Hartree Centre
Science & Technology Facilities Council

HPC in the UK: a changing landscape?

Iain Bethune,
Technical Programme Manager,
STFC Hartree Centre

iain.bethune@stfc.ac.uk
@ibethune

Who are we?

STFC's High Performance Computing, data analytics and cognitive technology centre

Provides businesses and researchers with access to powerful technologies, facilities and scientific computing expertise

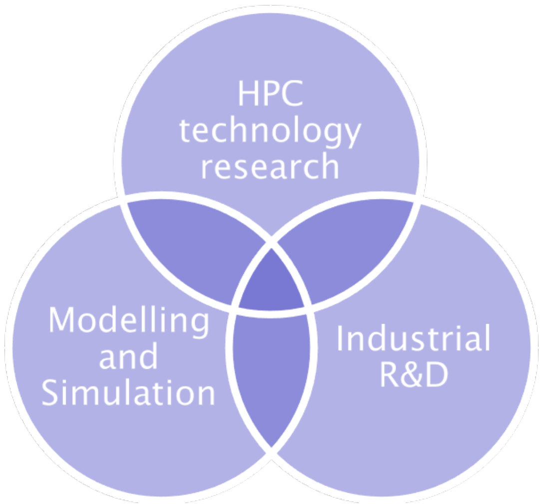
Hartree Centre

Science & Technology Facilities Council

What can we do?

- Hartree Centre builds on STFC's long heritage and expertise in High Performance Computing

- 2012 focus on economic impact through software & modelling

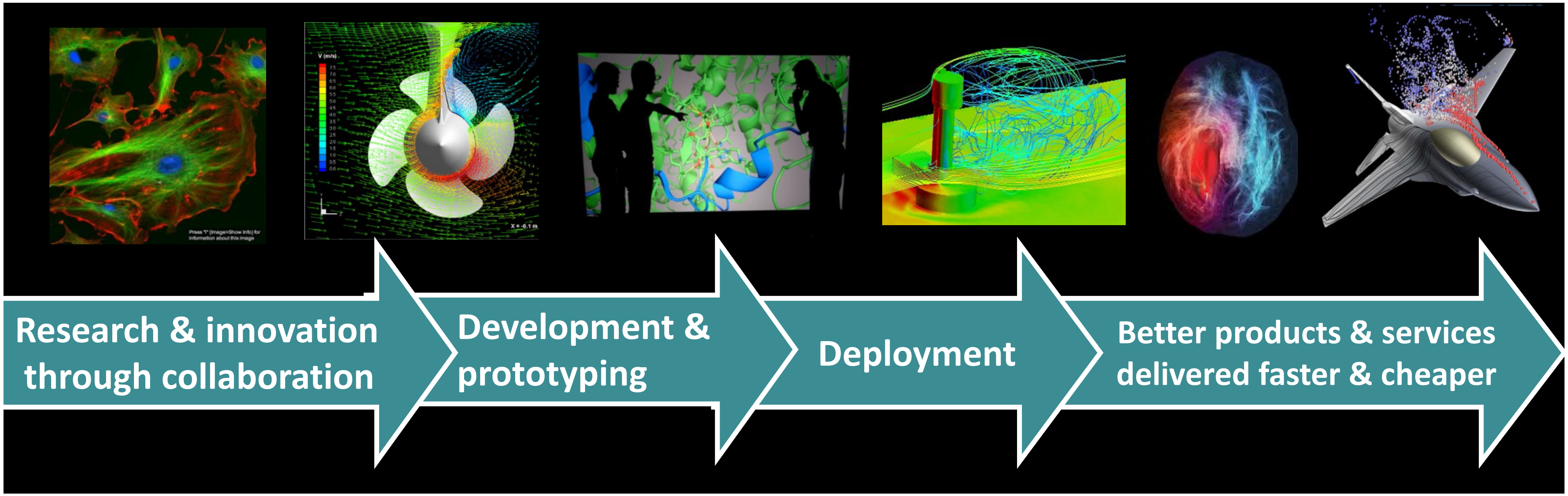


- 2015 major investment in collaborative research
  - Additional focus on data centric and cognitive computing
  - Embedded IBM Research Centre
  - Extended industrial & scientific reach

- Grown to ~50 STFC + 20 IBM staff

Our mission

To transform UK competitiveness by facilitating and accelerating the adoption of high performance computing, data-centric computing and cognitive technologies.



Project examples

- Engineering & Manufacturing
  - Vehicle Design & Testing
  - Consumer Electronics Design
  - Consumer Packaged Goods Products and Packaging
- Environment
  - Weather modeling
- Life Sciences
  - Genomics for better crop yields
  - Disease mapping
- Energy
  - Advanced Battery Cell Design
  - Efficient Well Head Oil extraction
- Financial Services
  - Risk Management
  - Service Modelling
- Transport
  - Network simulation

Converting physical experiments to a simulation workflow

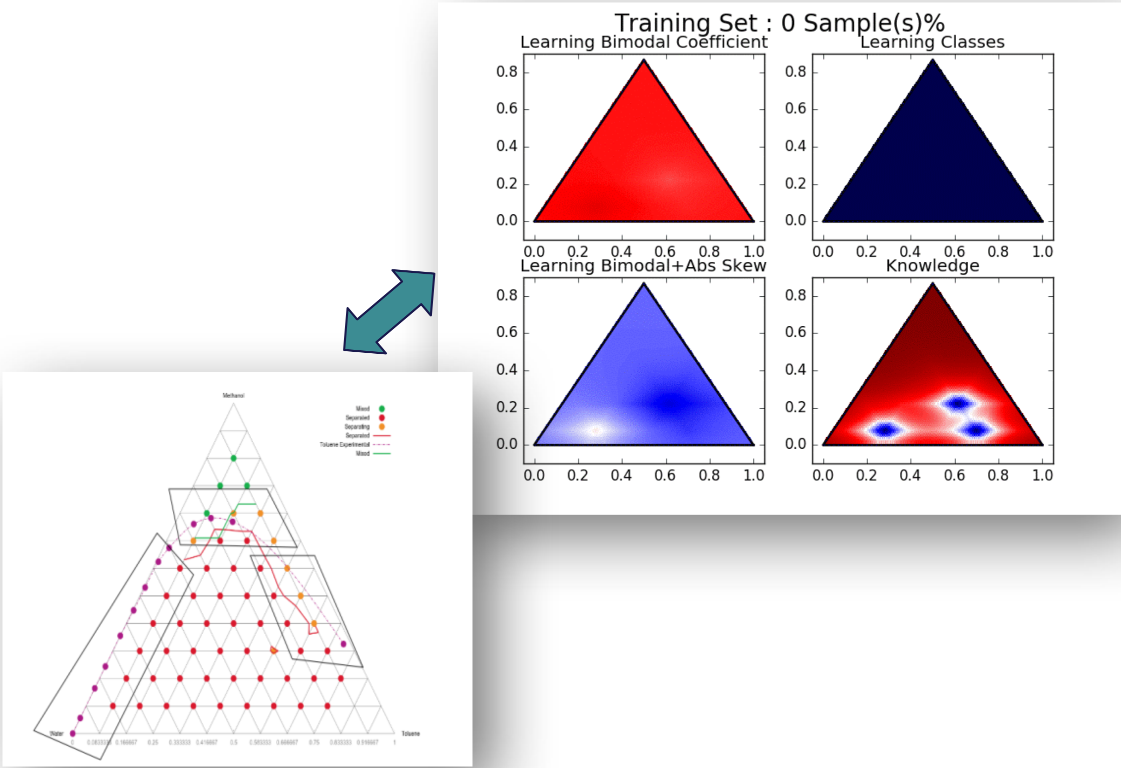
Reducing the innovation cycle and time to market

