Piotr Luszczek

Energy and Power Consumption Trends in HPC
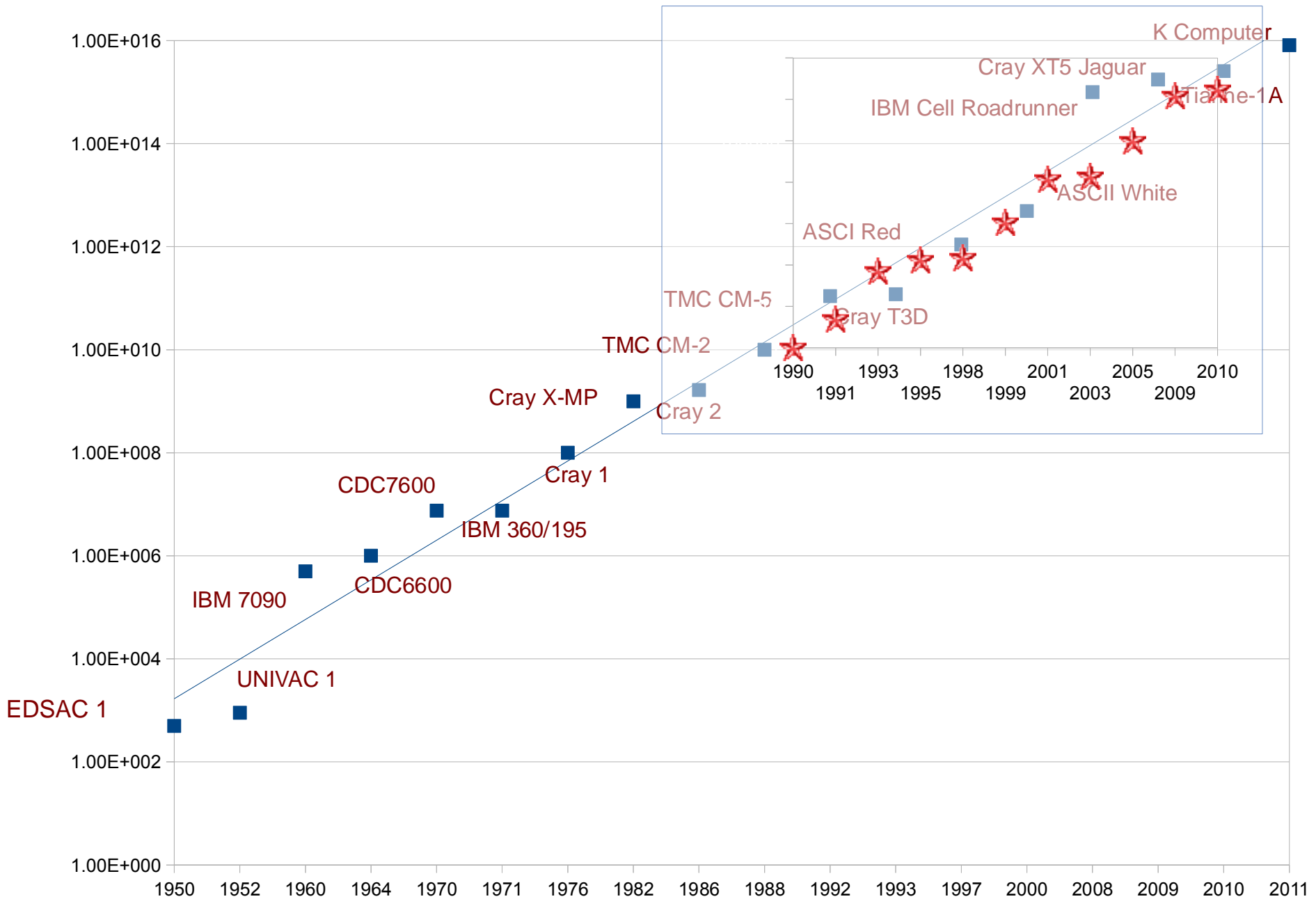
Fun Things to Do with (not so) Big Data

