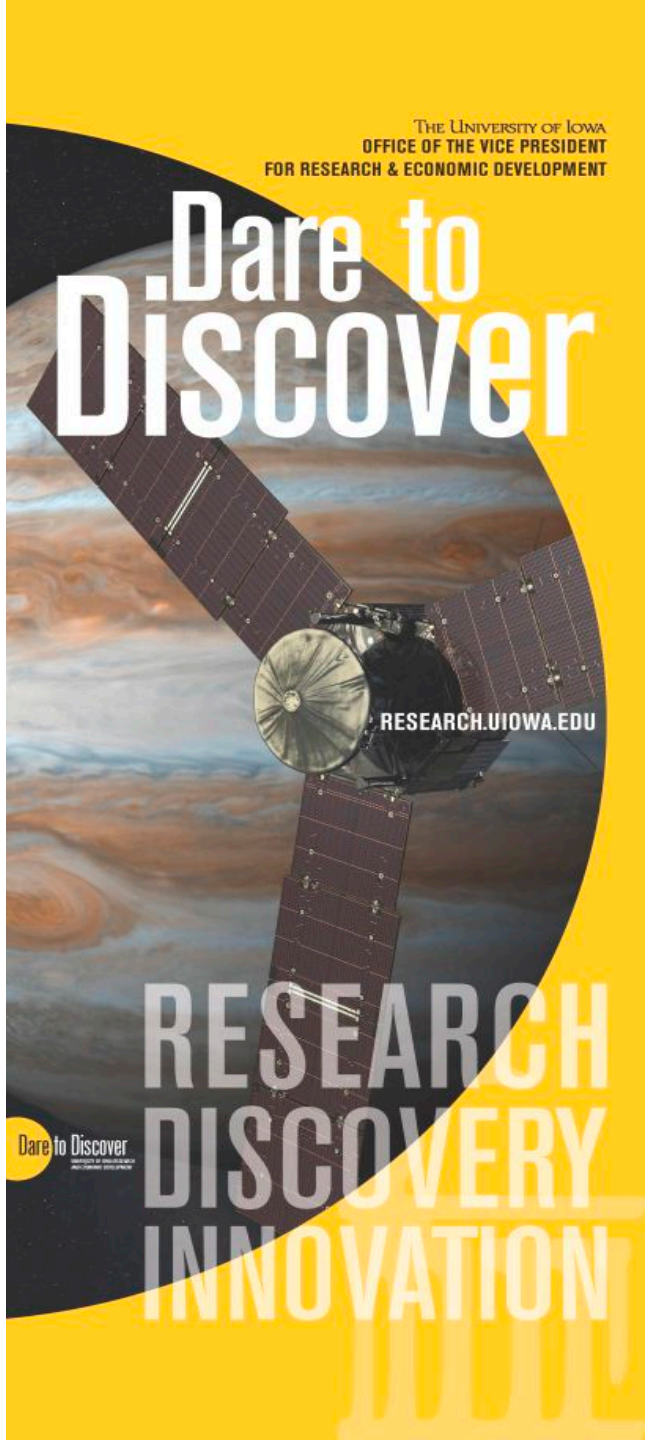
HPC, Energy Efficiency and TCO

Dan Reed
dan-reed@uiowa.edu
University of Iowa

November 13, 2016

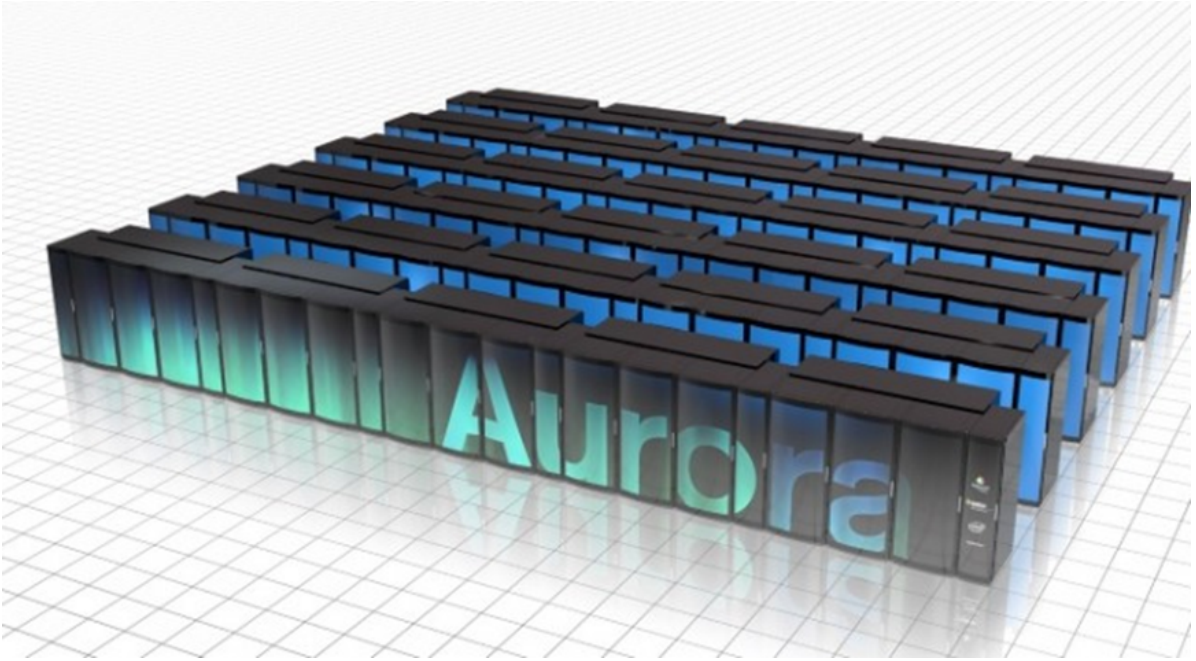


Rising cost of leading edge HPC

Cray X-MP/48 (1984) ~$15M plus disks (~$35M in current dollars)
Blue Waters (2013) **~$226M**