Putting the tools in the hands of the formulation chemist

Unilever – Computer Aided Formulation

Getting HPC to "work smart, not hard"

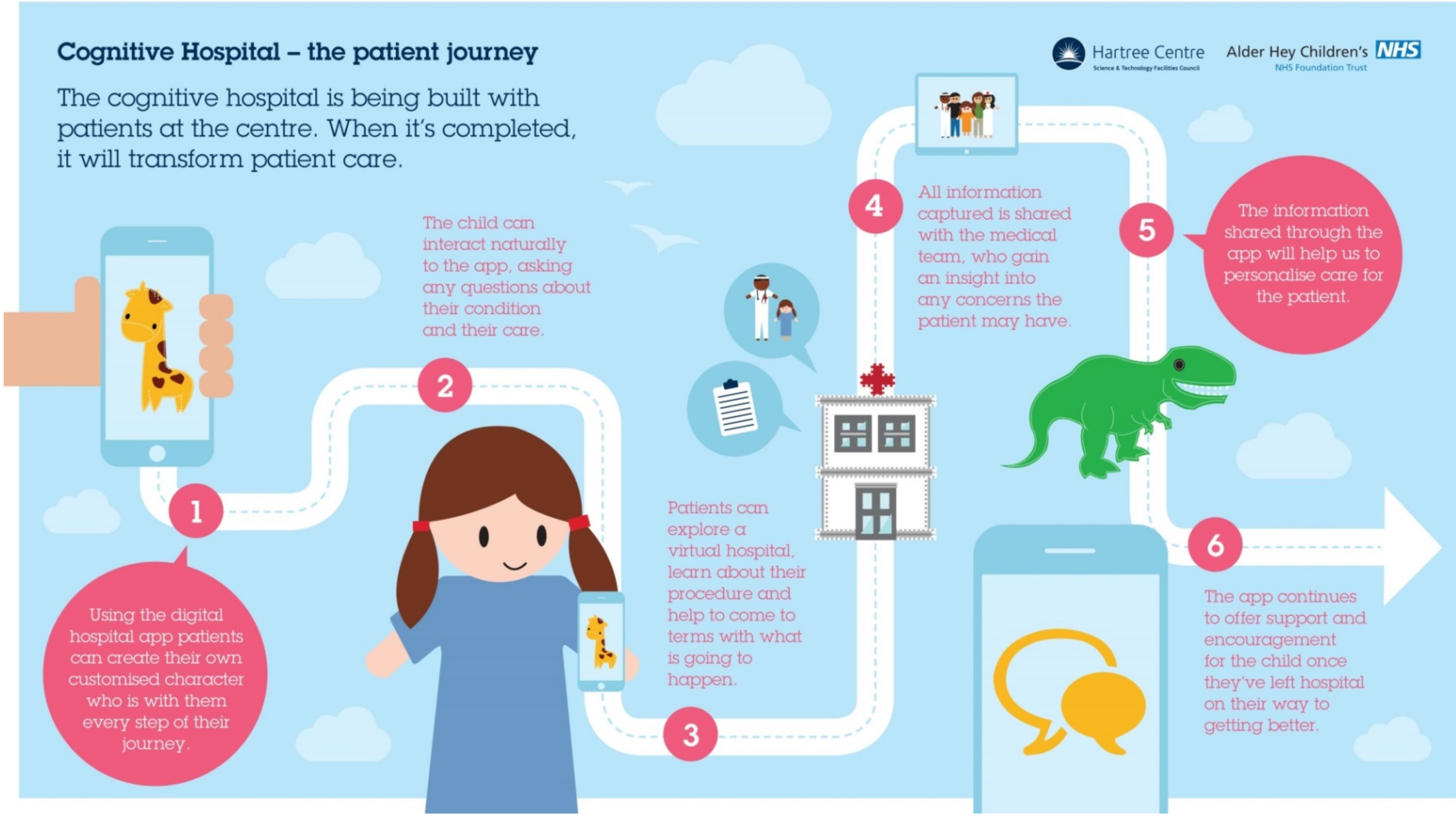
- Typically HPC development is focused on increased speed.
  - <u>The fastest calculation is the one which you don't run!</u>
- Can we use machine learning to make better decisions on **which simulations give the most value**?
- Can we use machine learning to improve resolution of information?



'Cognitive' workflow uses 1/3 of the calculations to achieve 4 orders of magnitude resolution increase

With our help, Alder Hey Children's Hospital is harnessing IBM Watson to enhance the patient experience.

Hartree Centre
Science & Technology Facilities Council


Cognitive Hospital – the patient journey
The cognitive hospital is being built with patients at the centre. When it's completed, it will transform patient care.
Hartree Centre
Alder Hey Children's NHS
NHS Foundation Trust
1
Using the digital hospital app patients can create their own customised character who is with them every step of their journey.
2
The child can interact naturally to the app, asking any questions about their condition and their care.
3
Patients can explore a virtual hospital, learn about their procedure and help to come to terms with what is going to happen.
4
All information captured is shared with the medical team, who gain an insight into any concerns the patient may have.
5
The information shared through the app will help us to personalise care for the patient.
6
The app continues to offer support and encouragement for the child once they've left hospital on their way to getting better.

Our machines

Intel platforms

Bull Sequana X1000 | 846 Xeon nodes + 840 KNL

Lenovo NeXtScale | 8,192 cores

Lenovo System x iDataPlex system | 2048 cores

Intel Xeon Phi | Knight's Corner

IBM big data analytics cluster | 288TB

IBM data centric platforms

IBM Power8 + NVLink + Tesla P100

IBM Power8 + Nvidia K80

Accelerated & emerging tech

Maxeler FPGA system

ARM 64-bit platform

Clustervision novel cooling demonstrator

Academic HPC platform

JADE NVIDIA DGX-1 Deep Learning System

New partnership with Atos Bull

General purpose HPC system to be called “Scafell Pike”

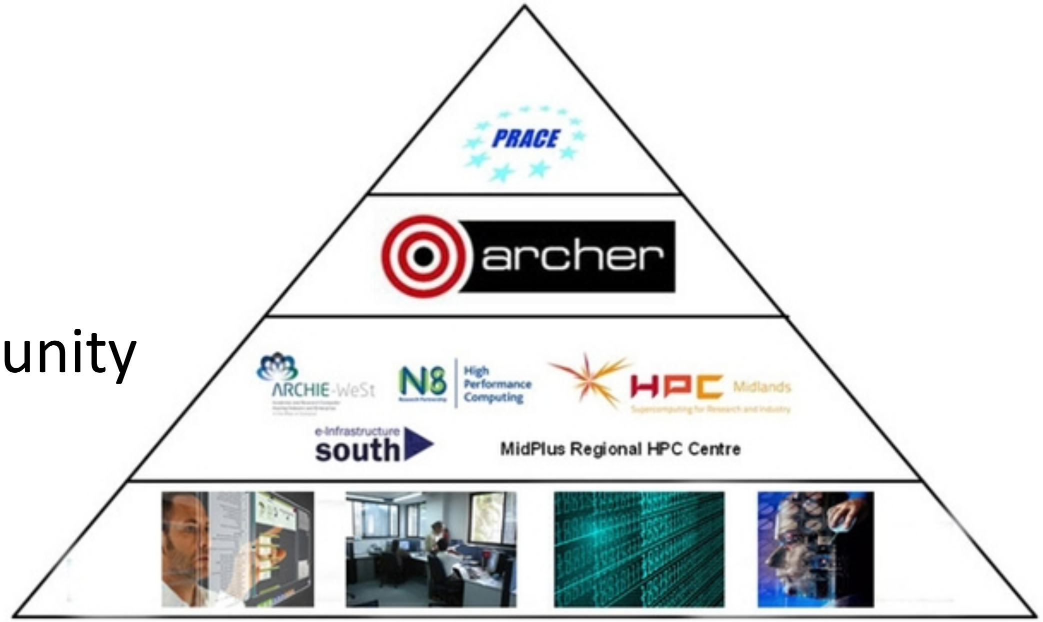
First Bull Sequana system in UK

One of the largest supercomputers in Europe (3.4 PFLOP/s estimated) and largest focusing primarily on industrial-led challenges

Academic HPC in the UK

National HPC Service(s)?