Innovative Computing Laboratory   University of Tennessee

5 Decades of the LINPACK: the LINPACK Law
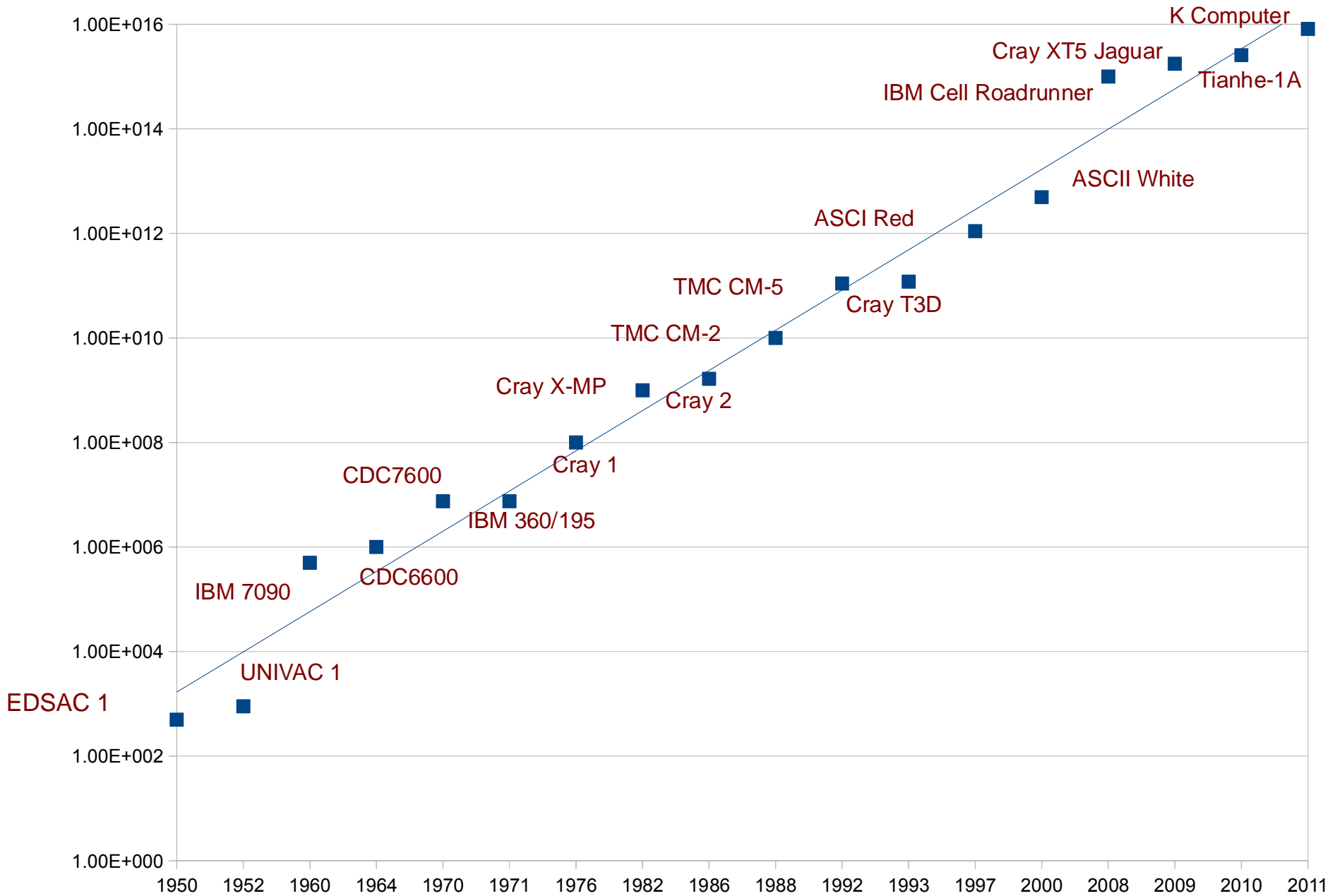
5 Decades of the LINPACK: the LINPACK Law

Exaflop by 2020?

