Bits/Joule

A Thermodynamics approach to High Performance Computing

Troy Benjegerdes,
Sandia CA Seminar, Dec, 2009
http://bitspjoule.org, <hozer@bitspjoule.org>

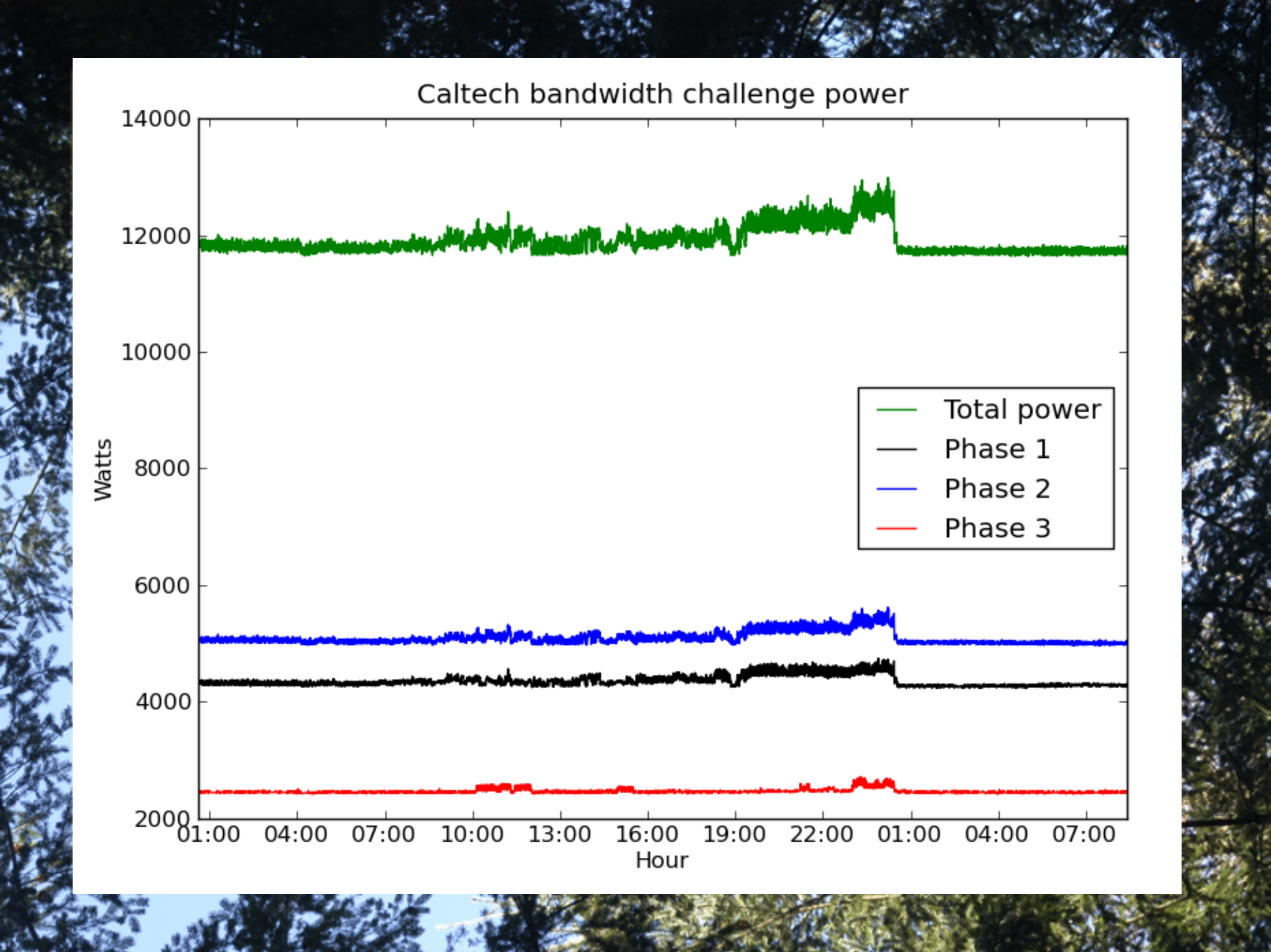
What is work?

