

An Exa-Op Data Center at $<10\text{MW}$ by 2020? Too Many Ops and Not Enough Energy

Christopher Batten

Computer Systems Laboratory
School of Electrical and Computer Engineering
Cornell University

IAP Workshop on Data Center and Cloud
Ithaca, NY
October 2013

How much energy/op for an Exa-Op at 10MW?

What is an “operation”?

- ▶ integer operation : embedded computing
- ▶ floating point operation : high-performance computing
- ▶ server side Java operation (ssj_op) : SPECpower for server apps

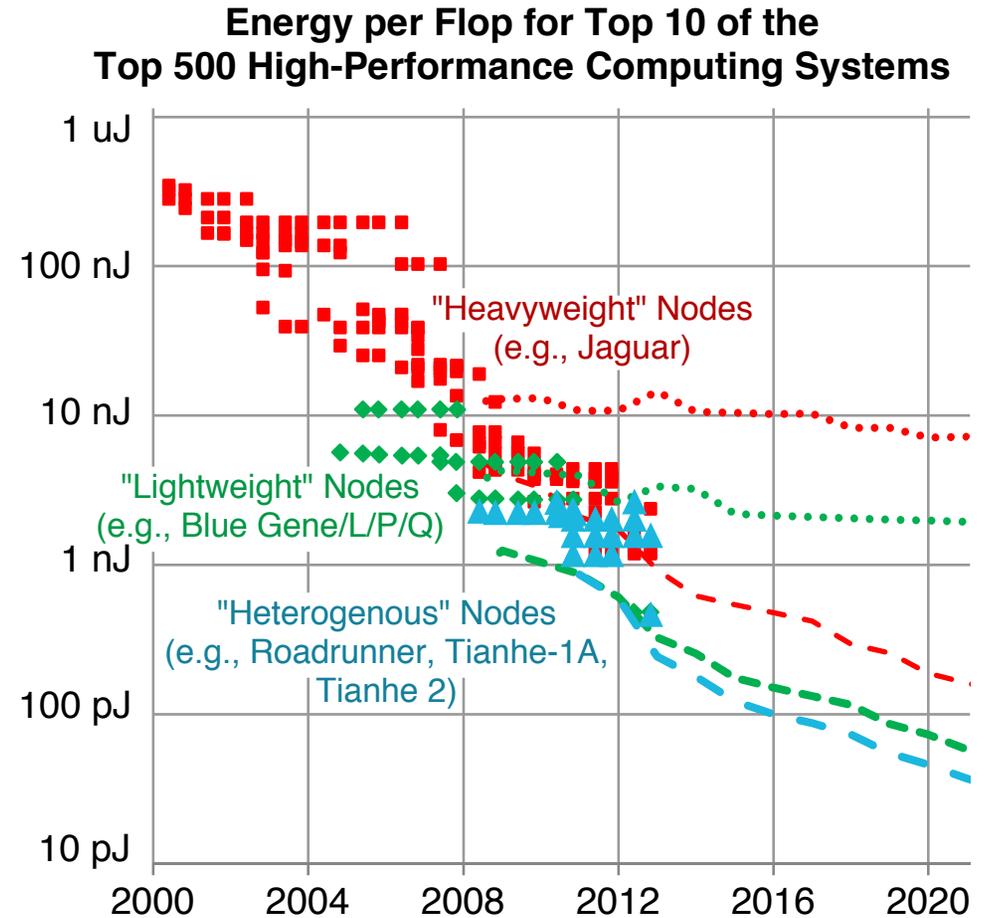
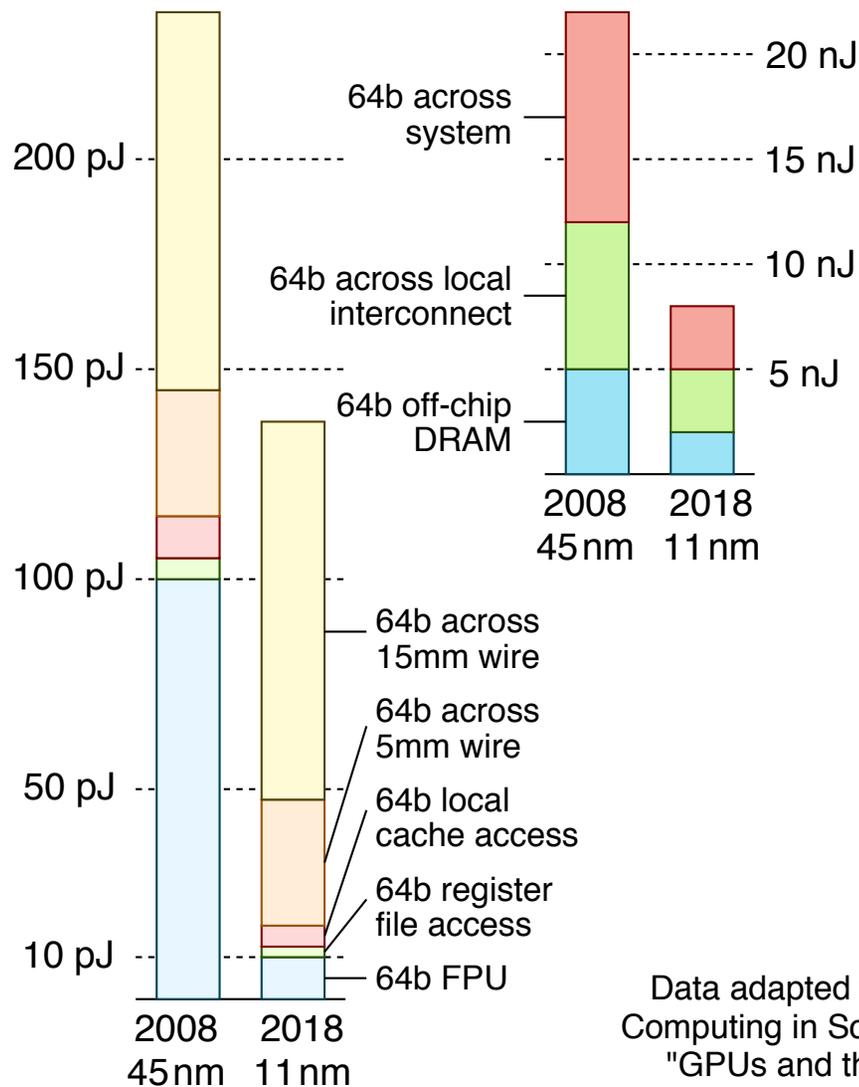
$$\frac{10 \text{ MW}}{1 \text{ Exaflop/s}} = \frac{10 \times 10^6 \text{ J/s}}{1 \times 10^{18} \text{ flop/s}} = 10 \text{ pJ/flop}$$