- Tier-0: International
- Tier-1: National
- Tier-2: Regional / Community
- Tier-3: Institutional/
  Departmental / Group



Hartree Centre
Science & Technology Facilities Council

2011

C/C++/Fortran + MPI is enough

10–100 cores for Tier–2
100–1000 cores for Tier–1

- PRACE Tier-0
  - JUGENE – 0.2 PF BlueGene/P
  - CURIE – 1.7 PF Intel Xeon
  - HERMIT – 1 PF AMD Opteron

- EPSRC Tier-1 (£118M)
  - HECToR – 0.8 PF AMD Opteron
- EPSRC Tier-2 (£10M)
  - HPC Midlands, Mid Plus, H8, ARCHIE-WeSt
    - 0.25 PF total, Intel Xeon

2017

- PRACE Tier-0
  - CURIE – 1.7 PF Intel Xeon Sandybridge
    - 9 PFLOP/s Skylake + KNL in mid-2018
  - MARCONI – 13 PF Intel Xeon Broadwell + KNL
  - Hazel Hen – 7.4 PF Intel Xeon Haswell
  - JUQUEEN – 5.9 PF BlueGene/Q Power A2
  - MareNostrum – 11 PF Intel Xeon
    - IBM Power + NVIDIA Volta, KNH, ARMv8 coming
  - Piz Daint – 25 PF Intel Xeon + NVIDIA Tesla P100
  - SuperMUC – 7.7 PF Intel Xeon Westmere/Haswell

24x compute in 6 years!

Hartree Centre
Science & Technology Facilities Council

2017

- EPSRC/NERC Tier-1 (£43M)
  - ARCHER – 2.6 PF Intel Xeon Ivy Bridge
    - 12 node KNL partition

- STFC Tier-1
  - DIRAC – 1.3 PF BlueGene/Q
    - Also data analytics services

3x compute!

< half price!

2017

- EPSRC Tier-2 (£20M investment 2016)
  - CSD3
    - 1.0 PF Intel Xeon Skylake, 0.5 PF Intel KNL, 1.2 PF NVIDIA P100
  - Thomas
    - 0.5 PF Intel Xeon Broadwell
  - JADE
    - 3.7 PF NVIDIA DGX-1 (Intel Xeon + Pascal)
  - HPC Midlands Plus
    - 0.5 PF Intel Xeon Broadwell + IBM Power8
  - Isambard GW4
    - ARMv8, x86, Xeon Phi KNL, NVIDIA Pascal
  - Cirrus
    - 0.3 PF Intel Xeon Broadwell

24x compute!

KNL, GPUs, ARM, Power8…

Hull Viper = ~0.2 PF!

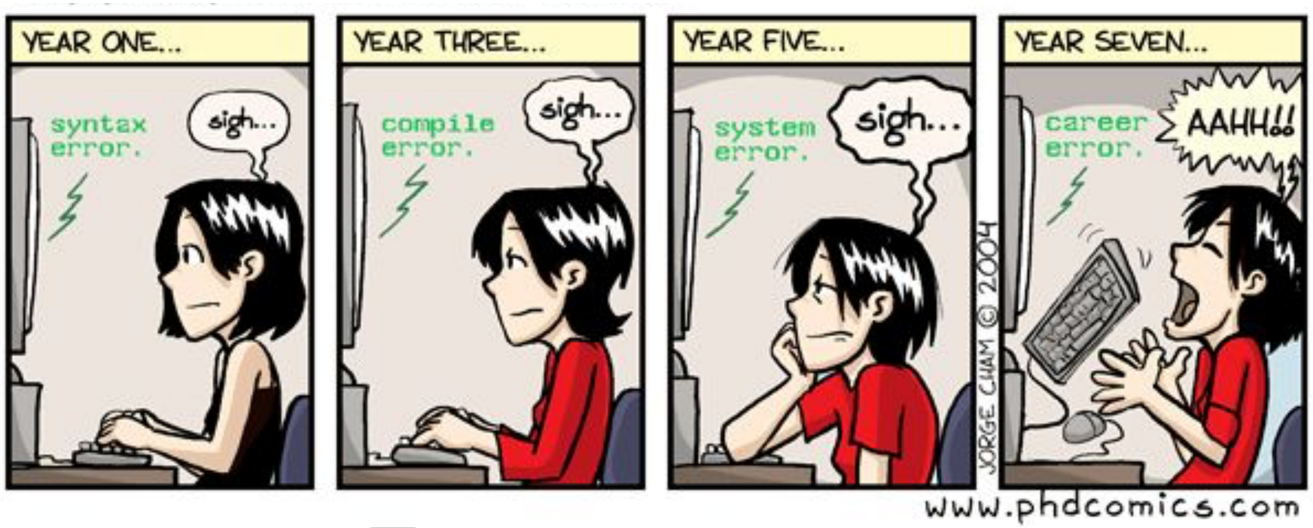
2017: Architecture Diversity

- Performance ↑ 🙂
- Cost / performance ↓ 😀
- Complexity & parallelism ↑ 😟
- Performance portability ↓ 😡

C/C++, Fortran, Python, R…
MPI+OpenMP/CUDA/OpenCL,
SIMD, NUMA…

Threads per node:10–10,000s

- Result: researchers need to spend more time writing, porting, maintaining code than doing research!



Hartree Centre
Science & Technology Facilities Council

Research Software Engineering

- HPC Centres have the expertise, but mainly focussed on Tier-1
  - Need skilled people, embedded in research groups / institutions
  - With up-to-date skills
  - Science literate
  - More than ‘just’ software engineers
  - With a recognised career path to drive excellence
- Research Software Engineer
  - First coined in 2012
  - Supported by Software Sustainability Institute
  - Now UK RSE association, EPSRC support...

Hartree Centre
Science & Technology Facilities Council

Research Software Engineering

- ~20 RSE posts associated with new Tier-2 sites
- RSE Groups springing up around the UK
  - UCL, Cambridge, Bristol, …
  - Many more posts in individual groups

We're hiring too!

