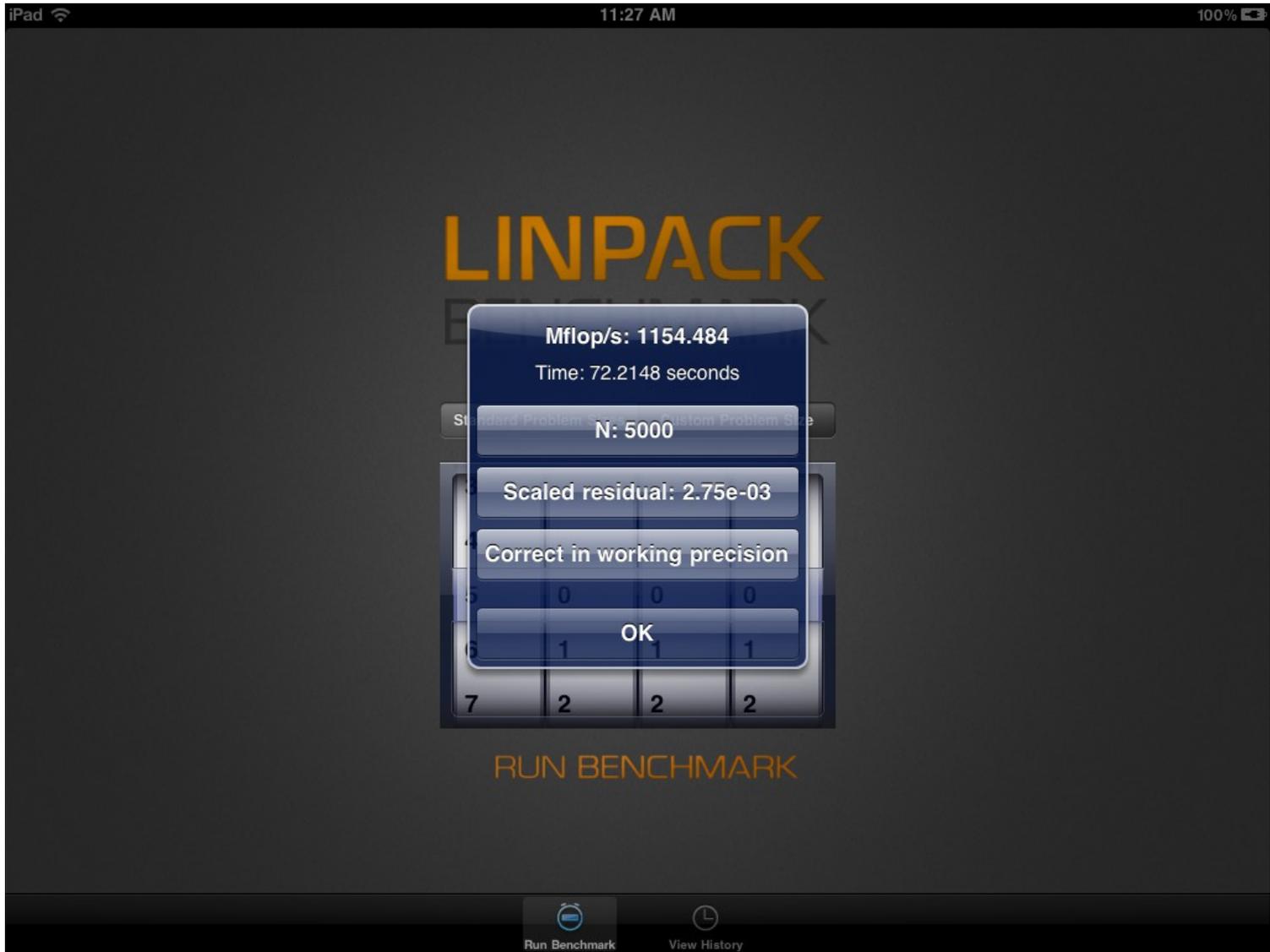
And at the Exascale...



And one more thing..