Collaboration of Oak Ridge, Argonne, and Lawrence Livermore (CORAL)

- LLNL and ORNL **$325M** (2014)
- Argonne **$200M** (2015)



Rising cost of leading edge HPC

National Supercomputer Center (Wuxi)

- Sunway TaihuLight, SW26010 (260C) + Custom
  - 10.6M cores cores, 93 PF (HPL), 125 PF (peak), 15.4 MW
  - 1.8B RMB **(~$280M** – building, hardware, software, …)
- Two other major efforts (NUDT and Sugon)

Post K system

- 50X (capability) and100X (capacity)
- 30-40 MW, ARM v8 extensions
- 130B yen **(~$1.29B**)

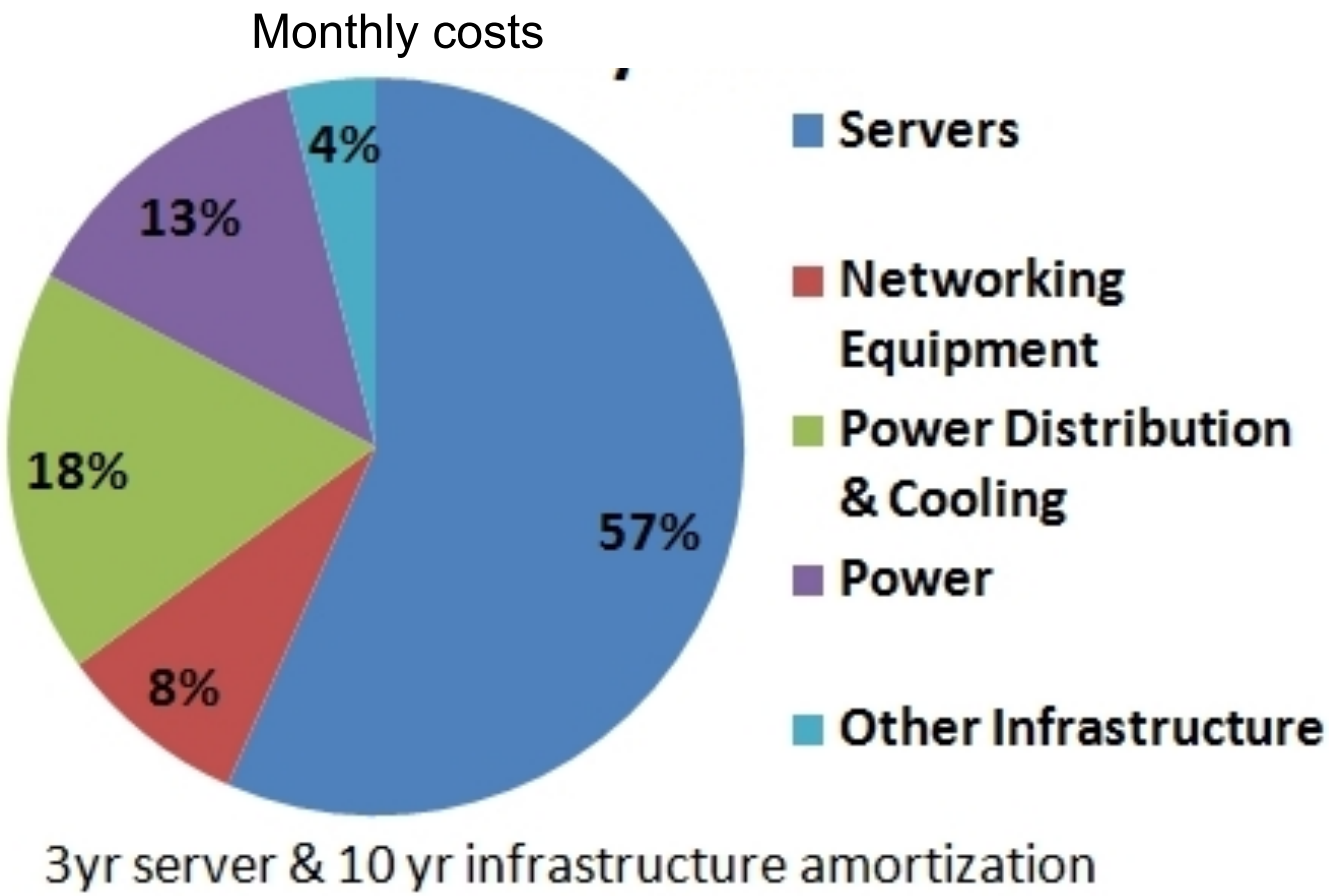
PRACE

- Advanced computing access

HPC apps acceleration	Post-K	FX100	FX10	K computer
FMA: Floating-point multiply and add	✔	✔	✔	✔
Math. acceleration primitives*	✔Enhanced	✔Enhanced	✔	✔
Inter-core barrier	✔	✔	✔	✔
Sector cache	✔Enhanced	✔Enhanced	✔	✔
Hardware "prefetch" assist	✔Enhanced	✔Enhanced	✔	✔

*Mathematical acceleration primitives include trigonometric functions, exponential functions, etc.