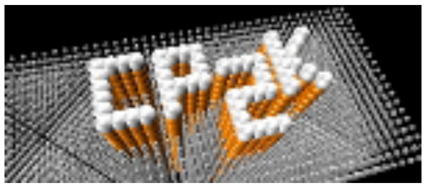
- PIs starting to see the value of including RSE support in grants
- Universities creating career pathways
- Growing number of RSEs
  - Skills development
  - Best practice / knowledge sharing
- Turning software into Impact

Virtuous circle

Hartree Centre
Science & Technology Facilities Council

Case study: CP2K

"CP2K is a program to perform atomistic and molecular simulations of solid state, liquid, molecular, and biological systems. It provides a general framework for different methods such as e.g., density functional theory (DFT) using a mixed Gaussian and plane waves approach (GPW) and classical pair and many-body potentials."



From www.cp2k.org (and original home page from 2004!)

- Open Source
  - GPL, Sourceforge SVN & Github
  - 1M LOC, ~2 commits per day
  - 10-20 core developers

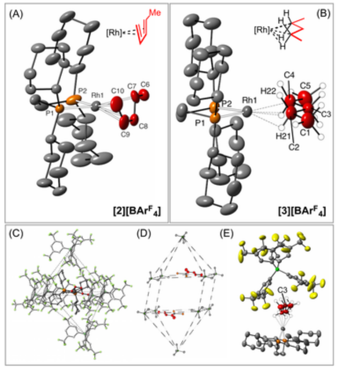
CP2K Applications

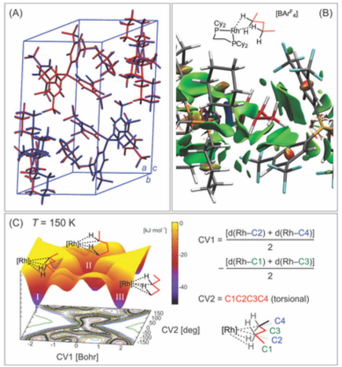
Alkane Complexes

International Edition: DOI: 10.1002/anie.201511269
German Edition: DOI: 10.1002/ange.201511269

A Rhodium–Pentane Sigma-Alkane Complex: Characterization in the Solid State by Experimental and Computational Techniques

F. Mark Chadwick⁺, Nicholas H. Rees, Andrew S. Weller, Tobias Krämer⁺, Marcella Iannuzzi, and Stuart A. Macgregor**



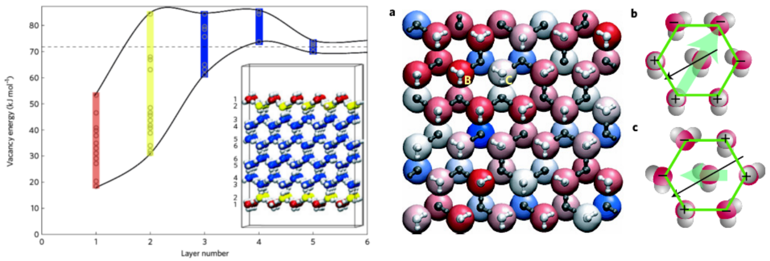


Large variation of vacancy formation energies in the surface of crystalline ice

M. Watkins[1,2,3], D. Pan[4], E. G. Wang[5], A. Michaelides[1,2,3], J. VandeVondele[6] and B. Slater[1,3]*

[1]Department of Chemistry, Christopher Ingold Building, 20 Gordon Street, University College London, London WC1H 0AJ, UK, [2]London Centre for Nanotechnology, University College London, London WC1H 0AJ, UK, [3]TYC@UCL, University College London, London WC1H 0AJ, UK, [4]Institute of Physics, Chinese Academy of Sciences, PO Box 603, Beijing 100190, China, [5]School of Physics, Peking University, Beijing 100871, China, [6]Institute of Physical Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. *e-mail: b.slater@ucl.ac.uk.

NATURE MATERIALS | VOL 10 | OCTOBER 2011



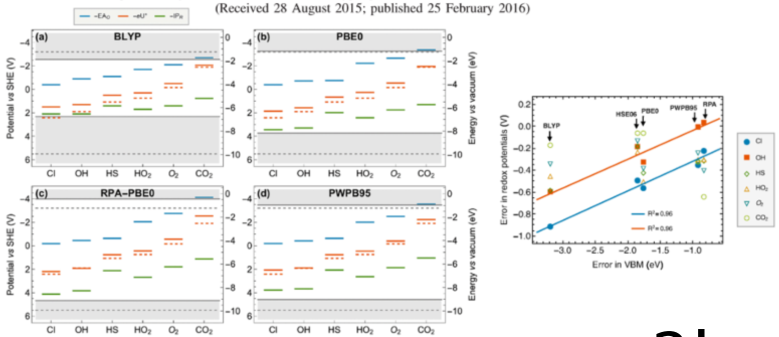
PRL 116, 086402 (2016) PHYSICAL REVIEW LETTERS week ending 26 FEBRUARY 2016

Calculation of Electrochemical Energy Levels in Water Using the Random Phase Approximation and a Double Hybrid Functional

Jun Cheng*
Collaborative Innovation Center of Chemistry for Energy Materials, State Key Laboratory of Physical Chemistry of Solid Surfaces, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, People's Republic of China and Department of Chemistry, University of Aberdeen, Aberdeen AB24 3UE, United Kingdom

Joost VandeVondele†
Department of Materials, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland

(Received 28 August 2015; published 25 February 2016)



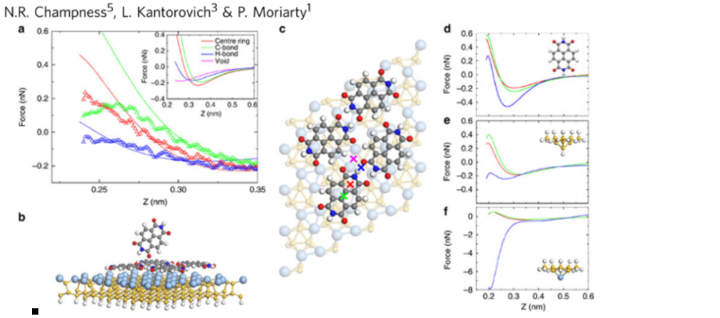
nature COMMUNICATIONS

ARTICLE

Received 6 Dec 2013 | Accepted 22 Apr 2014 | Published 30 May 2014 DOI: 10.1038/ncomms4931 OPEN

