OpenCL case study: Accelerating molecular docking on multi- and many-core computer architectures

Simon McIntosh-Smith

University of Bristol, UK

simonm@cs.bris.ac.uk

University of BRISTOL

Acknowledgements

- Richard Sessions, Amaurys Avila Ibarra, Jon Crisp
  - Authors of the BUDE docking code
  - School of Biochemistry, University of Bristol
- Terry Wilson
  - Original OpenCL port
  - Computer Science, University of Bristol

University of BRISTOL

Power-limited regimes

- Processor power consumption now has an upper bound (may even reduce over time)

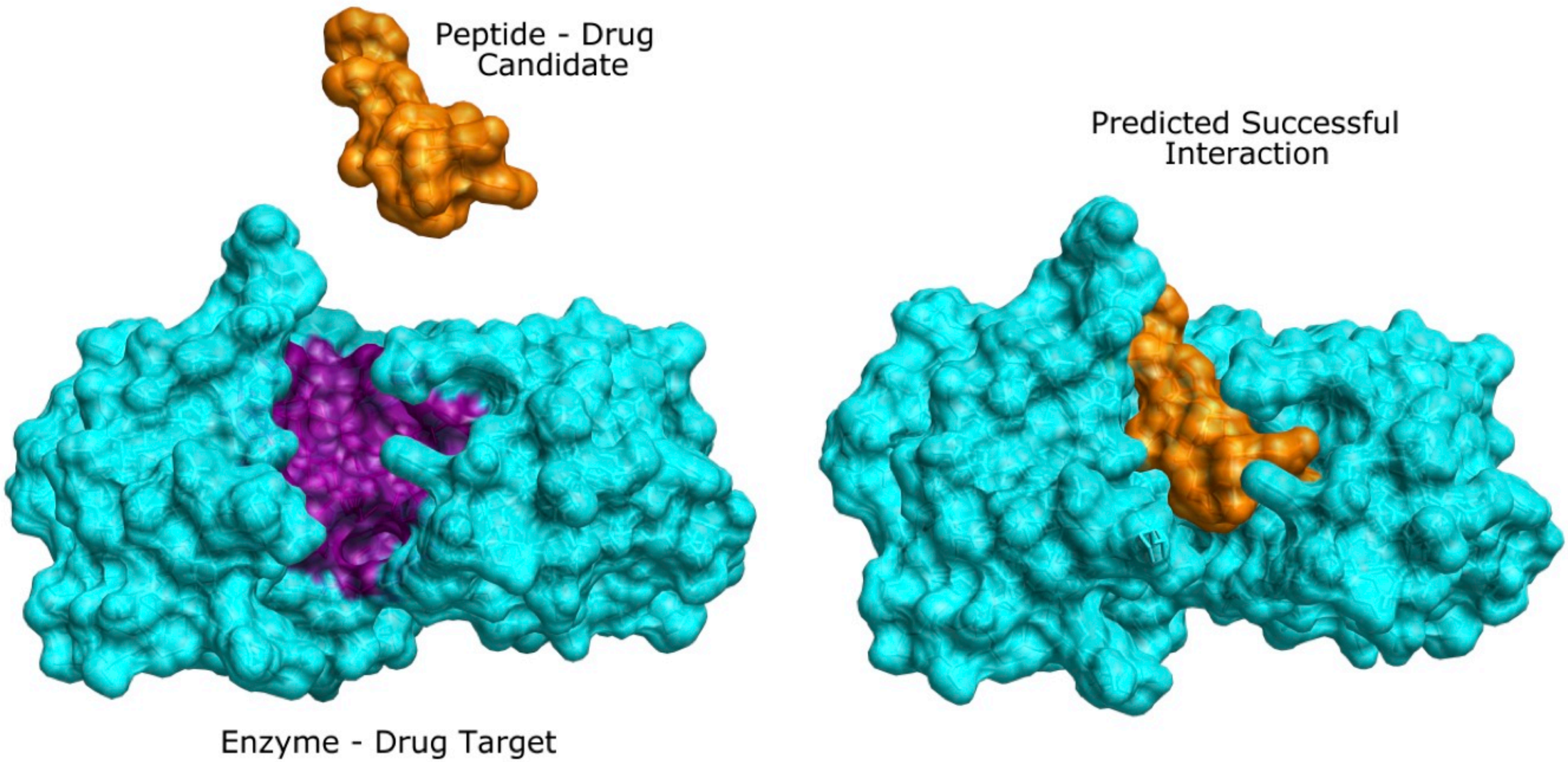
- Power consumption proportional to:
  - Clock frequency
  - Number of transistors (chip area)
    - Number of cores
  - Voltage squared