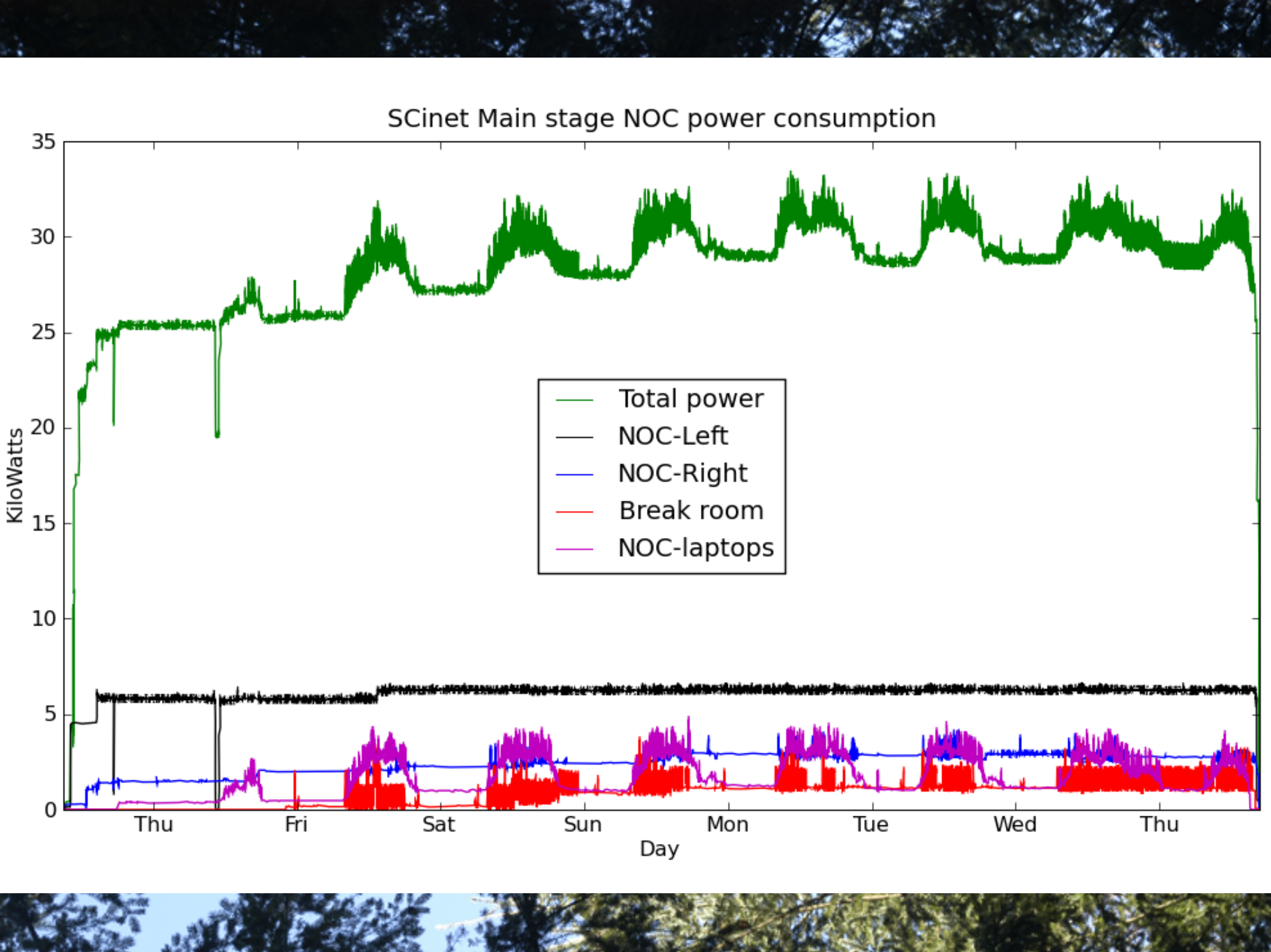
- Some specifics:
  - HPLinpack: Floating Point Operations
  - Physics/Chemistry: Joules (Newton-Meters)
  - Theoretical chemistry: Number of GAMESS/NWChem/MPQC jobs completed
  - Grad student: number of papers published
- Human work vs Thermodynamics work
  - We'll stick with Joules, the work done by an agent exerting a force F over distance D
    - Less messy value judgments here