- **Horst Simon's wager**
  - **A €2000 bet that it won't happen.**

- **Time to run**

Systems	2009	2015 +1/-0	2018 +1/-0
System peak	2 Peta	100-300 Peta	1 Exa
Power	6 MW	~15 MW	~20 MW
System memory	0.3 PB	5 PB	64 PB (+)
Node performance	125 GF	0.5 TF or 7 TF	2 TF or 10TF
Node memory BW	25 GB/s	0.2TB/s or 0.5TB/s	0.4TB/s or 1TB/s
Node concurrency	12	O(100)	O(1k) or 10k
Total Node Interconnect BW	3.5 GB/s	100-200 GB/s 10:1 vs memory bandwidth 2:1 alternative	200-400GB/s (1:4 or 1:8 from memory BW)
System size (nodes)	18,700	50,000 or 500,000	O(100,000) or O(1M)
Total concurrency	225,000	O(100,000,000) *O(10)-O(50) to hide latency	O(billion) * O(10) to O(100) for latency hiding
Storage	15 PB	150 PB	500-1000 PB (>10x system memory is min)
IO	0.2 TB	10 TB/s	60 TB/s (how long to drain the machine)
MTTI	days	O(1day)	O(1 day)

Let's Plugin the Numbers

- **Memory: 64 PB** $N \ = \ \sqrt{64\,\mathrm{PB}/8\,\mathrm{bytes}}$

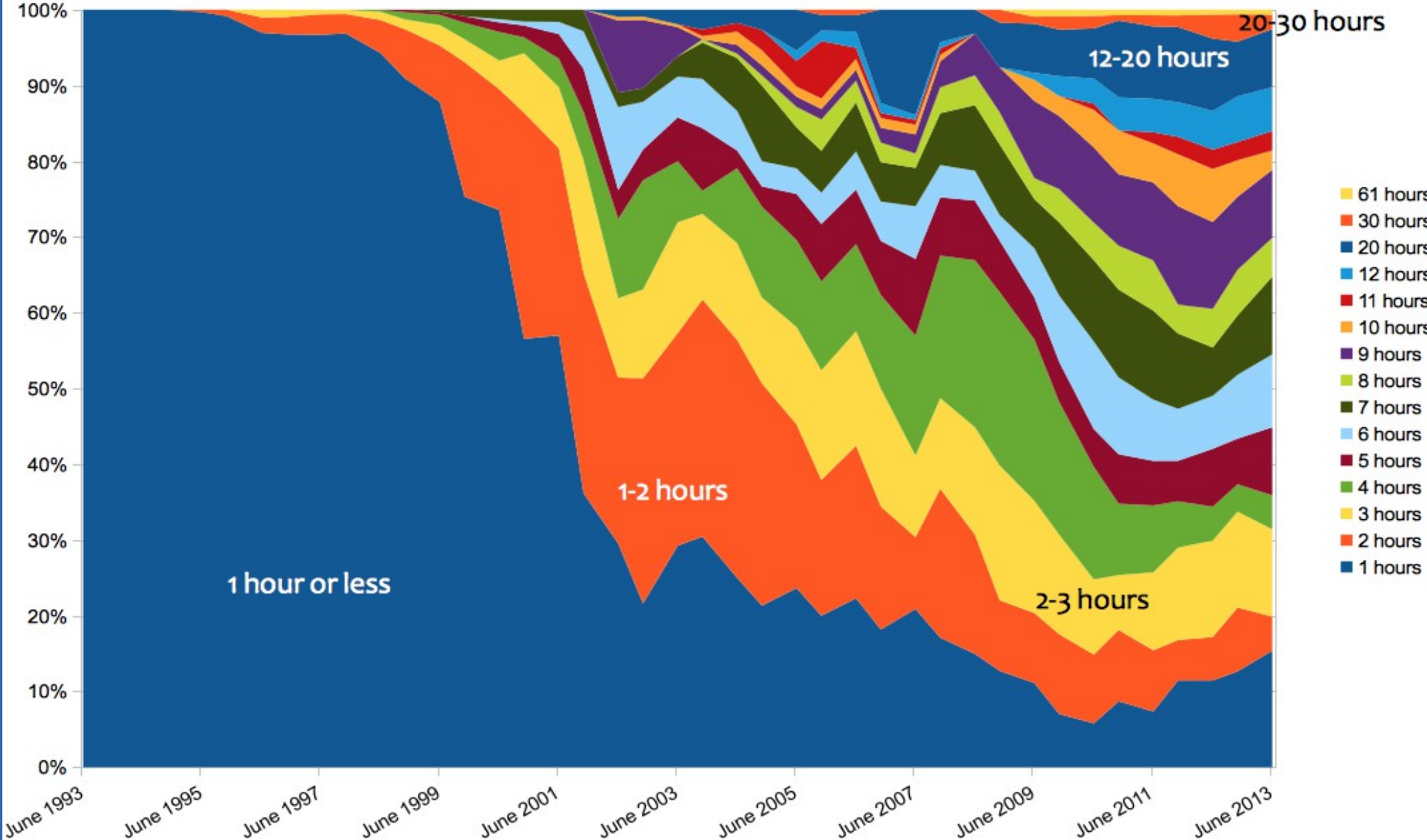 – **Problem size: ~90 million unknowns**