- When power has an upper bound, **“performance per watt = performance”**

- Driving growing interest in GPUs

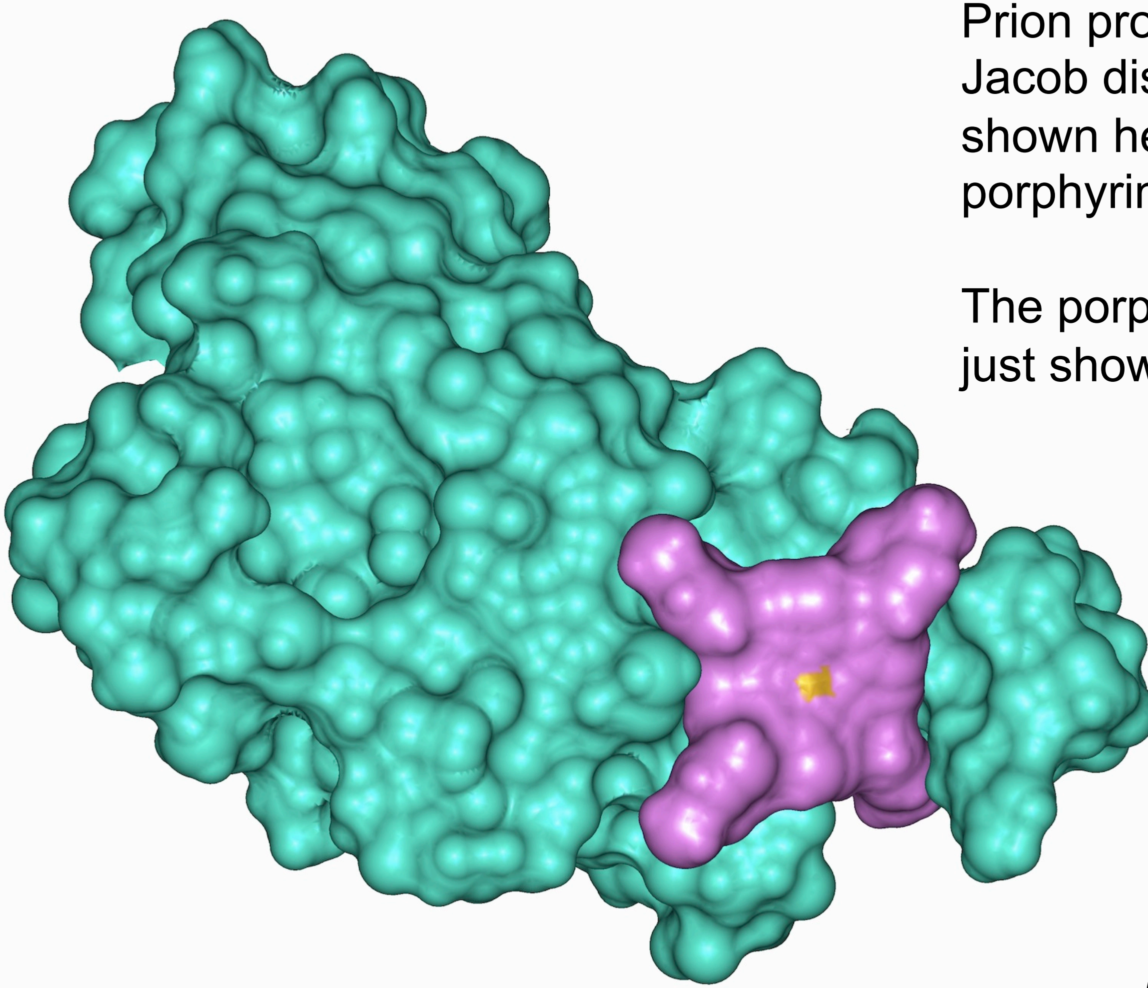
University of BRISTOL

Drug docking examples: Elastase inhibitors



University of BRISTOL

Prion disease



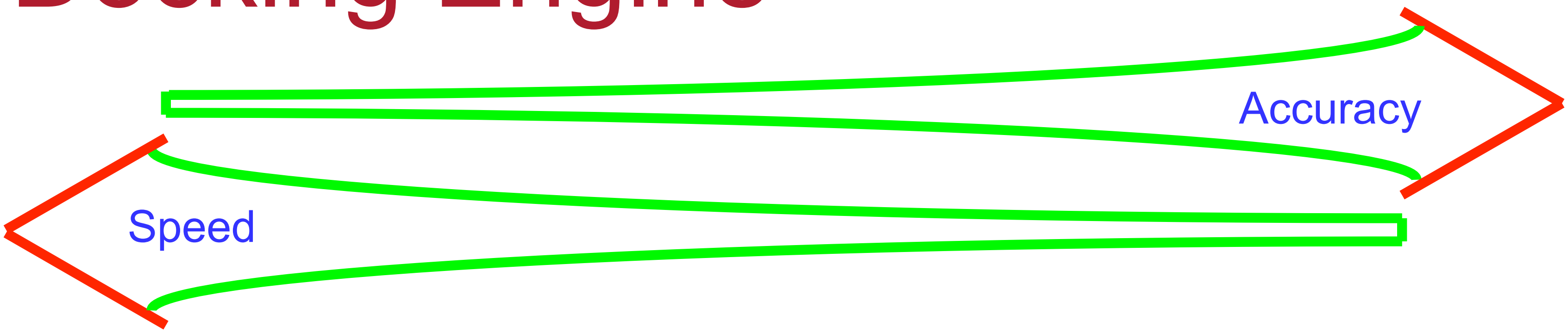
Prion protein behind Creutzfeld-Jacob disease in humans and shown here binding with a (pink) porphyrin-based ligand

The porphyrin's bound iron ion is just showing in yellow

1,719 atoms in the protein
53 atoms in the ligand

University of BRISTOL

BUDE: Bristol University Docking Engine



Typical docking scoring functions	Empirical Free Energy Forcefield **BUDE**	Free Energy calculations MM[1,2] QM/MM[3]

Entropy:			
solvation	**No**	**Yes**	**Yes**
configurational	**Approx**	**Approx**	**Yes**
Electrostatics	**?**	**Approx**	**Yes**
All atom	**No**	**Yes**	**Yes**
Explicit solvent	**No**	**No**	**Yes**

University of BRISTOL

1. MD Tyka, AR Clarke, RB Sessions, J. Phys. Chem. B 110 17212-20 (2006)
2. MD Tyka, RB Sessions, AR Clarke, J. Phys. Chem. B 111 9571-80 (2007)
3. CJ Woods, FR Manby, AJ Mulholland, J. Chem. Phys. 128 014109 (2008)