James Hamilton's observation (circa 2010)



Infrastructure matters!

http://perspectives.mvdirona.com/2010/09/overall-data-center-costs/

Belady's tent servers

Five HP DL585s in a tent (2007-2008)

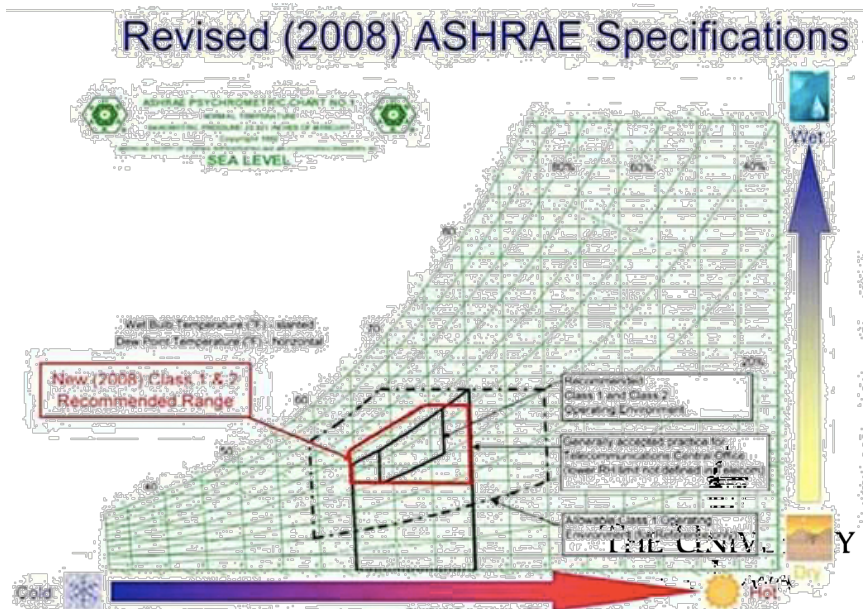
- Water dripped from the tent onto the rack
- A windstorm blew a section of the fence onto the rack
- An itinerant leaf was sucked onto the server fascia

