
Microsoft® Research
Faculty Summit
10
YEAR ANNIVERSARY

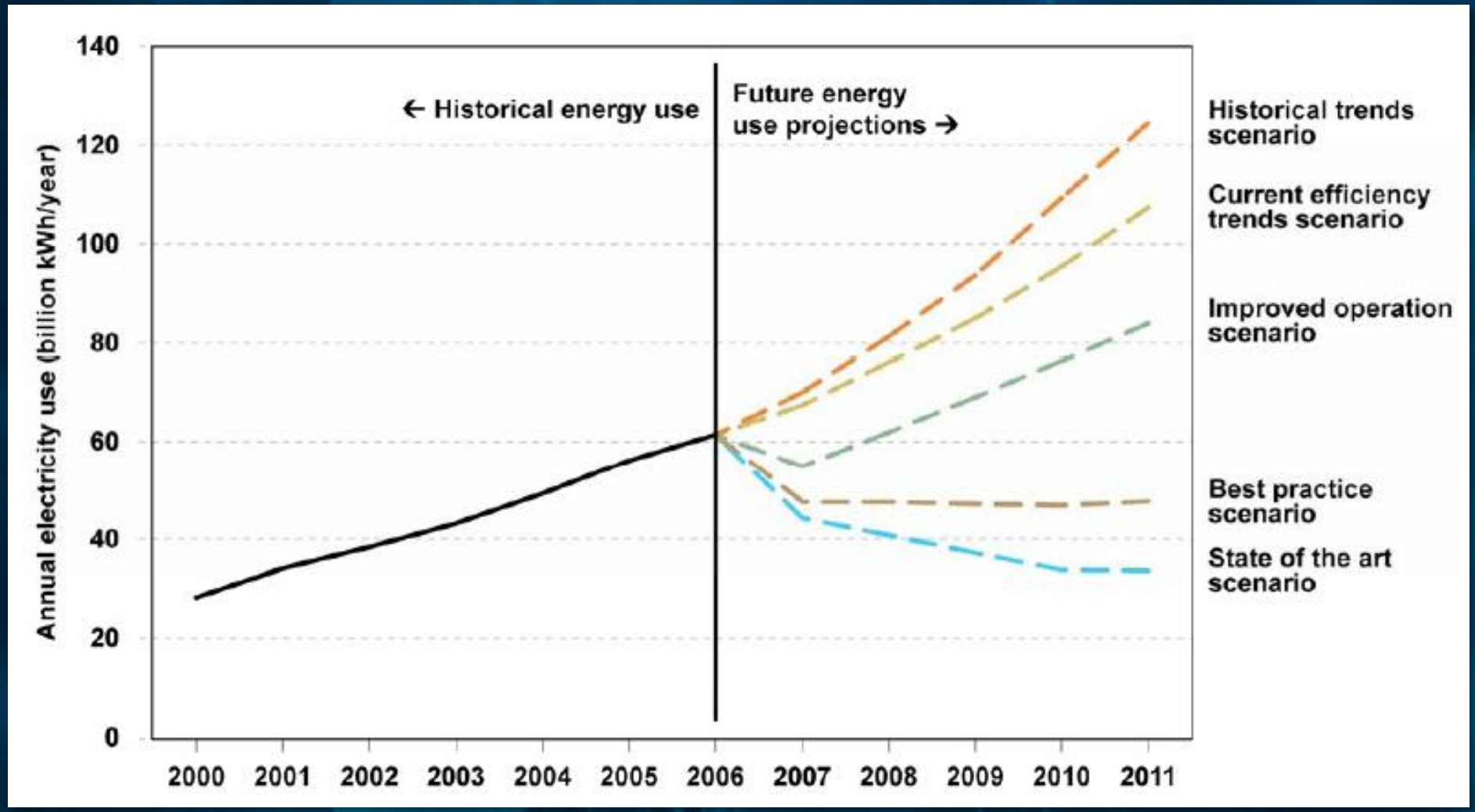
Microsoft Research

Energy-Efficient Computing: Emerging Technologies

Fred Chong
Director, Computer Engineering
Director, Greenscale Center for Energy-Efficient Computing
University of California at Santa Barbara