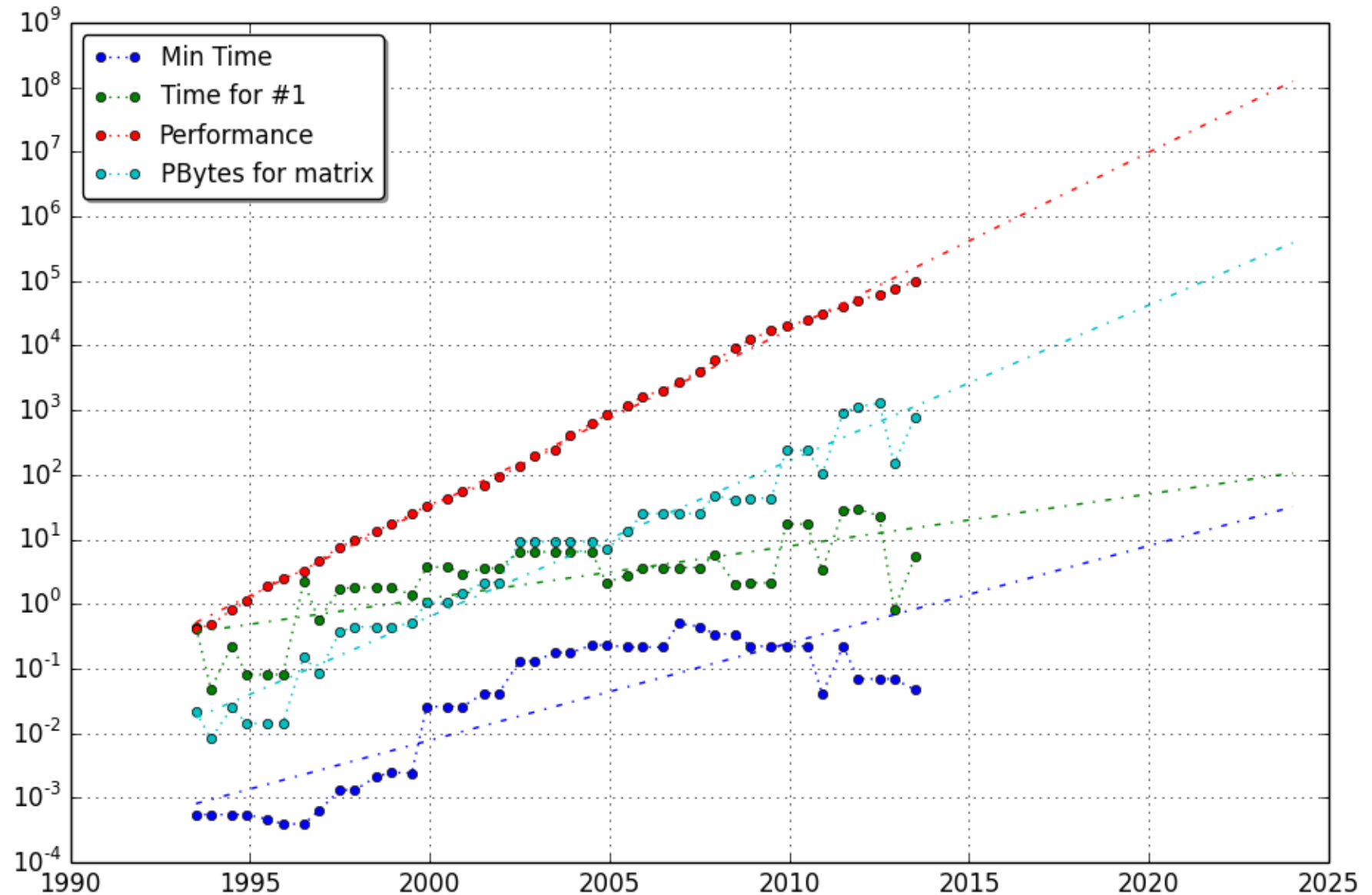- **Performance:**

  – **1 Eflop/s**

- **Time to run:** $\mathrm{time} \ \approx \ \dfrac{\dfrac{2}{3}N^3}{r_{max}}$