There were zero failures …

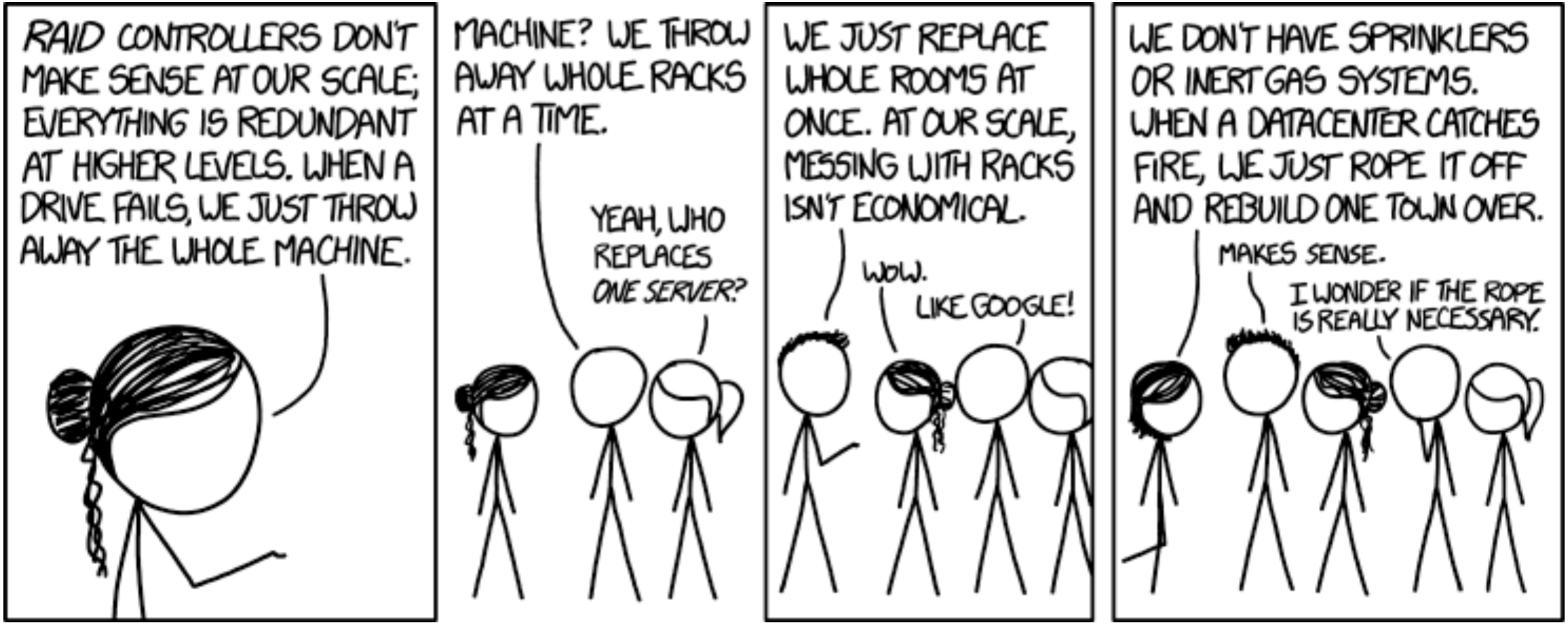
Led to a rethinking of psychometrics

- High temperature operation
- Air side economization

and energy consumption



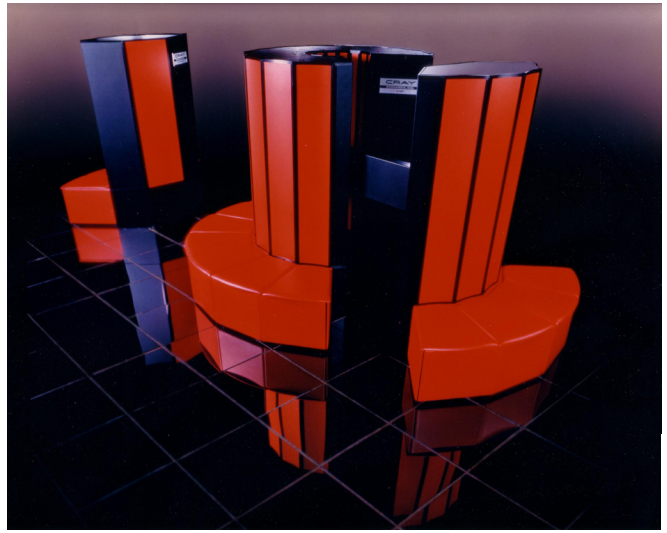
Yeah, big is really BIG



http://xkcd.com/1737/

… and the rope is an extra cost

Shifting cost balance: NCSA@30



$15M (1984)

»

~$1M (1984)

NCSA

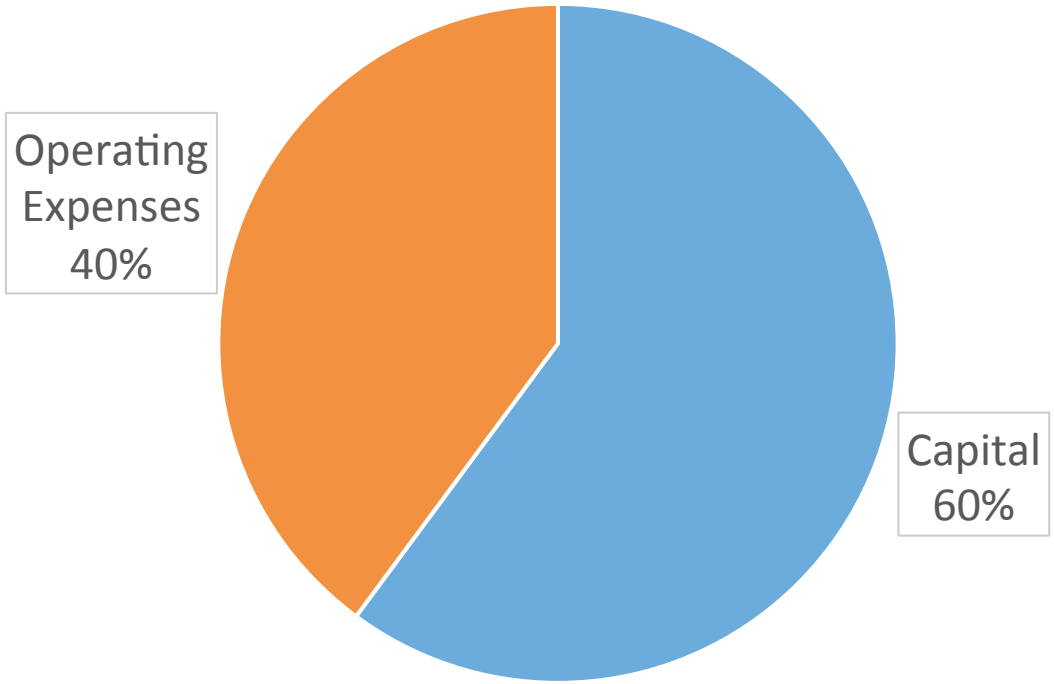
$226M (2013)

~$85M (2013)