It's not all Hype…

Microsoft Research



EPA report, Public Law 109-231, 2007

Maybe a little Hype…

Microsoft Research

Server

=

SUV

- More precisely:
  - 80 billion terawatt-hr / yr = 6 million SUVs in carbon production (10 mpg, 11K miles/yr)
- Computing is the fastest growing segment of carbon production
- *But actually an issue of energy productivity*

The First Steps Have Been Taken

Microsoft Research

$$\text{Efficiency} = \frac{\text{Computation}}{\text{Total Energy}} = \left(\frac{1}{\text{PUE}}\right) \times \left(\frac{1}{\text{SPUE}}\right) \times \left(\frac{\text{Computation}}{\text{Total Energy to Electronic Components}}\right)$$

[Barroso and Hoetzle 2009]

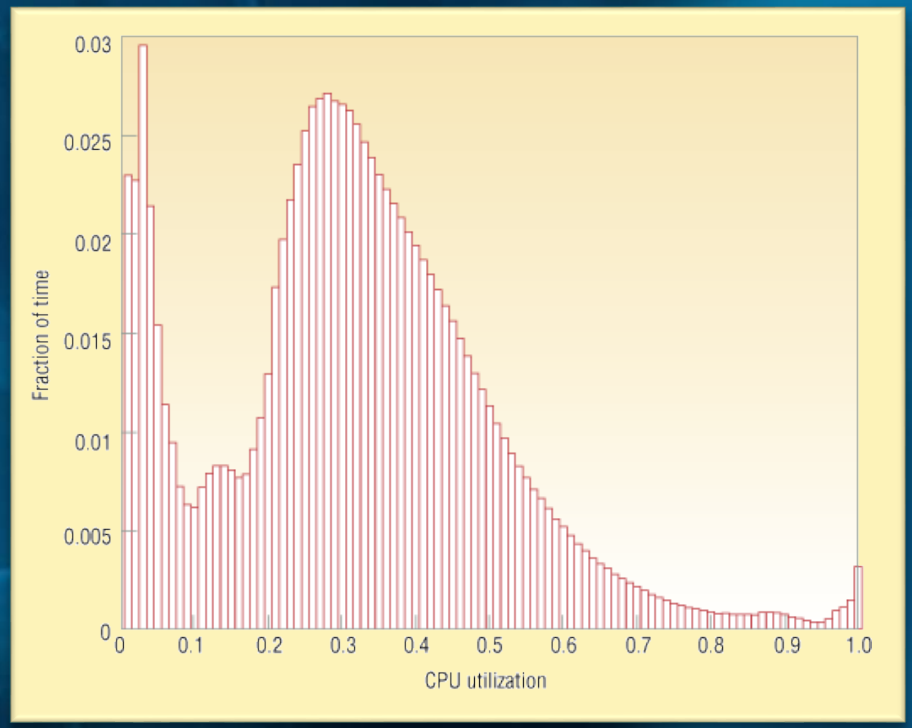
- Google reports infrastructure (PUE) and server (SPUE) efficiencies close to practical optimal (within 20%)
- Remaining gains in more:
  - *Efficient computation*
  - *Energy proportionality*

Google Cooling Tower

What's Left?

Microsoft Research

- More efficient computation
  - Parallelism
  - Communication
  - Storage hierarchy
- Energy proportionality
  - Infrastructure, power delivery, storage, networks, servers

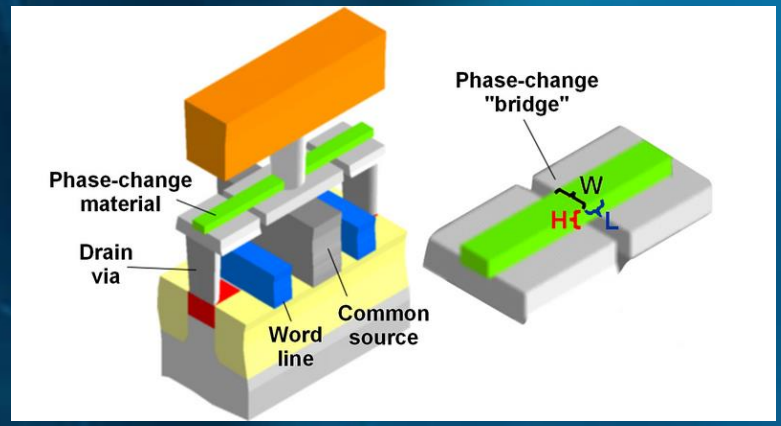


[Barroso and Hoetzle 2009]