$$\frac{100 \text{ W/socket}}{10 \text{ pJ/flop}} = 10 \text{ Teraflop/socket} \quad \frac{1 \text{ Exaflop/s}}{10 \text{ Teraflop/socket}} = 100,000 \text{ sockets}$$

$$\frac{10 \text{ Teraflop/socket}}{2 \text{ GHz}} = \frac{10 \times 10^{12} \text{ flop/s}}{2 \times 10^9 \text{ cycle/s}} = 5,000 \text{ FPUs/socket}$$

Assuming 4–16 FMACs/core, this requires 150–625 cores/socket

Energy per Operation Projections for HPC Systems



Data adapted from: **P. Kogge and J. Shalf**, "Exascale Computing Trends," Computing in Science and Engineering, Nov 2013; **S. Keckler, B. Dally, et al.**, "GPUs and the Future of Parallel Computing," EEE Micro, Sep/Oct 2011.

Is there hope?

- ▶ Does applying HPC projections to the data center and cloud make sense?
 - ▷ **Argument For:** Convergence on commercial off-the-shelf nodes and customized interconnects?
 - ▷ **Argument Against:** Data center and cloud workloads look significantly different than HPC workloads?
- ▶ An Exaop at 10MW may be possible sometime in the next decade (after 2025?), but it will take rethinking all layers of the “computing stack”