Mapping the force field of a hydrogen-bonded assembly

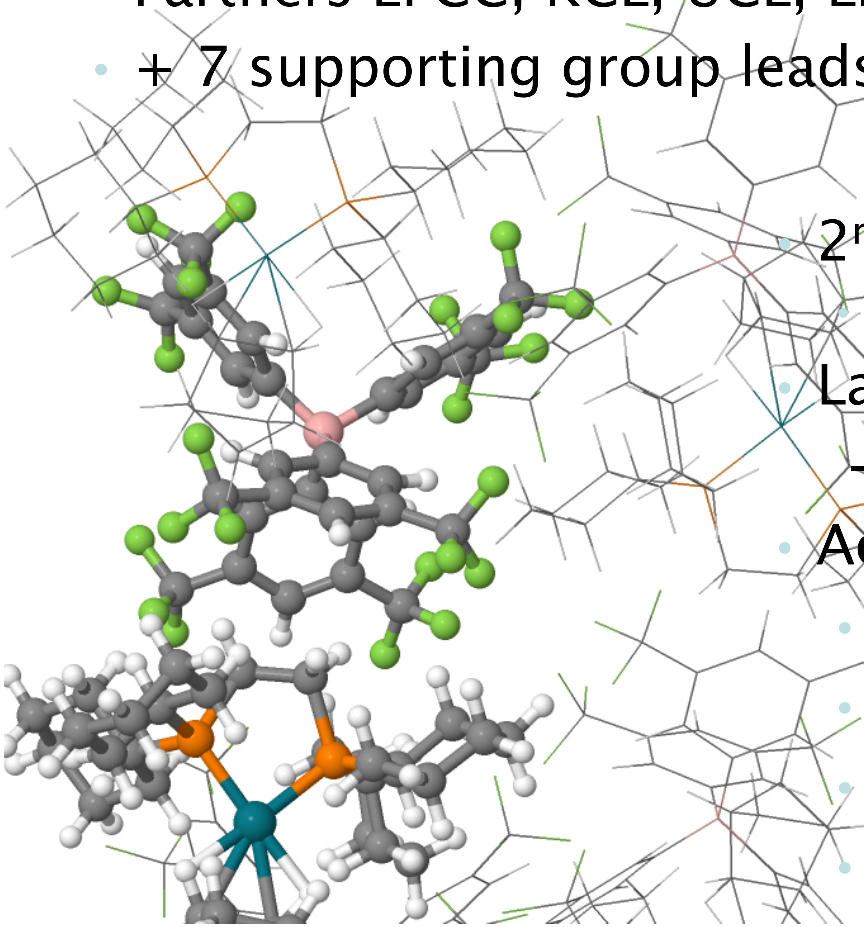
A.M. Sweetman[1,*], S.P. Jarvis[1,*], Hongqian Sang[2,3,*], I. Lekkas[1], P. Rahe[4], Yu Wang[2], Jianbo Wang[2], N.R. Champness[5], L. Kantorovich[3] & P. Moriarty[1]



www.cp2k.org/science

CP2K-UK

- 2013-2018 EPSRC Software for the Future
- Led by Hartree Centre
- Partners EPCC, KCL, UCL, Lincoln
- + 7 supporting group leads



- 2rd ranked code on ARCHER
  - Growing usage and interest
- Large feature set + excellent performance
  - -> complexity hump for new users / devs
- Activities:
  - User group meetings (annual)
  - Usability, New Functionality
  - Training
  - Spin-off projects

CP2K Overview

- QUICKSTEP DFT: Gaussian and Plane Waves Method
  (VandeVondele *et al*, Comp. Phys. Comm., 2005)
  - Advantages of atom-centred basis (primary)
    - Density, Overlap, KS matrices are sparse

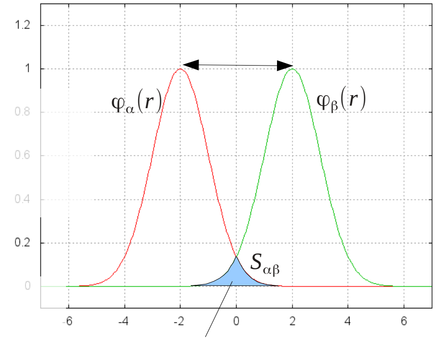
Sparse Matrices!

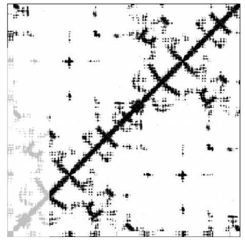
  - Advantages of plane-wave basis (auxiliary)
    - Efficient computation of Hartree potential

Distributed 3D multigrids!

  - Efficient mapping between basis sets
    - -> Construction of the KS Matrix is ~O(n)

3D FFT! Gaussian collocation / integration!





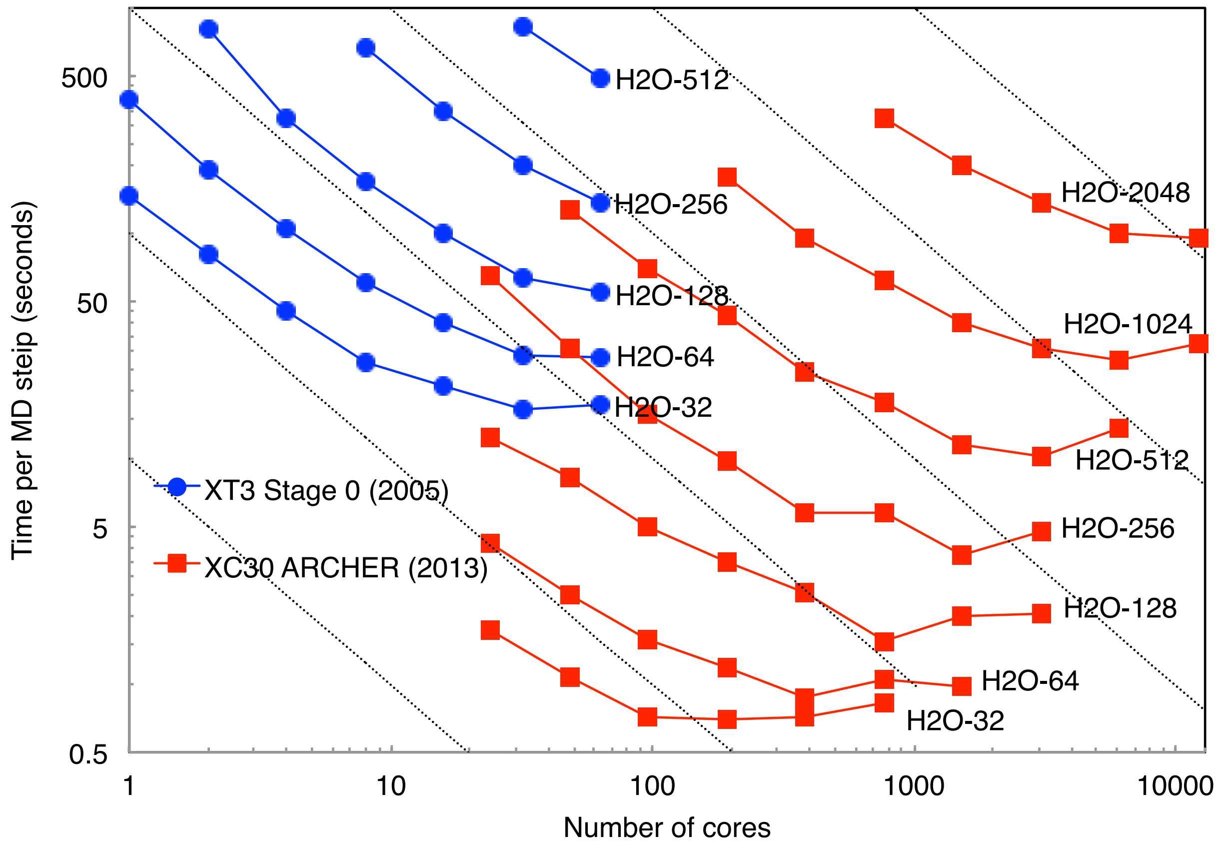
- Orbital Transformation Method (VandeVondele & Hutter, J. Chem. Phys., 2003)
  - Replacement for traditional diagonalisation/density mixing(non-metallic systems only)
    - » Cubic scaling but ~10% cost

Dense Matrices! (ScaLAPACK / ELPA)

Hartree Centre
Science & Technology Facilities Council

Water benchmarks

- *ab initio* MD of various sizes of water boxes
- Production quality settings
- 84x single node speed-up in 8 years!
- Scaling and peak perf up 10-20x



Ref: "*CP2K Performance from Cray XT3 to XC30*", **IB** et al, Proceedings of Cray User Group 2014