Persistent Memory

Microsoft Research

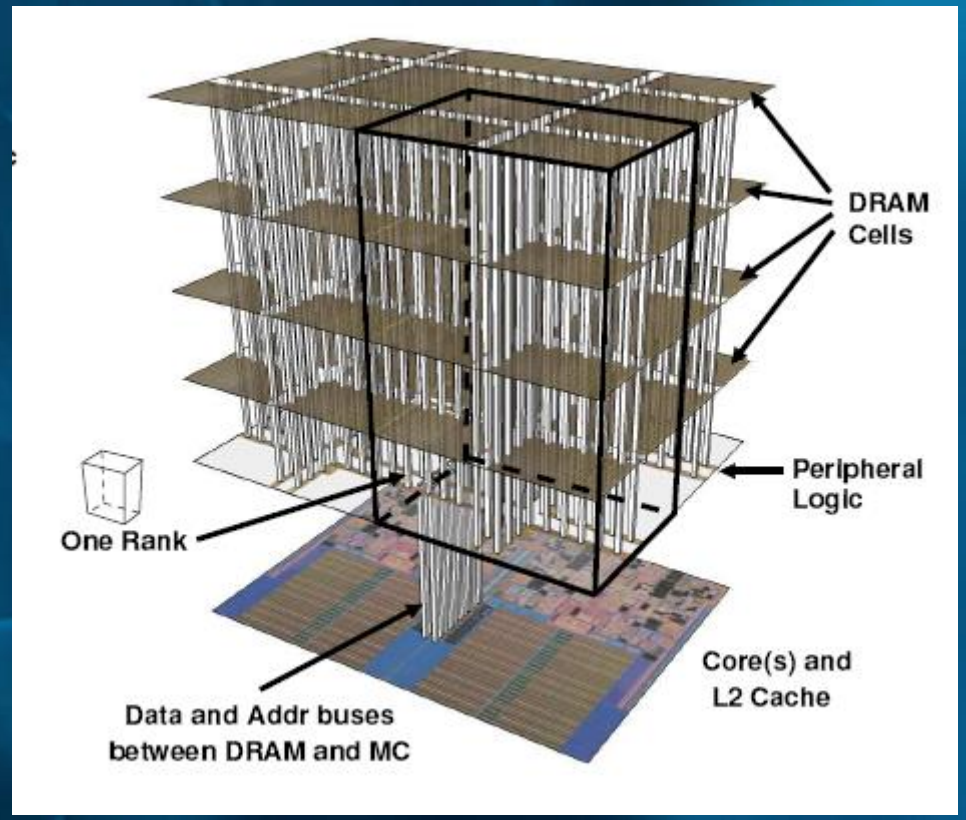
- Substantially better energy proportionality
  - But distributed system software/protocols must be restructured
- FLASH
  - 2.5X server performance/watt [Swanson ASPLOS09]
  - Lots of work on durability
  - Optically-assisted FLASH?
- Phase-change memory
  - Same active energy, 20% slower than DRAM [Lee ISCA 09]