Empirical Free Energy Function (atom-atom)

$$\Delta G_{\text{ligand binding}} = \sum_{i=1}^{N_{\text{protein}}} \sum_{j=1}^{N_{\text{ligand}}} f(x_i, x_j)$$

Parameterised using experimental data†

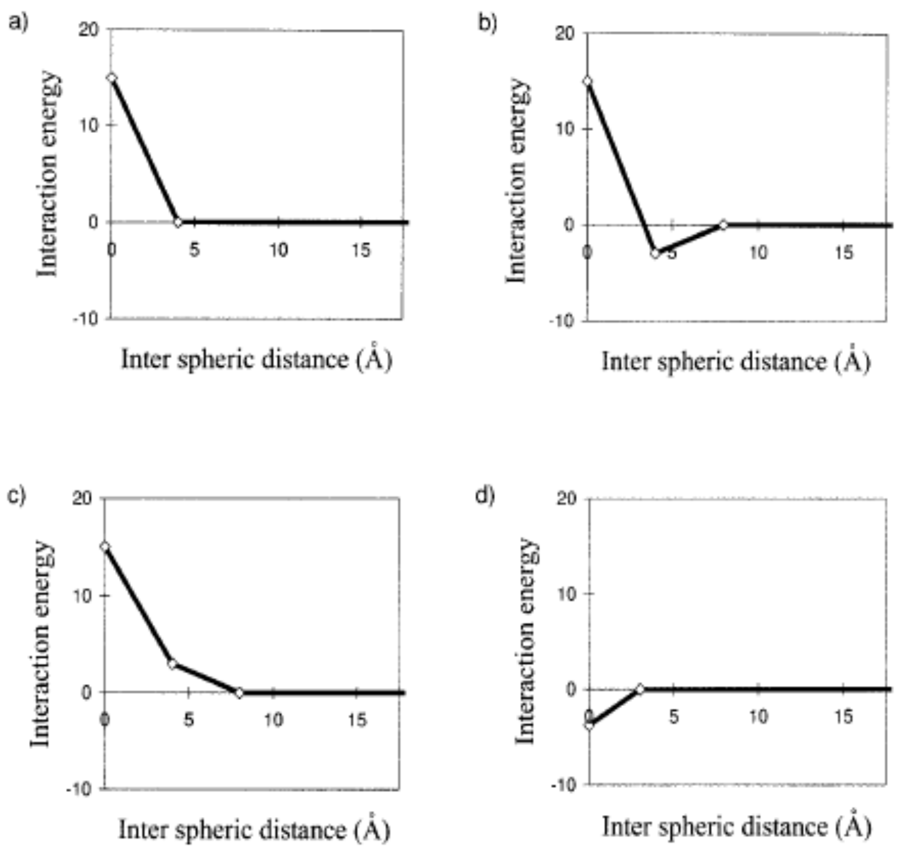
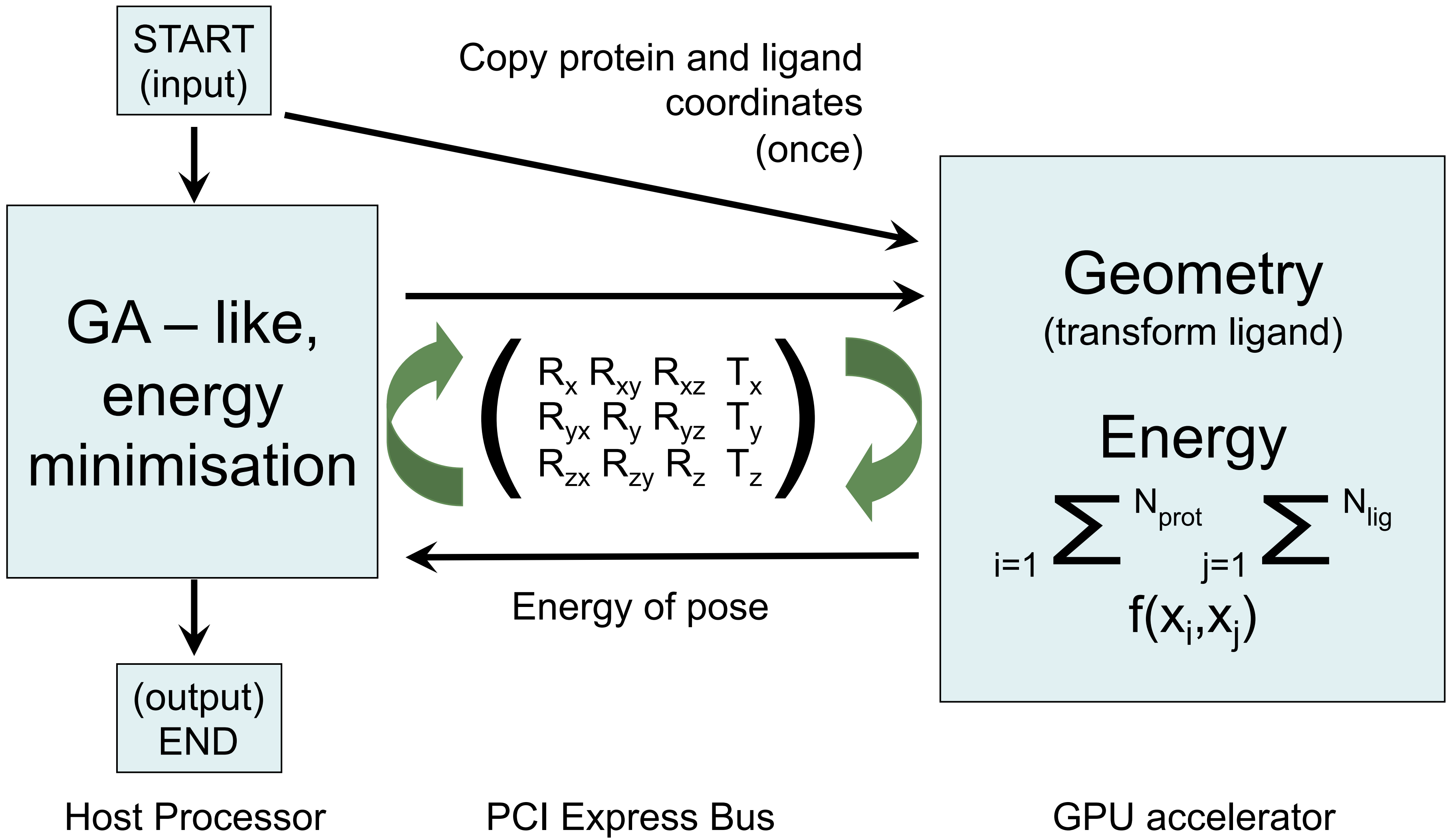