Hartree Centre
Science & Technology Facilities Council

Algorithm development for CPUs

- MPI Load balancing, communication optimisation (2008/9)
- OpenMP parallelism (2009/10)
  - 3D FFT, grid operations, matrix/matrix multiplication
- Optimised small block matrix multiplications (2011/12)
- Memory-efficient algorithms (2015)

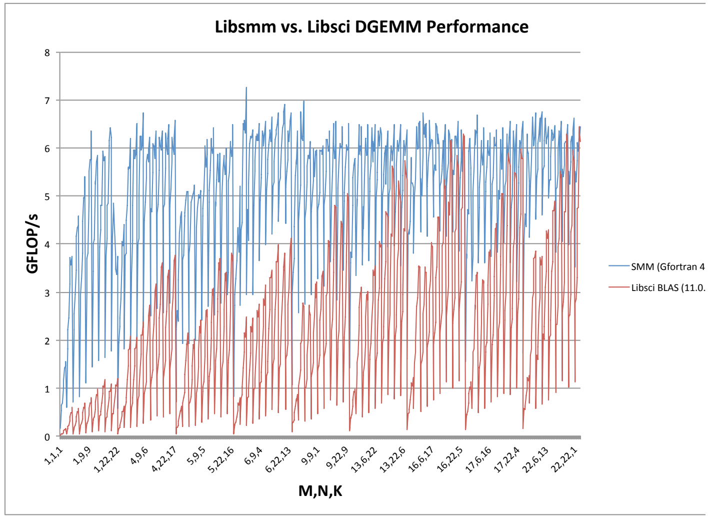


Figure 5: Comparing performance of SMM and Libsci BLAS for block sizes up to 22,22

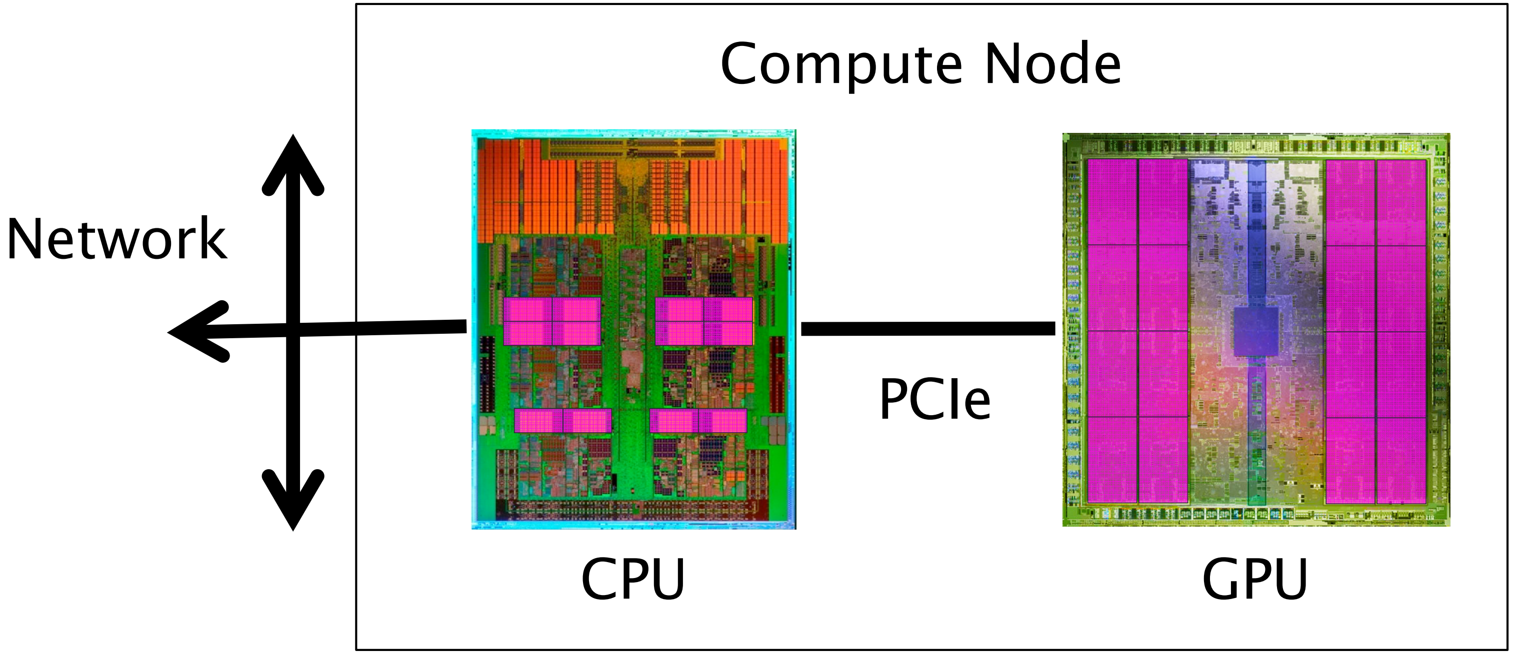
Nodes of ARCHER	2	4	8	16	32	48	64	96
Old Algorithm (millisec)	26	32	51	153	389	1140	1864	5406
New Algorithm (millisec)	17	20	34	69	115	171	305	607
Speedup	1.53	1.60	1.50	2.22	3.38	6.67	6.11	8.91

Table 1: Time in `optimize_load_list`

Saving 3.3GB memory per node!