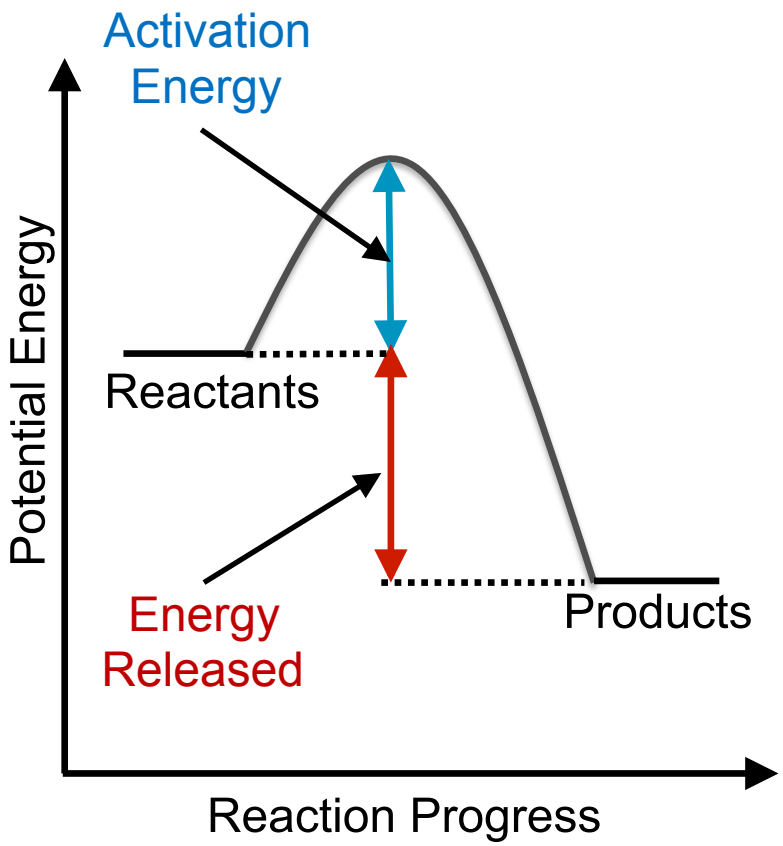
The University of Iowa

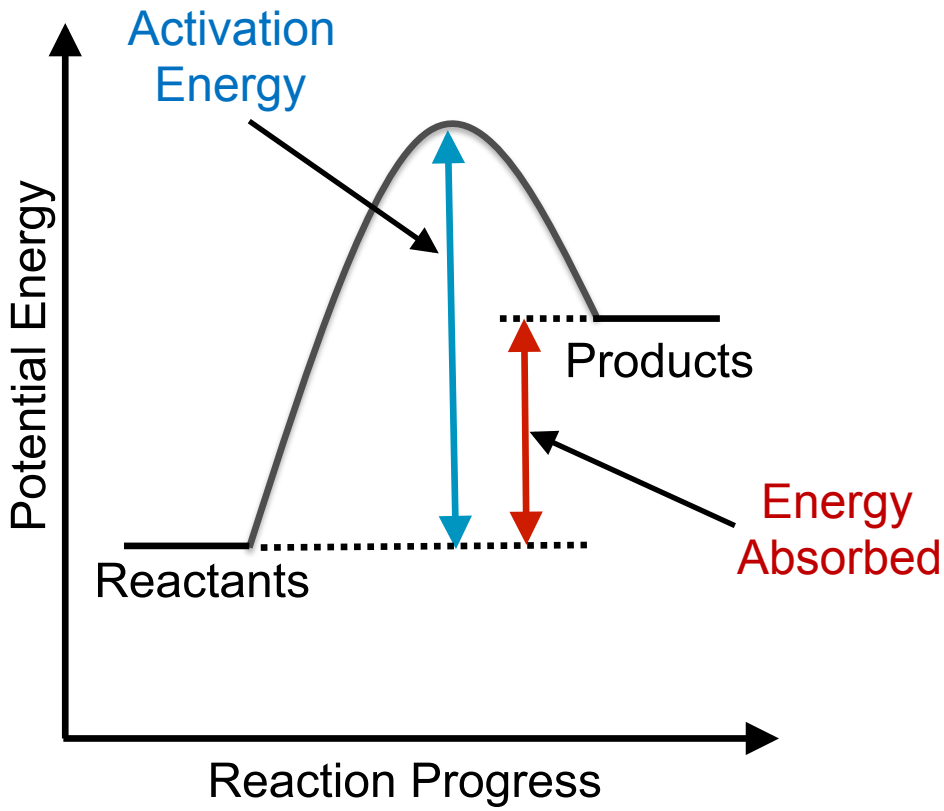
Blue Waters capital and operating expenses

Nominal five year lifetime



Endothermic and exothermic processes





Infrastructure can make money, cost money, or break even

Must include business needs/lifetime –can mean "throwaway" infrastructure

Total cost of ownership (TCO)

TCO components

- Computing system
- Software (licenses, etc)
- Technical staff (operations, support, development)
- ***Energy and utilities (power, cooling)***
- Physical plant
- Debt service and cost of money
- Depreciation (equipment, facilities)

Replacement cost

- Rarely included in academic/government TCO models
- Essential in commercial TCO models