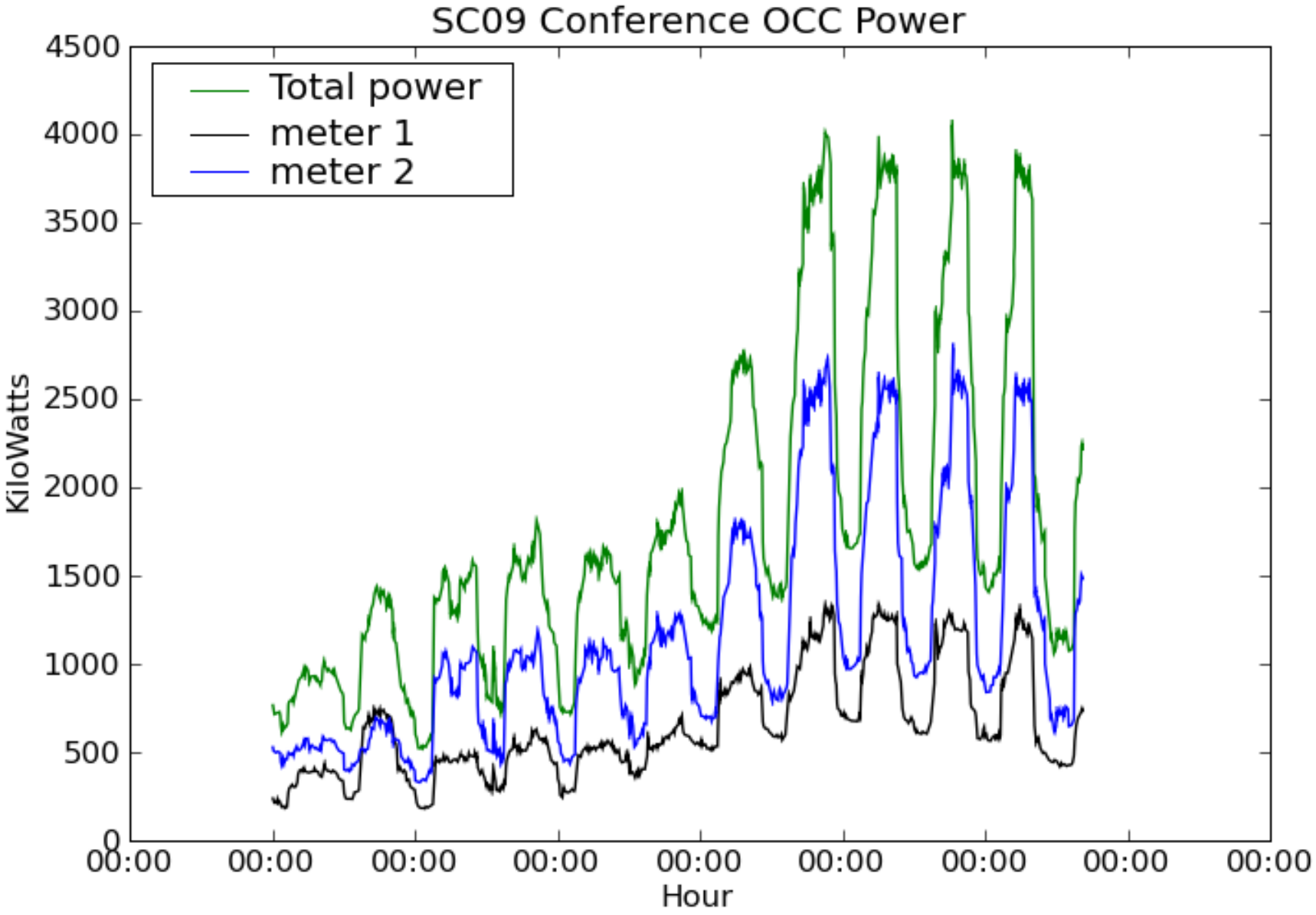
Benchmarking this presentation

- Energy(work) budget
- ~72 kJ for laptop
  - 20 watt-hours
- ~9.1 GJ for air travel
  - 6500 km, 1.4 MJ/km
- ~10 Joules for whiteboard
  - F * D

- Bit budget
- 1 MB pdf
  - 4Mbits
- 40 Gbytes of face-time (*WAG)
  - 320Gbits
- ~320/9.1
  → 34 Bits/Joule


Caltech bandwidth challenge power
14000
12000
10000
8000
6000
4000
2000
Watts
01:00
04:00
07:00
10:00
13:00
16:00
19:00
22:00
01:00
04:00
07:00
Hour
Total power
Phase 1
Phase 2
Phase 3


SCinet Main stage NOC power consumption
KiloWatts
Day
Total power
NOC-Left
NOC-Right
Break room
NOC-laptops
0
5
10
15
20
25
30
35
Thu
Fri
Sat
Sun
Mon
Tue
Wed
Thu


SC09 Conference OCC Power
Total power
meter 1
meter 2
KiloWatts
4500
4000
3500
3000
2500
2000
1500
1000
500
0
00:00
00:00
00:00
00:00
00:00
00:00
00:00
00:00
00:00
Hour

Bits/Joule & FLOPS

- Is Bits/Joule a superset of FLOPs/Second?
- [1] Lower bound interconnect latency is d/C (distance divided by speed of light)
- Any conceivable computing technology converts electricity to computations, with heat as a co-product (okay, anything I can think of)
- Smaller distance → higher Watts/m^3
- [2] Computation (aka FLOPS) will (or are) bound by heat dissipation (thermodynamic) considerations

To Exascale, and Beyond

- Due to [1] and [2], a thermodynamic work based benchmark may be in order
- Bits moved, divided by energy (joules) used.
  - Bits in/out of floating point unit is a linear function of FLOPs
- Performance of an FPGA on double precision floating point matrix mutiply is limited by bits in/out of the FPGA
- Some Exascale applications are inherently limited by storage bandwidth

A thought experiment..

- Bits/Joule of
  - Turing machine
  - Hand calculation
  - Babbage difference engine
  - Supercomputing 2009 conference
    - CalTech Bandwidth Challenge
  - LHC experiment(s)
  - Jaguar/Roadrunner
  - Google/DOE/ESnet