3D Integration

Microsoft Research

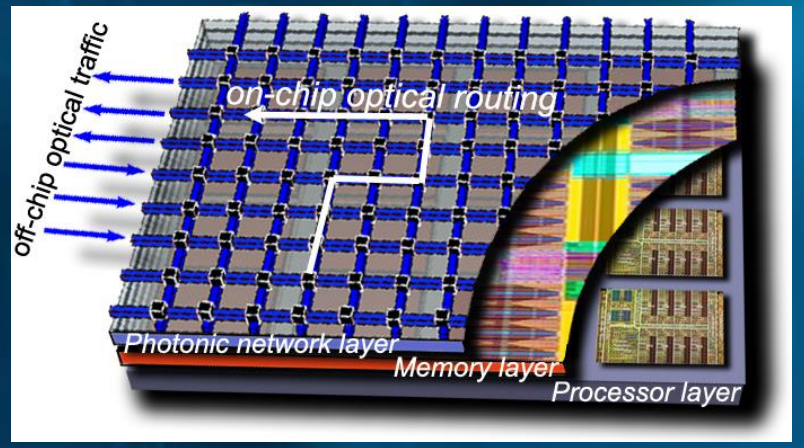
- Through-Silicon Vias (TSVs)
- 3D DRAM
  - 2X performance [Loh ISCA08]
  - 10X capacity?
- Power dissipation?
  - Vapor / liquid cooling
  - Superlattice microrefrigeration (600 W/cm^2, 35% efficiency)
  - Carbon nanotube thermal routing
- Combine different technologies
  - Logic, DRAM, Optics

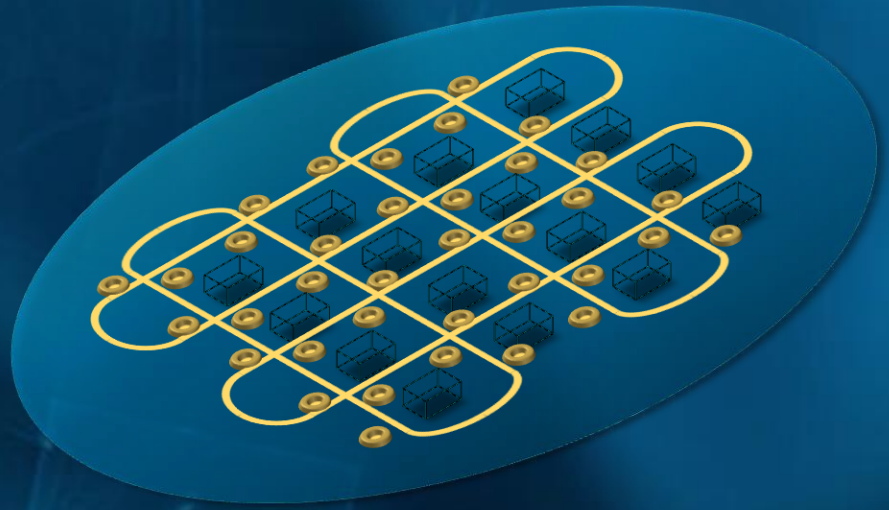


Silicon Photonics

Microsoft Research

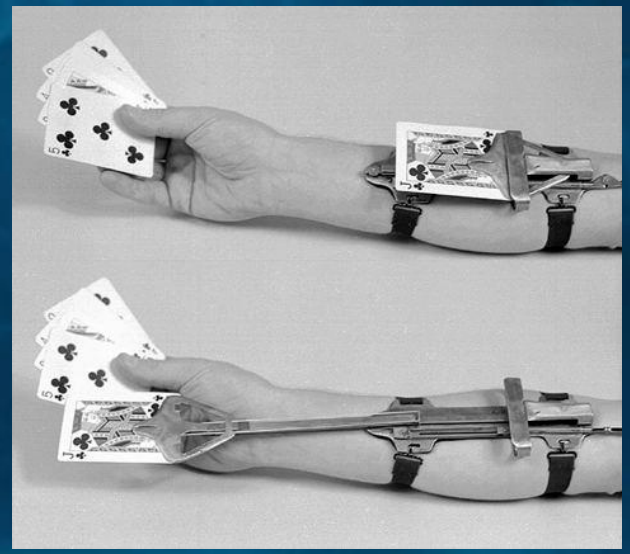
- Server backplane
  - 10 Tbit/s [Beals Appl. Phys A 2009]
- On-chip or on-wafer [Vantrease ISCA08] [Krishnamoorthy Proc IEEE 2009]
  - 1 Tbit/s per frequency per link
  - 100s of frequencies
  - 10-100s Tbytes/s bisection





Changing the Game

Microsoft Research



- Energy-efficient computation
  - Will likely exploit parallelism
  - Will need game-changing communication support
  - Will need game-changing storage technologies