making fair comparisons more challenging

Commercial viability metrics

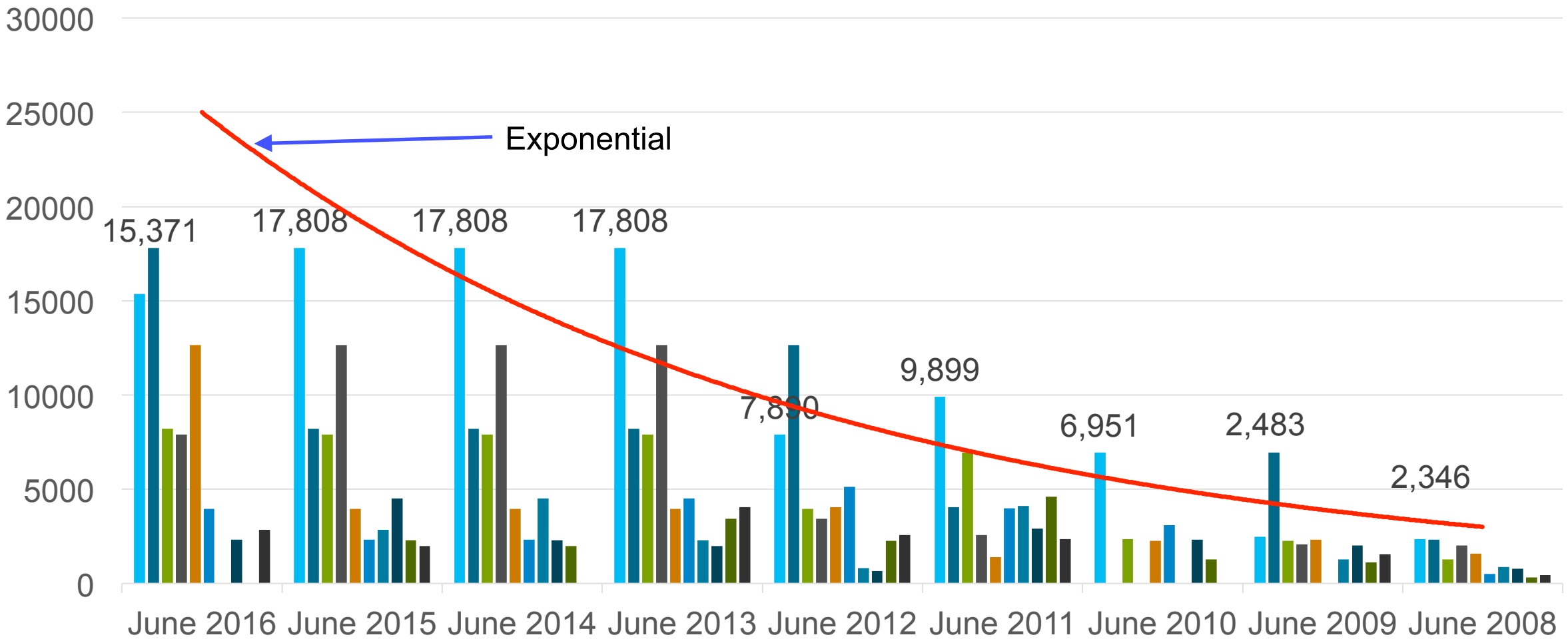
- NPV (net present value)
- IRR (internal rate of return

$$NPV = \sum_{t=1}^{T} C_t / (1+r)^t - C_0$$

where

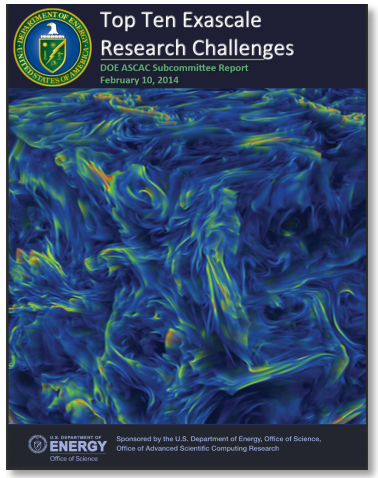
- C_t is net cash flow during period t
- C_0 is the total initial investment cost
- r is the discount rate (interest rate)
- t is the number of periods

TOP500 power consumption (top 10)



THE UNIVERSITY OF IOWA

Top ten exascale challenges (DOE)



Energy efficiency (circuits, power, cooling)

Interconnects (performance & energy efficient data movement)

Memory (*capacity and bandwidth*)

System software (scalable, power and resilience aware)

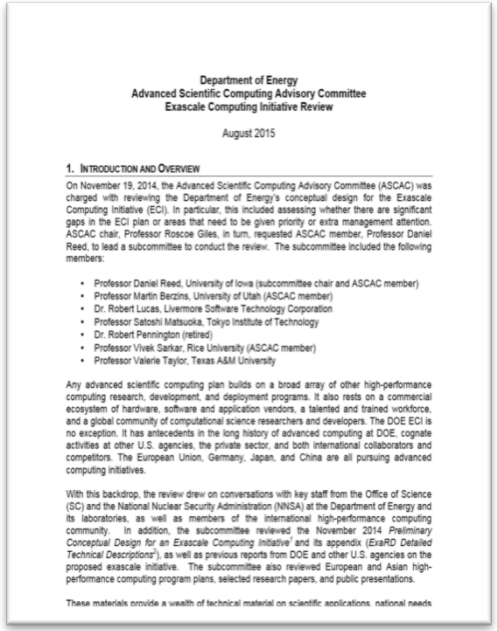
Programming systems (massive parallelism, data locality, resilience)

Data management (*data volume, velocity and diversity*)

Algorithms (mathematical optimization and uncertainty quantification)

Resilience and correctness (faults, reproducibility, verification)

Scientific productivity (software engineering tools and environments)



Department of Energy
Advanced Scientific Computing Advisory Committee
Exascale Computing Initiative Review

August 2015

1. Introduction and Overview

On November 19, 2014, the Advanced Scientific Computing Advisory Committee (ASCAC) was charged with reviewing the Department of Energy's conceptual design for the Exascale Computing Initiative (ECI). In particular, this included assessing whether there are significant gaps in the ECI plan or areas that need to be given priority or extra management attention. ASCAC chair, Professor Roscoe Giles, in turn, requested ASCAC member, Professor Daniel Reed, to lead a subcommittee to conduct the review. The subcommittee included the following members:

- Professor Daniel Reed, University of Iowa (subcommittee chair and ASCAC member)
- Professor Martin Berzins, University of Utah (ASCAC member)
- Dr. Robert Lucas, Livermore Software Technology Corporation
- Professor Satoshi Matsuoka, Tokyo Institute of Technology
- Dr. Robert Pennington (retired)
- Professor Vivek Sarkar, Rice University (ASCAC member)
- Professor Valerie Taylor, Texas A&M University