Fig. 1. Inter-residue sphere-sphere interaction energy functions of the force field. a: Between two polar spheres, or between a backbone sphere and any other non hydrogen-bonding sphere. b: Between two non-polar spheres. c: Between a non-polar sphere and a polar sphere. d: Between a hydrogen bond donor sphere and a hydrogen bond acceptor sphere.

University of BRISTOL

† N. Gibbs, A.R. Clarke & R.B. Sessions, "Ab-initio Protein Folding using Physicochemical Potentials and a Simplified Off-Lattice Model", Proteins 43:186-202,2001

BUDE Acceleration with OpenCL



Systems benchmarked

High-end:

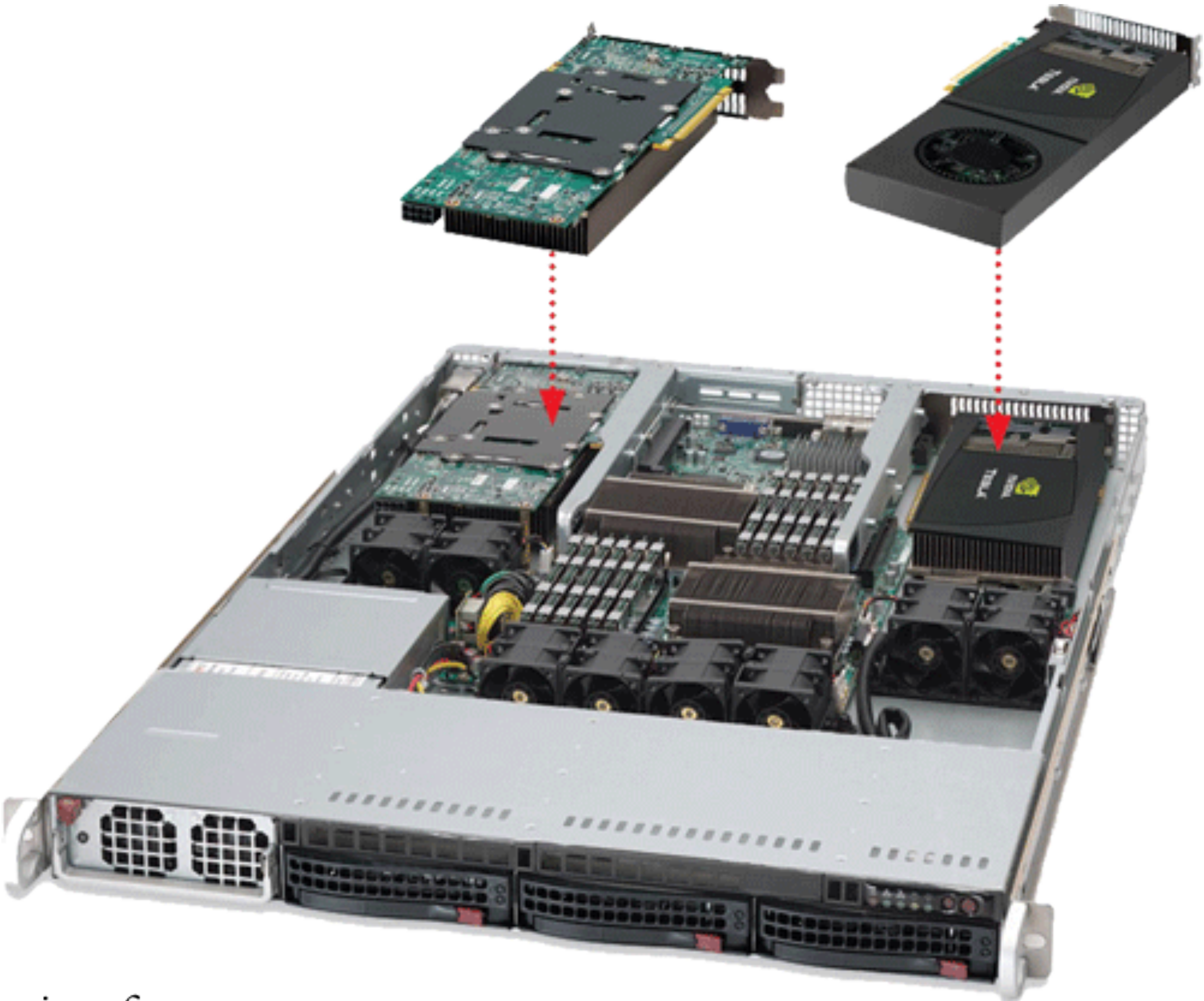
- Supermicro 1U dual GPU server
- Two Intel 5500 series 2.4 GHz Xeon ‘Nehalem’ quad-core processors
- 24 GBytes of DRAM
- Two Nvidia C2050 ‘Fermi’ GPUs ***or***
- Two AMD ‘Cypress’ FirePro V7800s