  – **5.5 days!**

How Long a TOP500 Run Takes
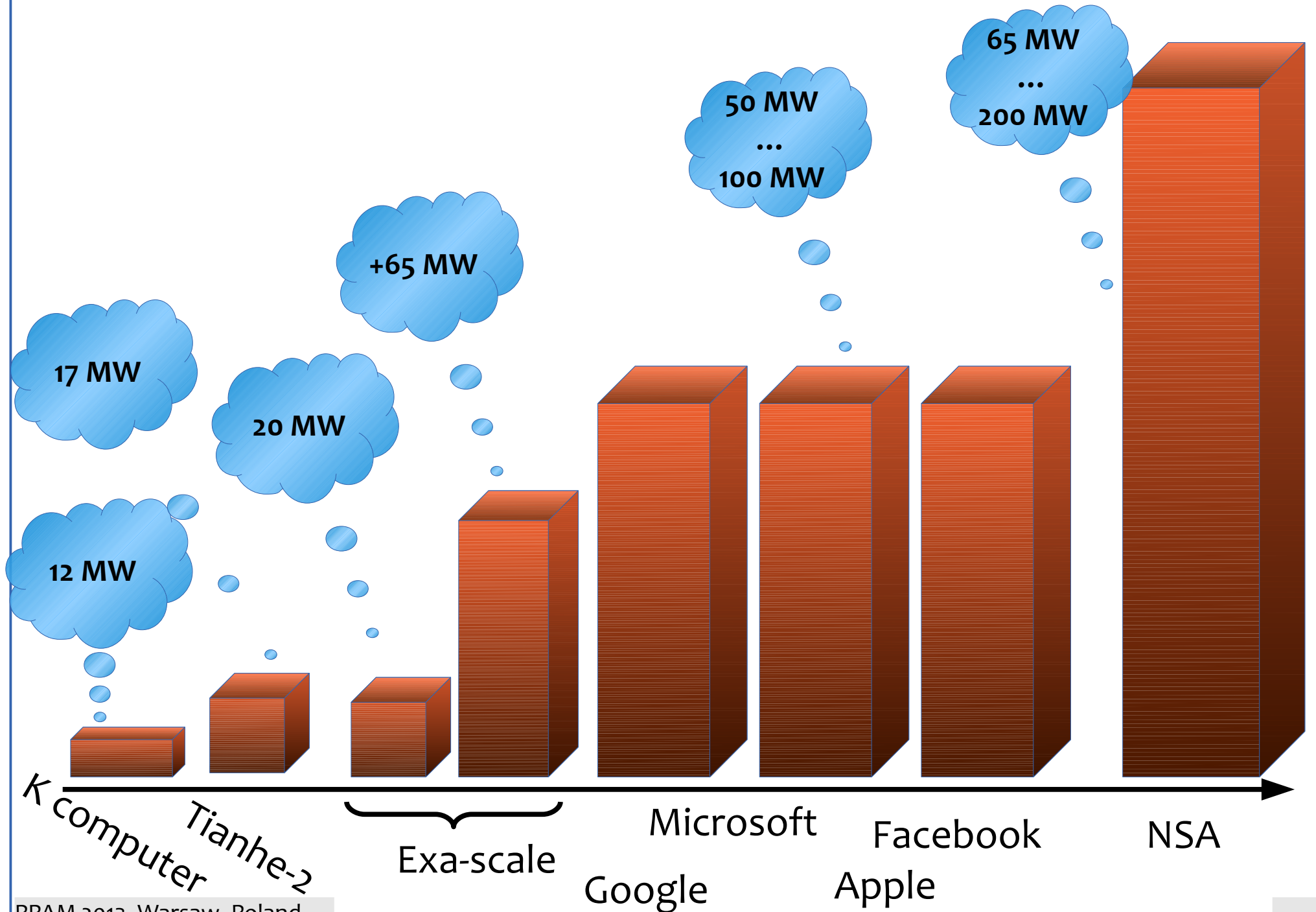
Projecting Execution Time to Exascale

Number Two Obstacle?