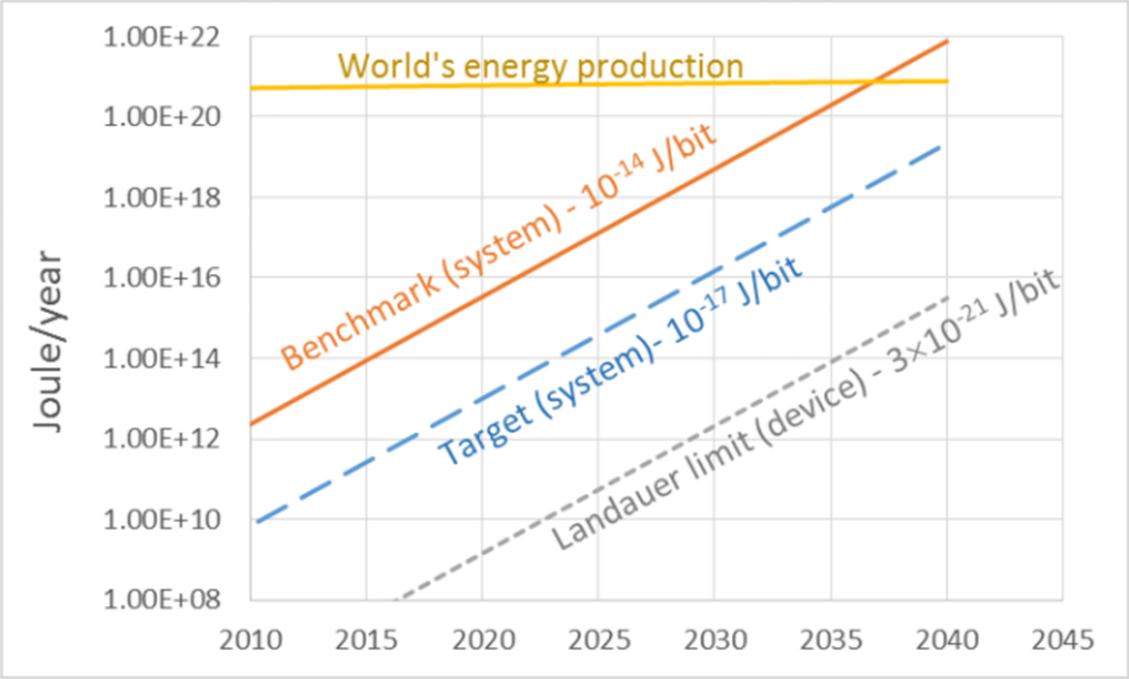
Any advanced scientific computing plan builds on a broad array of other high-performance computing research, development, and deployment programs. It also rests on a commercial ecosystem of hardware, software and application vendors, a talented and trained workforce, and a global community of computational science researchers and developers. The DOE ECI is no exception. It has antecedents in the long history of advanced computing at DOE, cognate activities at other U.S. agencies, the private sector, and both international collaborators and competitors. The European Union, Germany, Japan, and China are all pursuing advanced computing initiatives.

With this backdrop, the review drew on conversations with key staff from the Office of Science (SC) and the National Nuclear Security Administration (NNSA) at the Department of Energy and its laboratories, as well as members of the international high-performance computing community. In addition, the subcommittee reviewed the November 2014 *Preliminary Conceptual Design for an Exascale Computing Initiative*[1] and its appendix (*ExaRD Detailed Technical Descriptions*[2]), as well as previous reports from DOE and other U.S. agencies on the proposed exascale initiative. The subcommittee also reviewed European and Asian high-performance computing program plans, selected research papers, and public presentations.

These materials provide a wealth of technical material on scientific applications, national needs

http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/20140210/Top10reportFEB14.pdf

Abstraction and efficiency



100 picojoules can

- Run a Cortex M0 for 10 cycles
- Write 1 bit of FLASH or 300 bits of SRAM or DRAM
- Send 5 bits across LPDDR4
- Transmit 2 bits of UWB data
- Transmit 0.02 bits over Bluetooth LE

Remember

- 1 KW is 3.6 *megajoules*

Abstraction layers have been key to success …

- Physics, logic, ISAs, operating systems, runtimes, libraries, applications
- Ops/joule drives specialization but software favors generality

… but may now be our albatross

Sources: ARM, SIA/SRC (Rebooting the IT Revolution)

THE UNIVERSITY OF IOWA

We need a conceptual breakthrough …

Ecosystems in collision, driven by ops/joule and business models

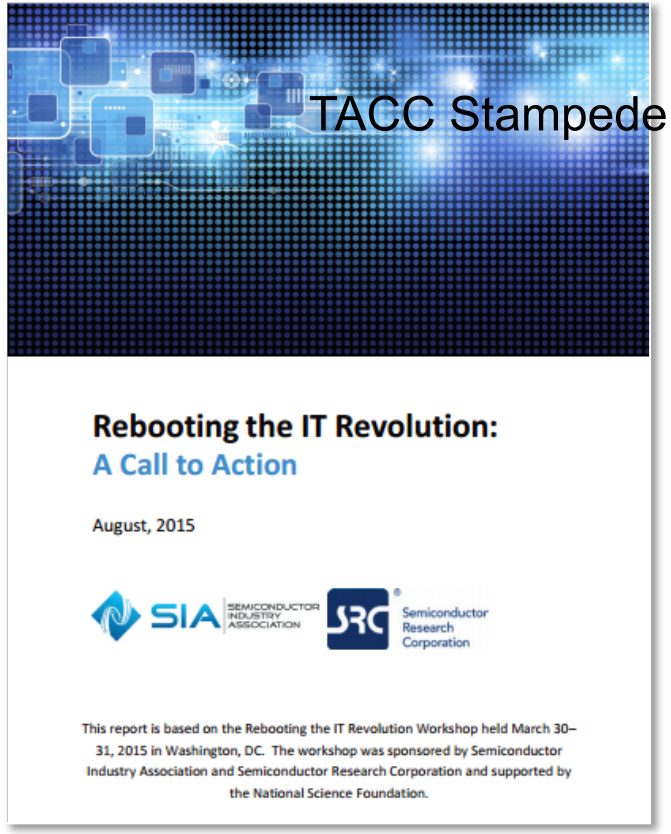
- General purpose vs. accelerators
- Special purpose vs. radical alternatives
- ***Parts volume vs. margins and value add***