Medium-end:

- Workstation with 1 CPU & 1 GPU
- Intel E8500 3.16 GHz dual core CPU
- 4 GBytes of DRAM
- Previous generation Nvidia consumer-level GPU, the GTX280

University of BRISTOL

Supermicro GPU server



University of BRISTOL

Systems benchmarked

Middle-end:

- Workstation based on a 3-core AMD 2.8 GHz Phenom II X3 720
- 4 GBytes of DRAM
- **No GPU!**

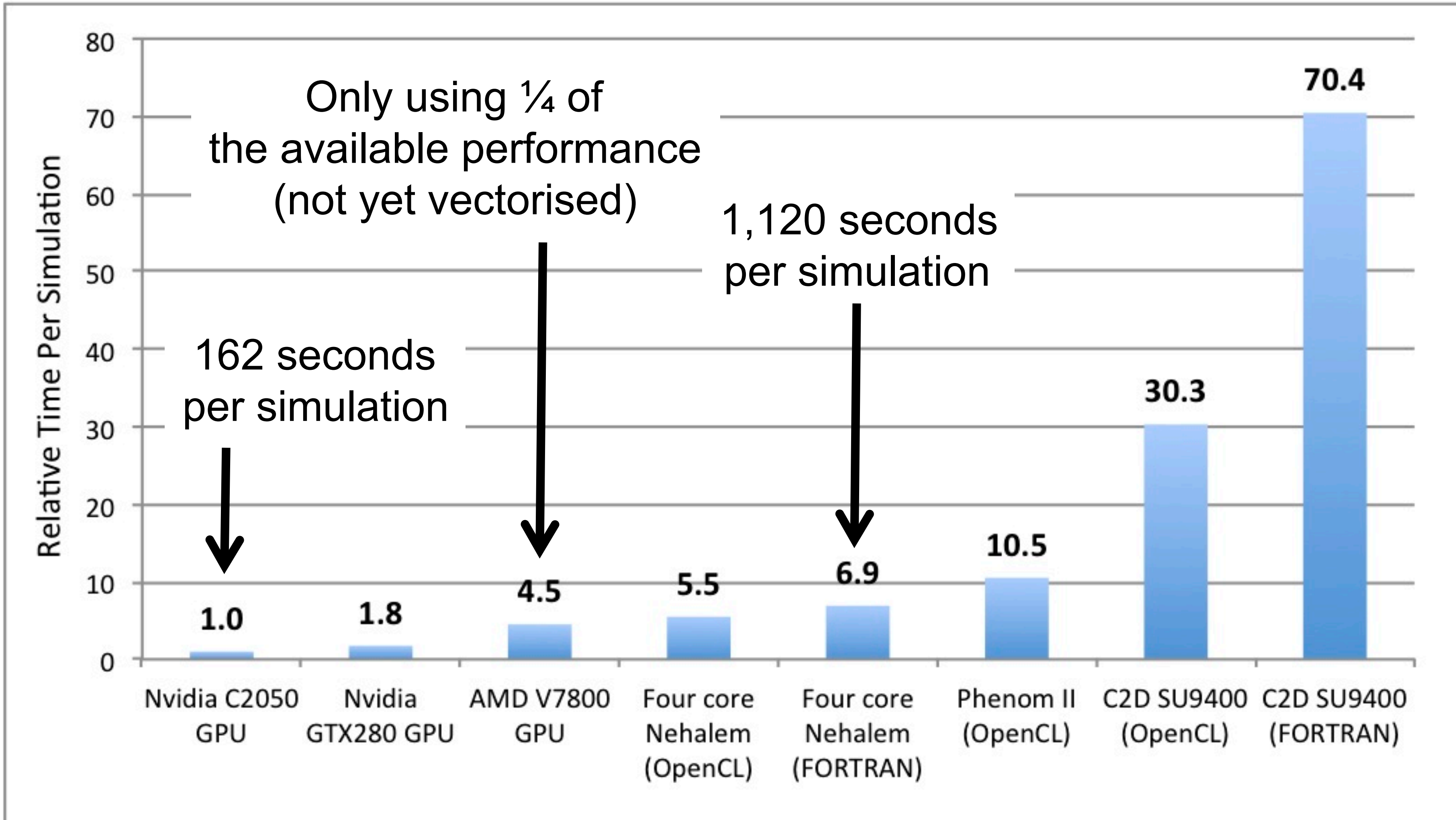
Low-end:

- Laptop based on an Intel Core2Duo SU9400 ‘Penryn’ 1.4 GHz CPU
- 4 GBytes of DRAM
- **No GPU!**

Benchmarking methodology

- Use the same power measurement equipment for all the systems under test
- Watts Up? Pro meter
- +/- 1.5% accuracy
- Measures complete system power 'at the wall'
- User-definable sampling rate
- Using a real problem with BUDE
- Run as fast as possible on all available resources (i.e. all cores or all GPUs simultaneously)
- Removed GPUs from the systems when benchmarking host performance