5 Decades of the LINPACK

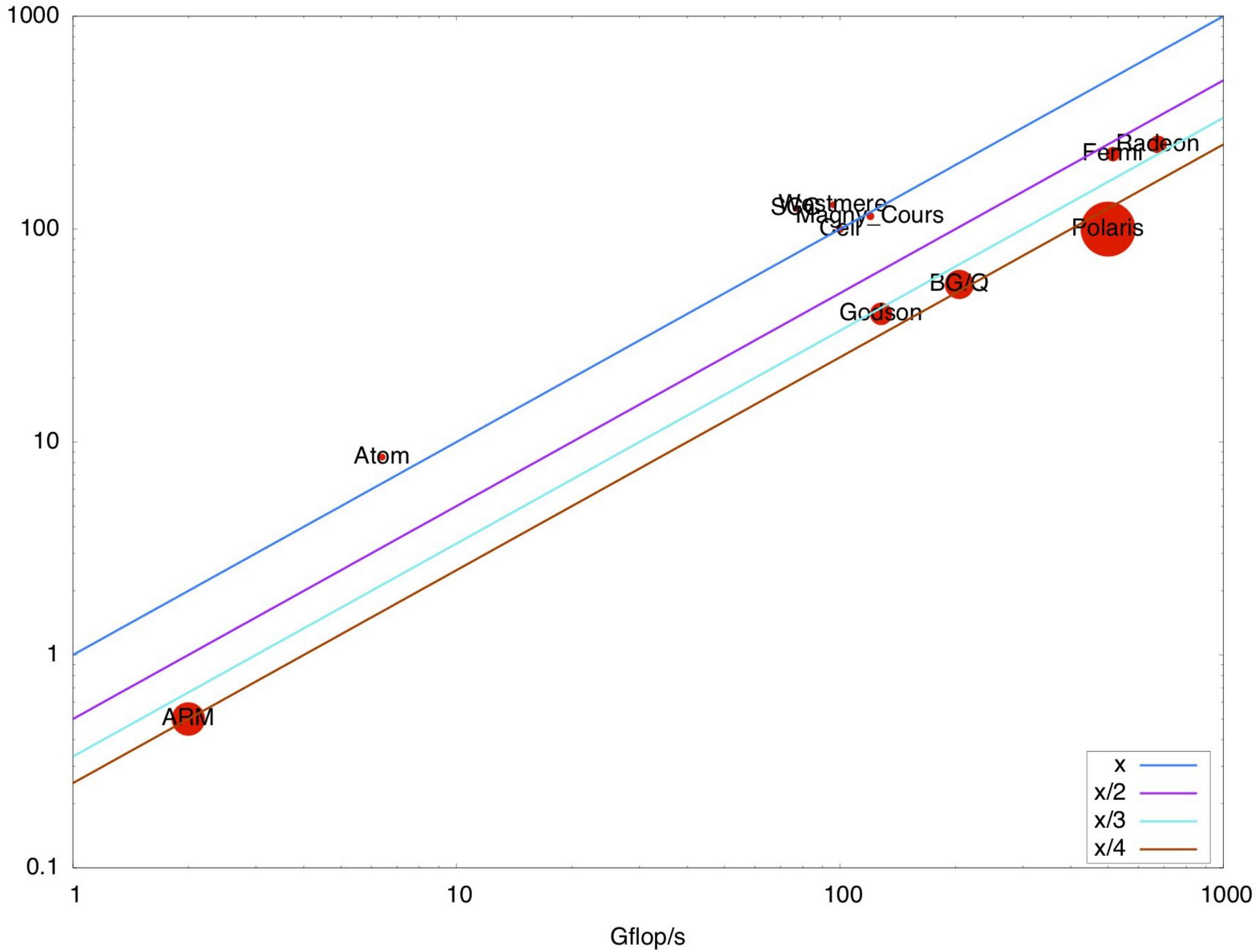




Choose Your Device

	DP Gflop/s	million transistors	TDP Watt	nm	Cores	GHz	Gflop/Watt
Westmere-EX	96	2600	130	32	10	2.4	0.74
Atom	6.4	50	8.5	45	2	1.6	0.75
Magny Cours	120	2000	115	45	12	2.5	1.04
Intel Polaris	500	100	100	65	80	4	5.0
Intel SCC	76.8	1300	125	45	48	1.6	0.6
Tilera TILE64				90	64	0.9	
NVIDIA Fermi	515.2	3000	225	25	16	1.15	2.29
AMD Radeon	675	2640	250	25	32	0.88	2.35
IBM Cell	100.0	200	100	90	8	3.2	1.0
Godson-3B	128	582.6	40	65	8	1	3.20
Blue Gene/Q	204.8	1400	55		16/1+1	1.6	3.72
ARM A9 Cortex	2	26	0.5	60	2	1	4.0