- **Power**
  - **Total power consumption**
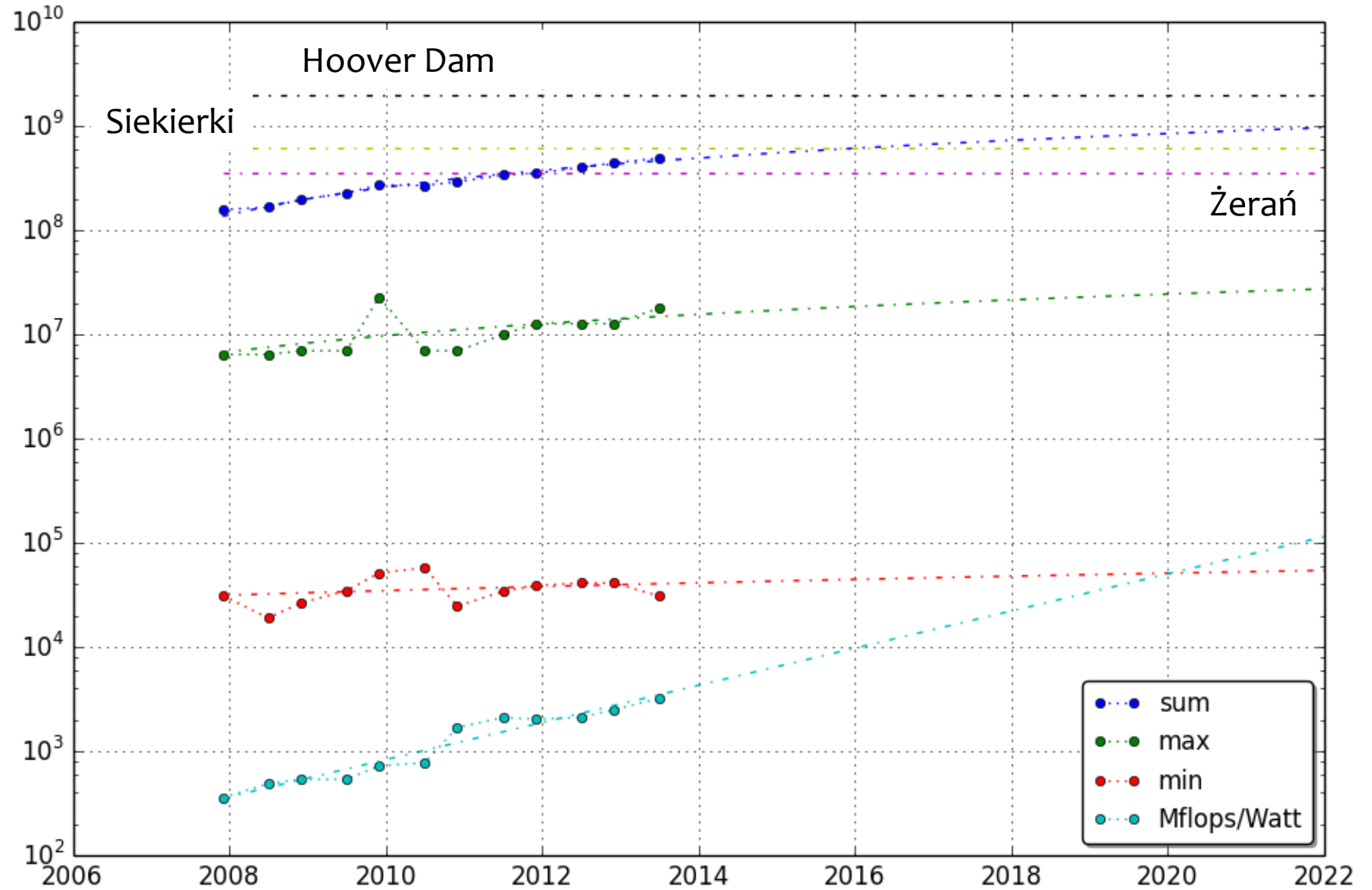 - **Gains in power efficiency**