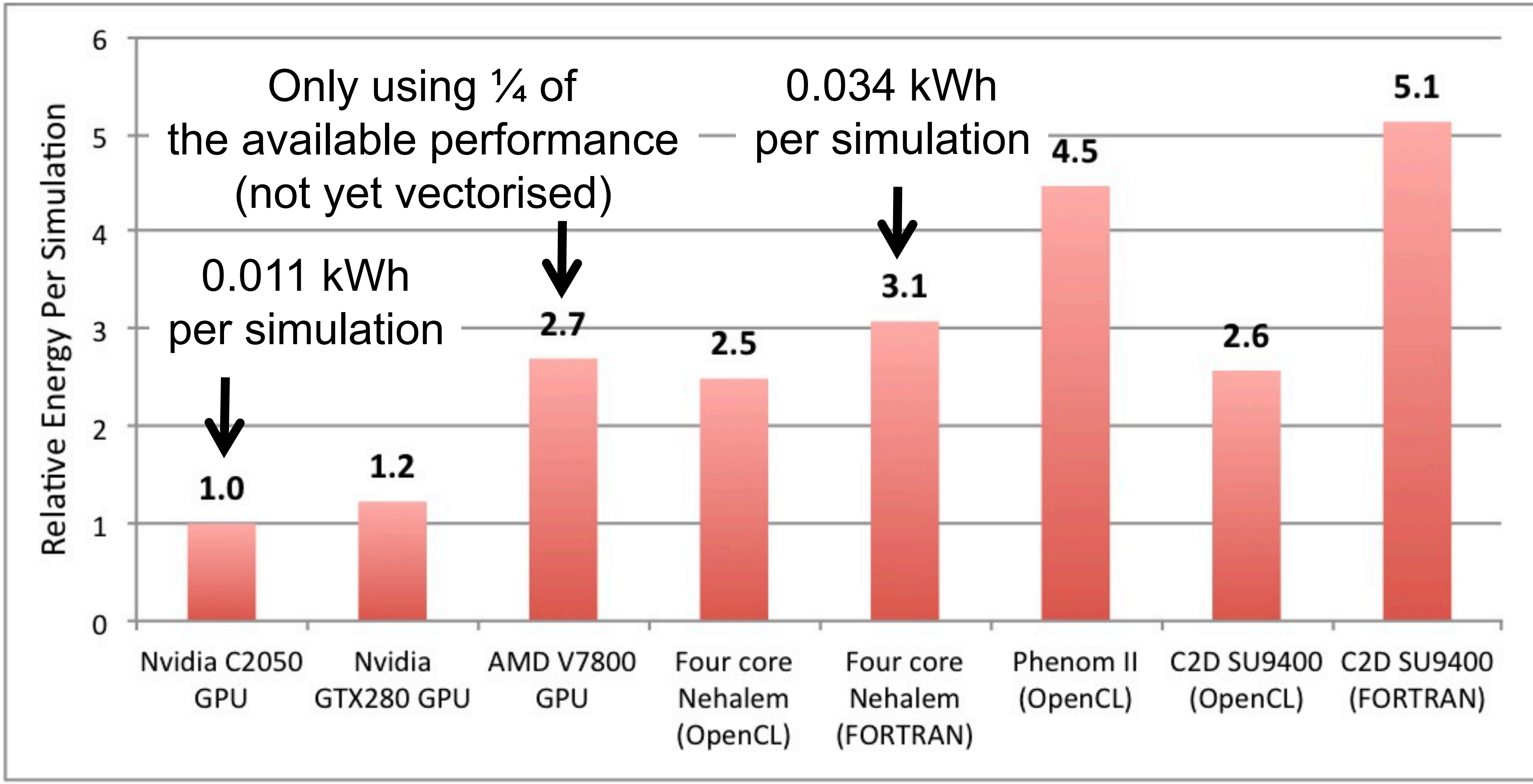
Relative performance



University of BRISTOL

Less than 2 days to screen a library of 1 million drug candidates on 1000 GPUs

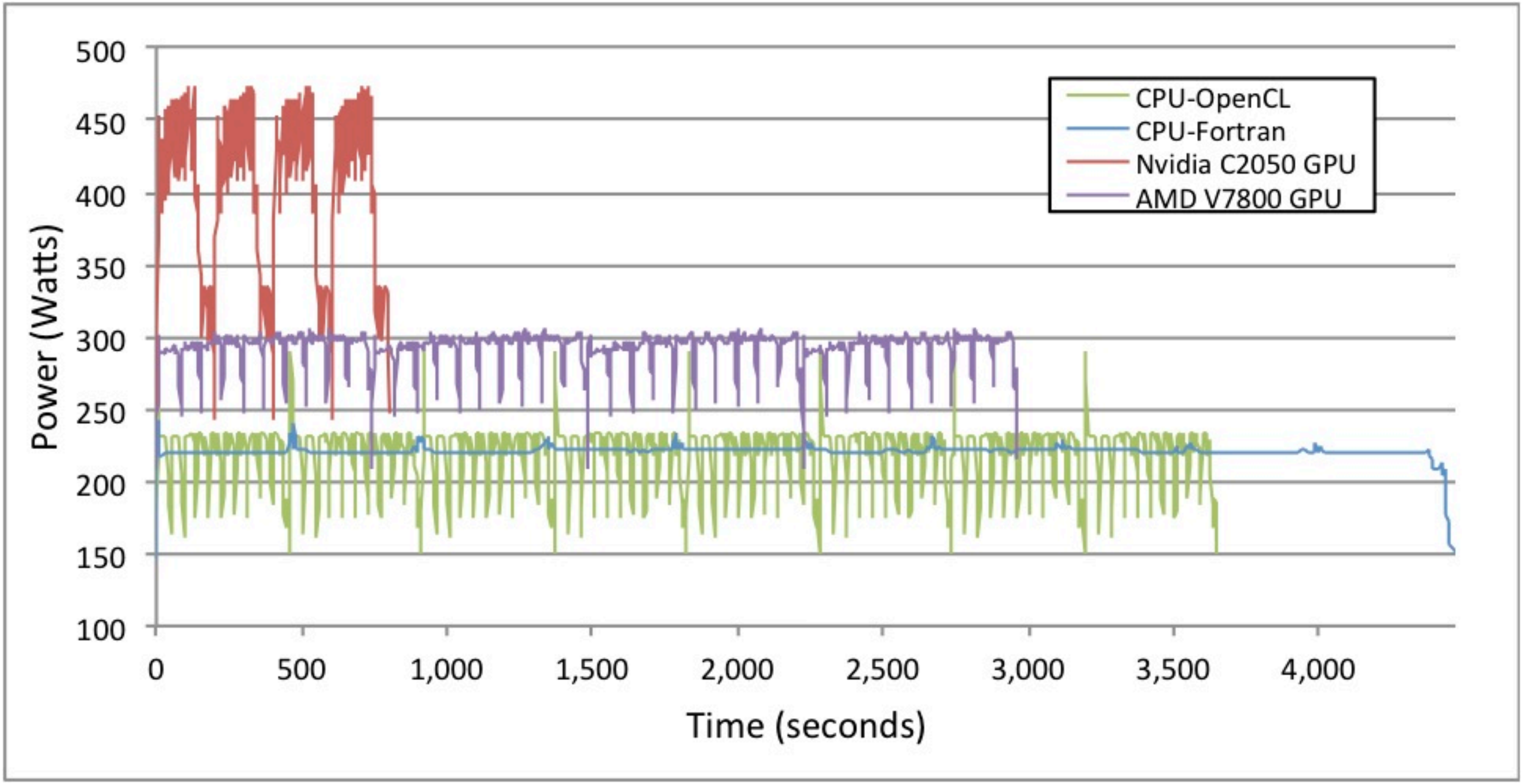
Relative energy efficiency



0.011 kWh = 0.16 pence per simulation
1 million simulations → £1,600 on energy for one experiment