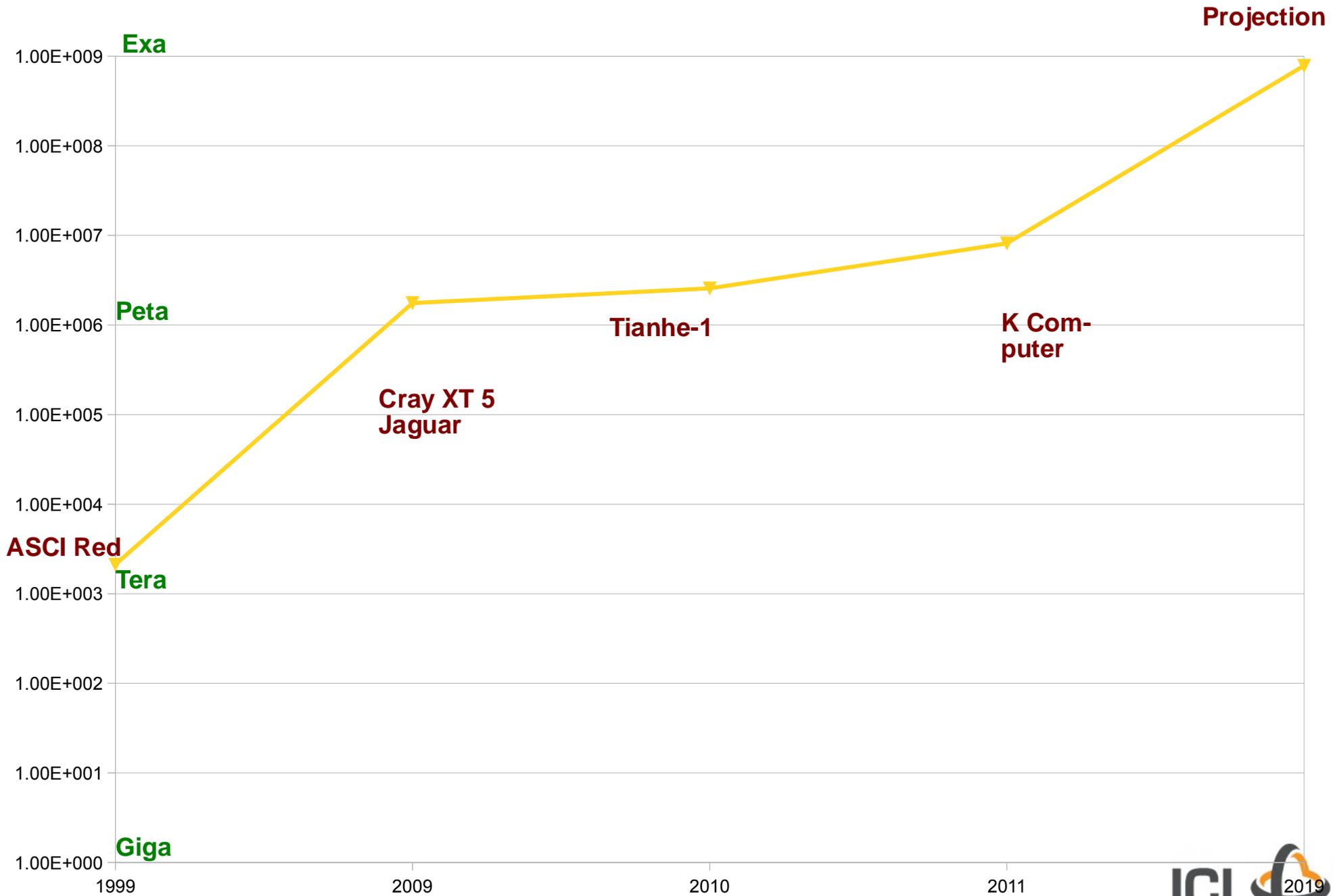
And visually...