Power Consumption

- **Power consumption is an issue throughout the chip industry**

- **Supercomputing is hit hard from all sides**

  - **CPU/GPU/Accelerator/Coprocessor**

  - **Memory**

  - **Interconnect**

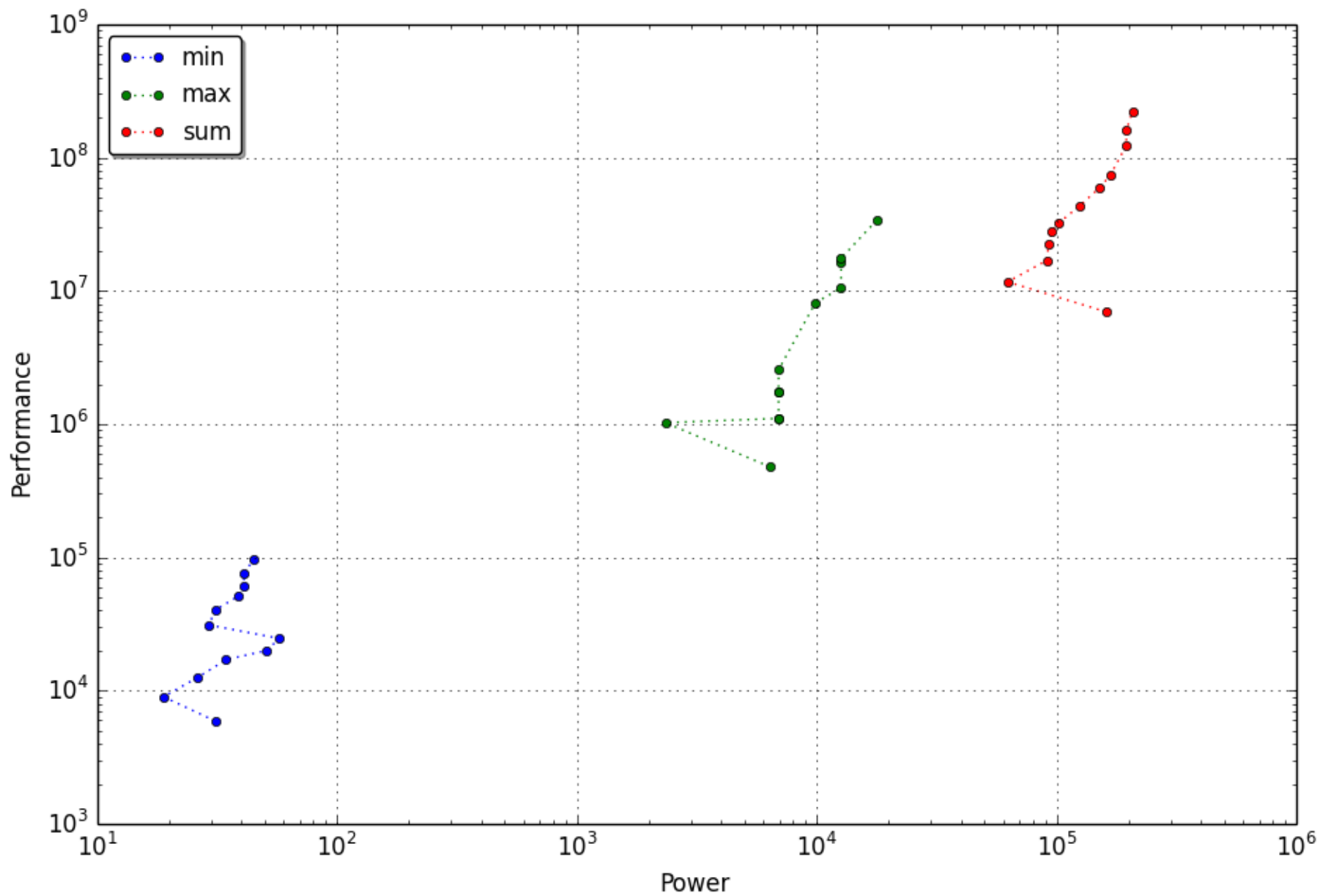- **Power consumption projections for the exascale system continue evolving**

Power Consumption in Perspective

17 MW

12 MW

20 MW

+65 MW

50 MW ... 100 MW

65 MW ... 200 MW

K computer

Tianhe-2

Exa-scale

Microsoft
Google

Facebook
Apple

NSA

Power Trends Towards Exascale

Another Way to Look at Power and Performance

How to Measure the Cost of the System

- **Initial system cost has always been of questionable value**
  - **Often, part of a deal with the vendor and the price cannot be disclosed**
    - **Large supercomputer installations create good "bragging rights" opportunity**
  - **Vendor might ship early and/or development parts which are discounted**
    - **Getting the system up and running for a press release**
      - **The early systems are "ASAP quality"**
- **Life time costs have become prominent over the years**
  - **Energy is the biggest component**

Energy Trends