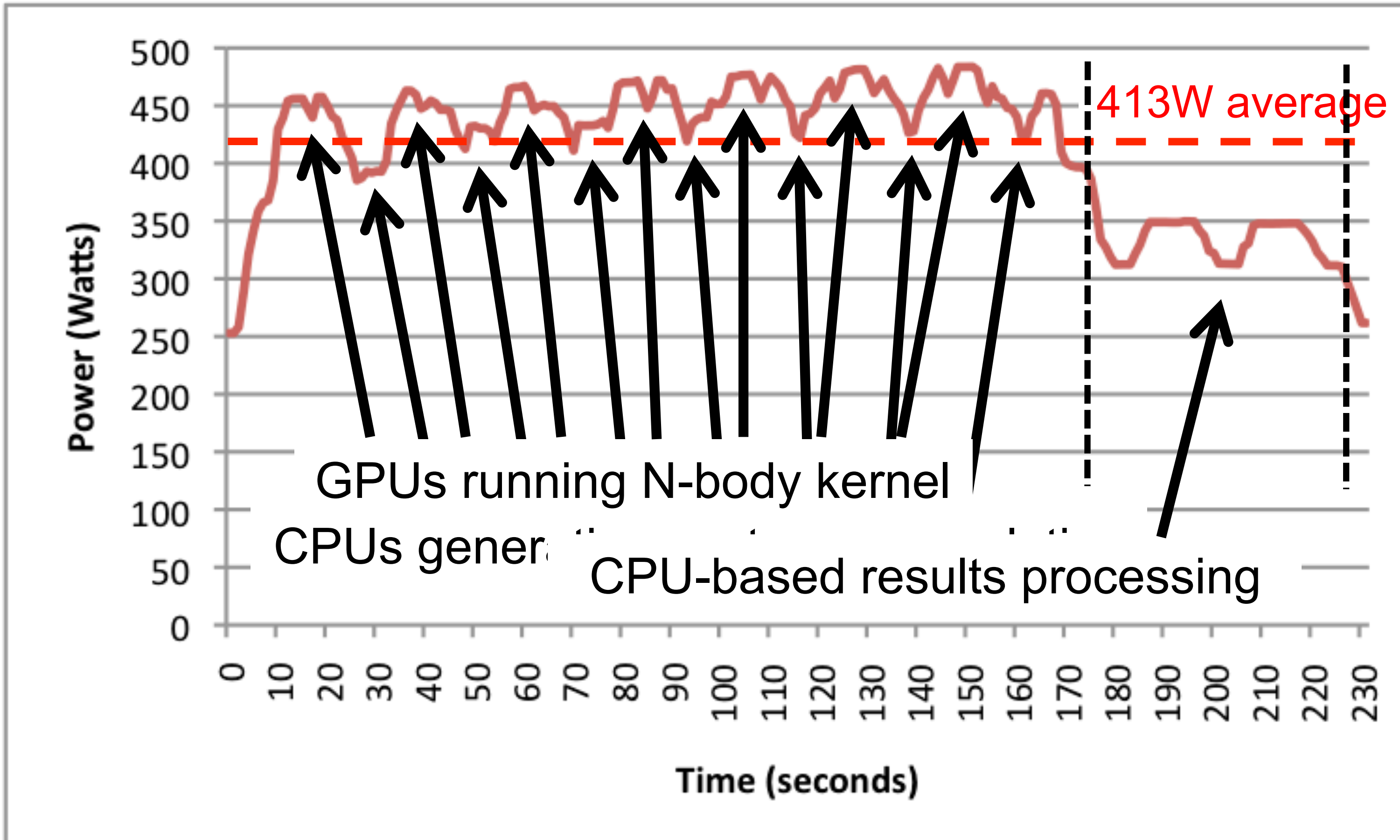
University of BRISTOL

Power consumption profiles



Time to complete 8 simulations using all resources simultaneously

Dual C2050 energy profile



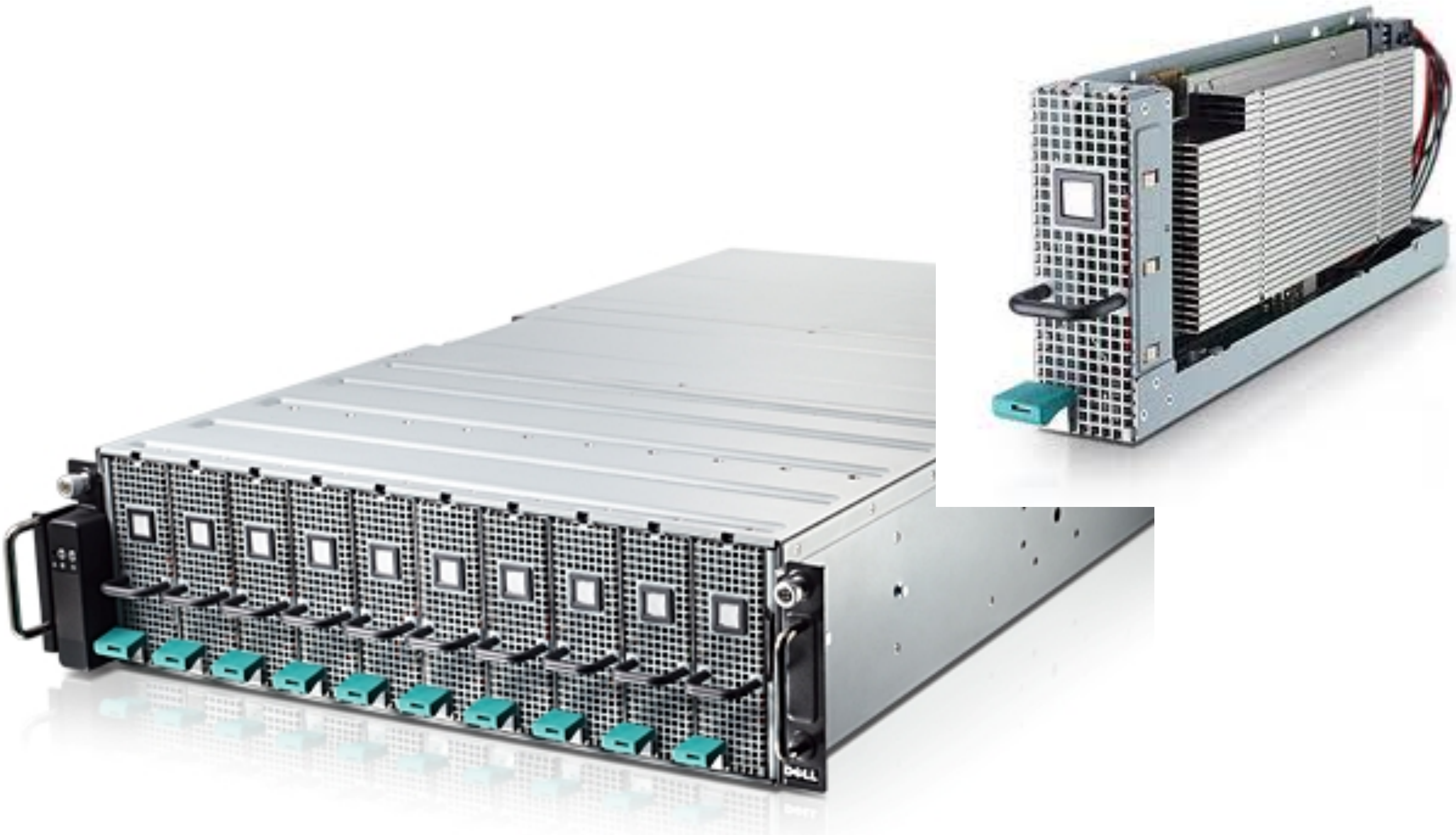
Conclusions

- Energy efficiency will eventually become ***the*** first order consideration driving performance
- Possible to measure metrics for per simulation $$$
- Hard to accurately *compare* energy consumption
- GPUs can lead to big increases in performance per watt, not just performance
- OpenCL can work just as well for multi-core CPUs

It's possible to screen libraries of millions of molecules against complex targets using highly accurate methods in a weekend using 10 racks costing < £2M

University of BRISTOL

I want to try one of these…



University of BRISTOL

References

- S. McIntosh-Smith, T. Wilson, A. A. Ibarra, R. B. Sessions, and J. Crisp. *Benchmarking energy efficiency, power costs and carbon benchmarking energy efficiency, power costs and carbon emissions on heterogeneous systems.* The Computer Journal, 2011 (to appear).

- S. McIntosh-Smith, T. Wilson, J. Crisp, A. A. Ibarra, and R. B. Sessions. *Energy-aware metrics for benchmarking heterogeneous systems.* SIGMETRICS Perform. Eval. Rev., 38:88–94, March 2011.

University of BRISTOL