The serious action is at the two extremes …

- Cloud data centers and exascale HPC
- Mobile and the Internet of Things (IoT)

But the solutions are probably not the same

- Quantum computing and/or superconducting devices in data centers
- Neuromorphic and analog in mobile devices



… and the best way is to try lots of things

THE UNIVERSITY OF IOWA

Culture and economics

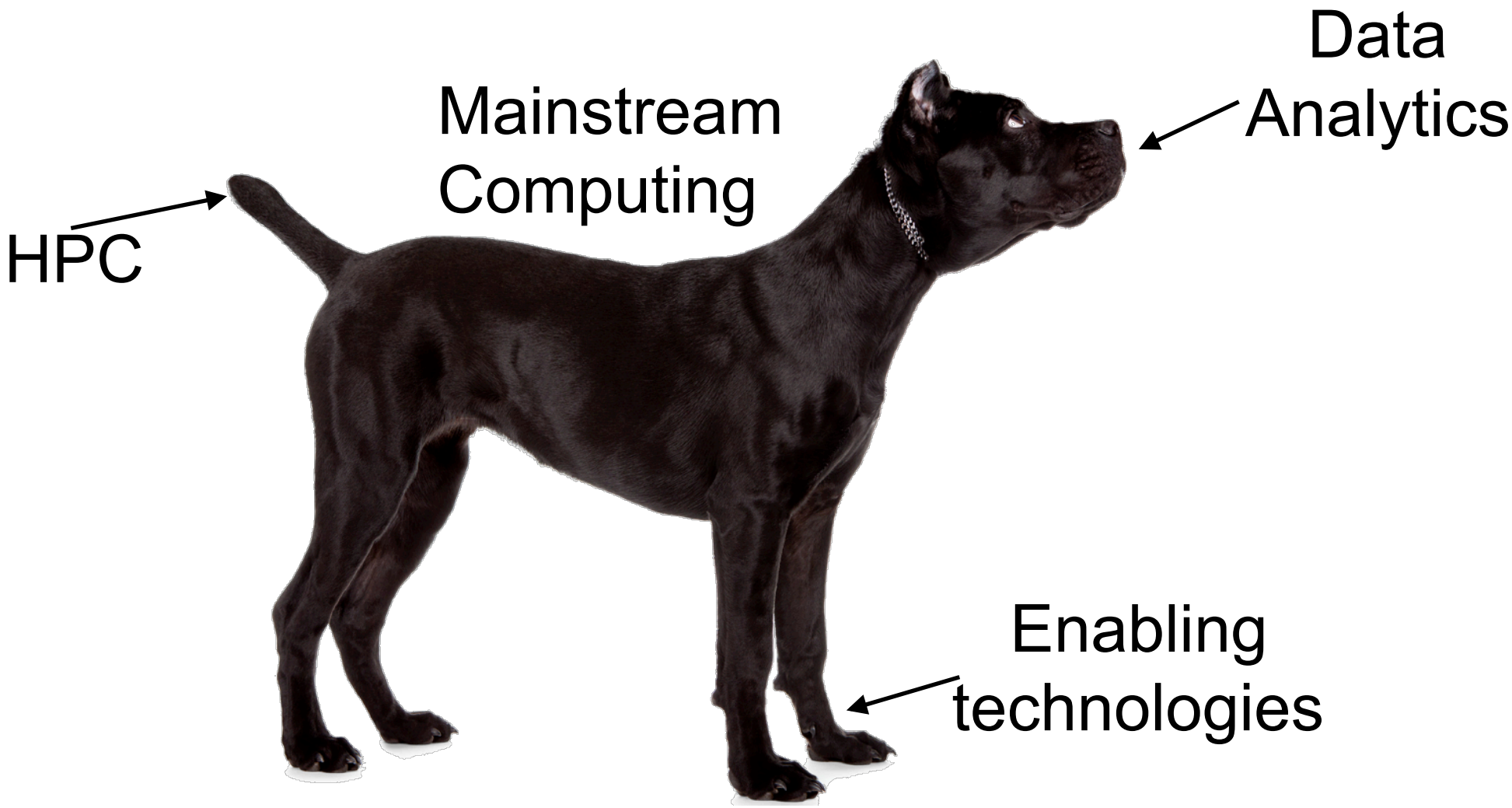
Industry
- Capital is cheap (look at interest rates)
- Labor is expensive
- ROI drives behavior

Academia and government
- Capital is expensive
- Labor is mostly cheap
- Other metrics drive success

To change the game, change the metrics …
- Infrastructure, personnel, social and political

Put another way, where you draw the bounding box shapes the answer

Follow the money and the users …



… or the money and the users may not follow YOU

THE UNIVERSITY OF IOWA

Discussion