Let's try to choose a device

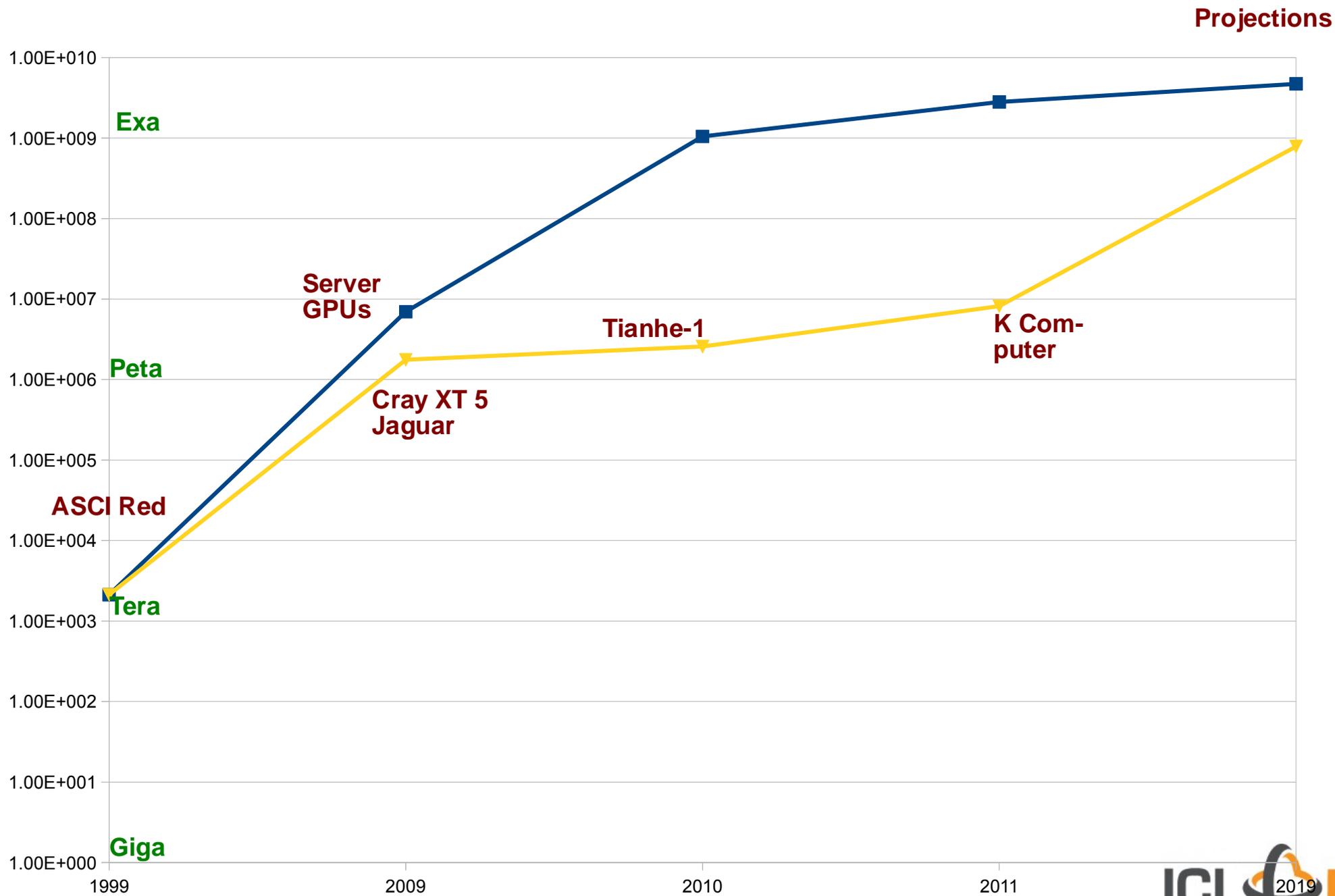
Performance Projections Based on Current Tech.

Keep doing what have been doing



Performance Projections Based on Current Tech.

Just Use GPUs



Performance Projections Based on Current Tech.

Go mobile!