Hartree Centre
Science & Technology Facilities Council

Adapting to GPU

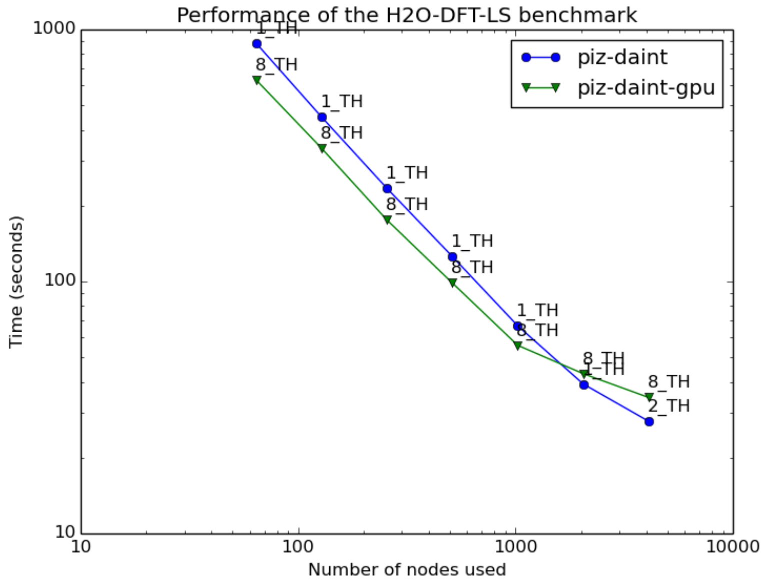


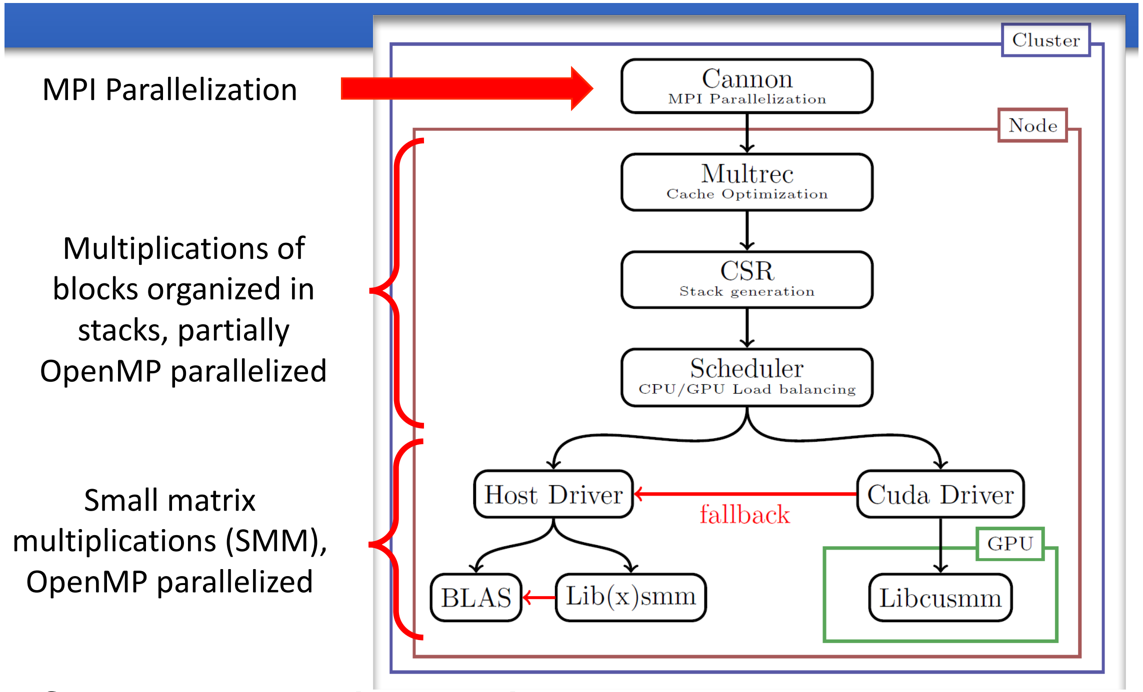
- CPU socket – 0.3 TFLOP/s, ~50 GB/s, 64+ GB DDR
- GPU socket – 1-5 TFLOP/s, ~700+ GB/s, 8-16 GB GDDR
- PCIe x16 – 32 GB/s
- 10,000s threads needed
- Programmability!

Hartree Centre
Science & Technology Facilities Council

Adapting to GPU

- CUDA kernels for SMM (and FFT)
- Latency hiding
- Load balance GPU & CPU
- Result: 25% speedup

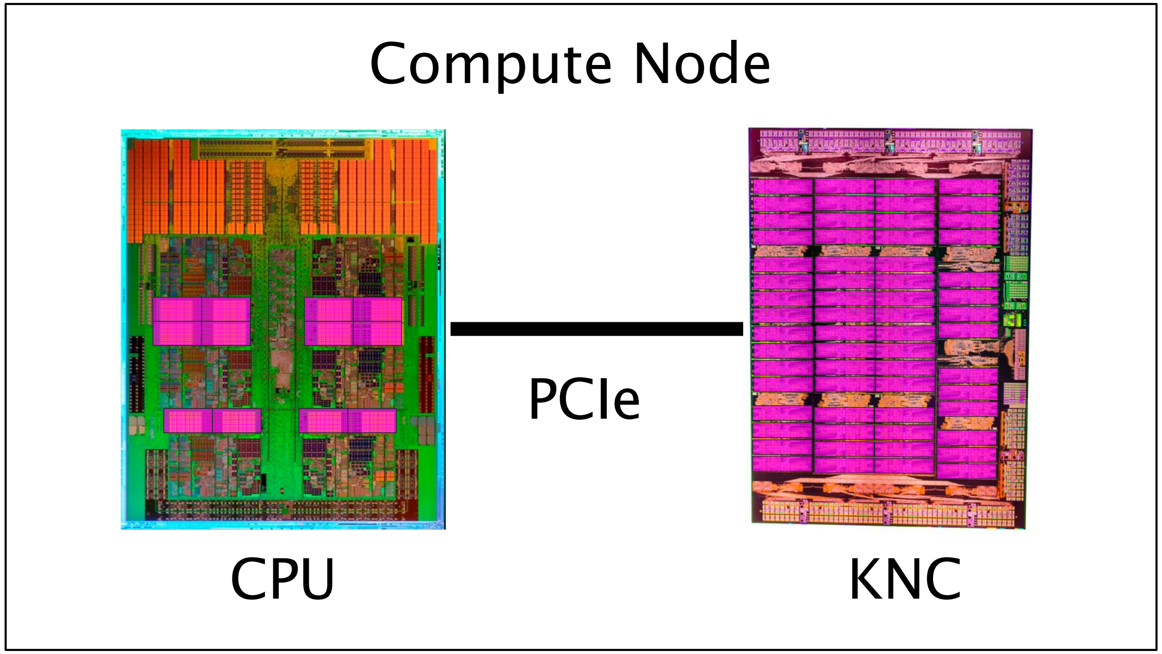




Ref: *"GPU-accelerated Sparse Matrix-matrix Multiplication for Linear Scaling DFT"*, Schütt et al, Electronic Structure Calculations on GPUs (2016)

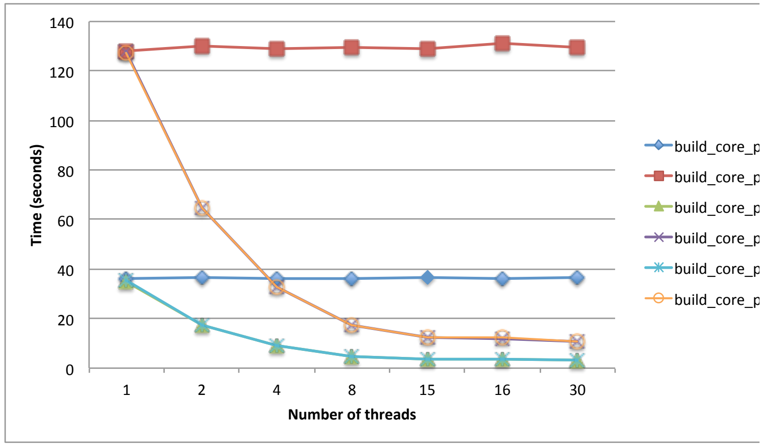
Hartree Centre
Science & Technology Facilities Council

Adapting to Xeon Phi KNC



- CPU socket – 0.3 TFLOP/s, ~50 GB/s, 64+ GB DDR
- KNC socket – 1.2 TFLOP/s, 352 GB/s, 16 GB GDDR
- PCIe x16 – 32 GB/s
- 240 threads needed
- Less parallelism, easy to program!

Adapting to Xeon Phi KNC

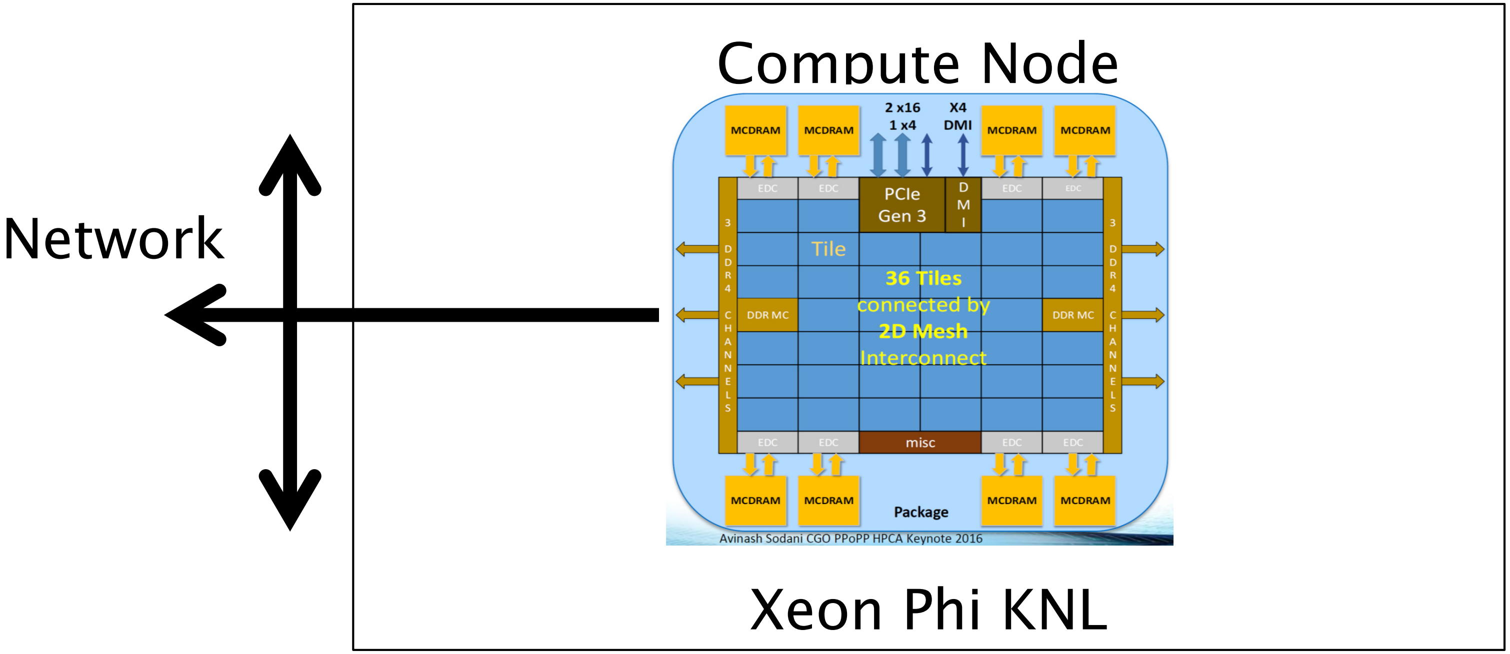


- Requires excellent vectorisation (hard)
- Requires Intel compiler suite (tricky)
- Requires scaling to 240 threads while fitting into 16 GB (very hard)
- P54C cores (from 1993!) exposed by complex logic, branching, function calls
- Lots of work on efficient OpenMP, memory reduction…
- Result: KNC in native mode 4-8x slower than Sandy Bridge Xeon!

Refs: *"Evaluating CP2K on Exascale Hardware: Intel Xeon Phi"*, *"Optimising CP2K for the Intel Xeon Phi"* FR + **IB**, PRACE white papers, 2013

Hartree Centre
Science & Technology Facilities Council

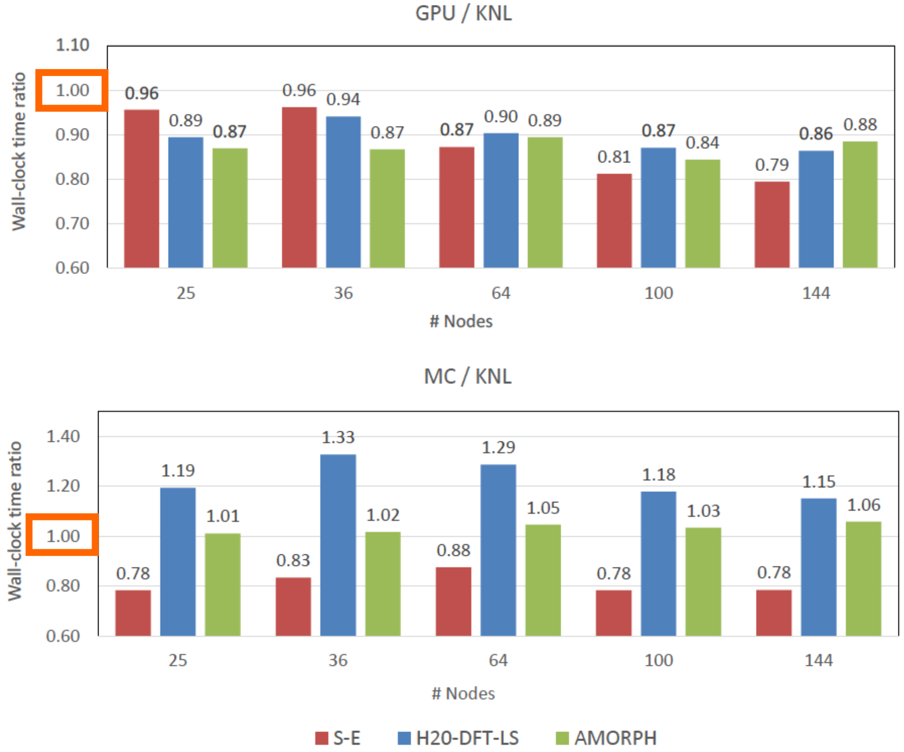
Adapting to Xeon Phi KNL



- KNL socket – ~3 TFLOP/s, ~450+ GB/s (HBM)
  - 96 GB DDR + 12 GB HBM
- PCIe x16 – 32 GB/s
- 128/256 threads needed
- No offload model!

Adapting to Xeon Phi KNL

>1 ➔ KNL faster
<1 ➔ KNL slower



No KNL-specific tweaks

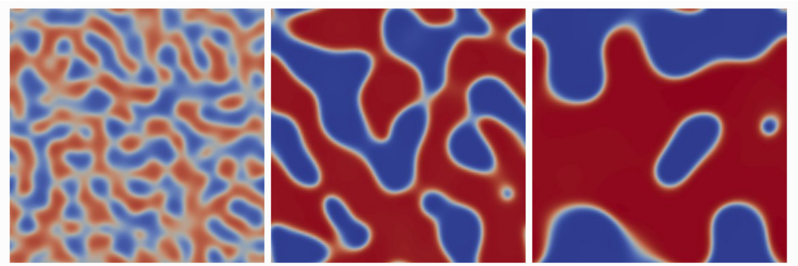
- S-E
  - Small blocks size
  - Dominated by stacks preparation and communications
- H2O-DFT-LS
  - Large blocks size
  - Communication-bound
- AMORPH
  - Medium blocks sizes
  - Computation-bound

Ref: *"Porting of the DBCSR library for Sparse Matrix-Matrix Multipication to Intel Xeon Phi systems"*, **IB** et al, ParCo 2017

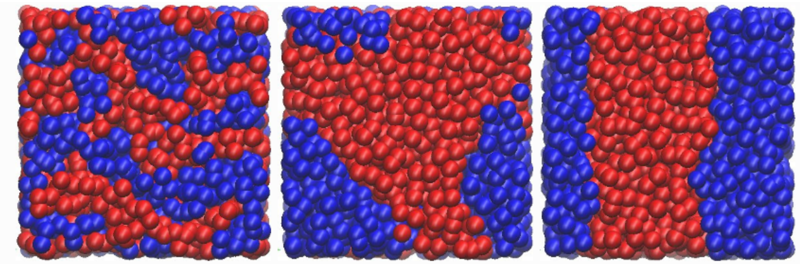
Hartree Centre
Science & Technology Facilities Council

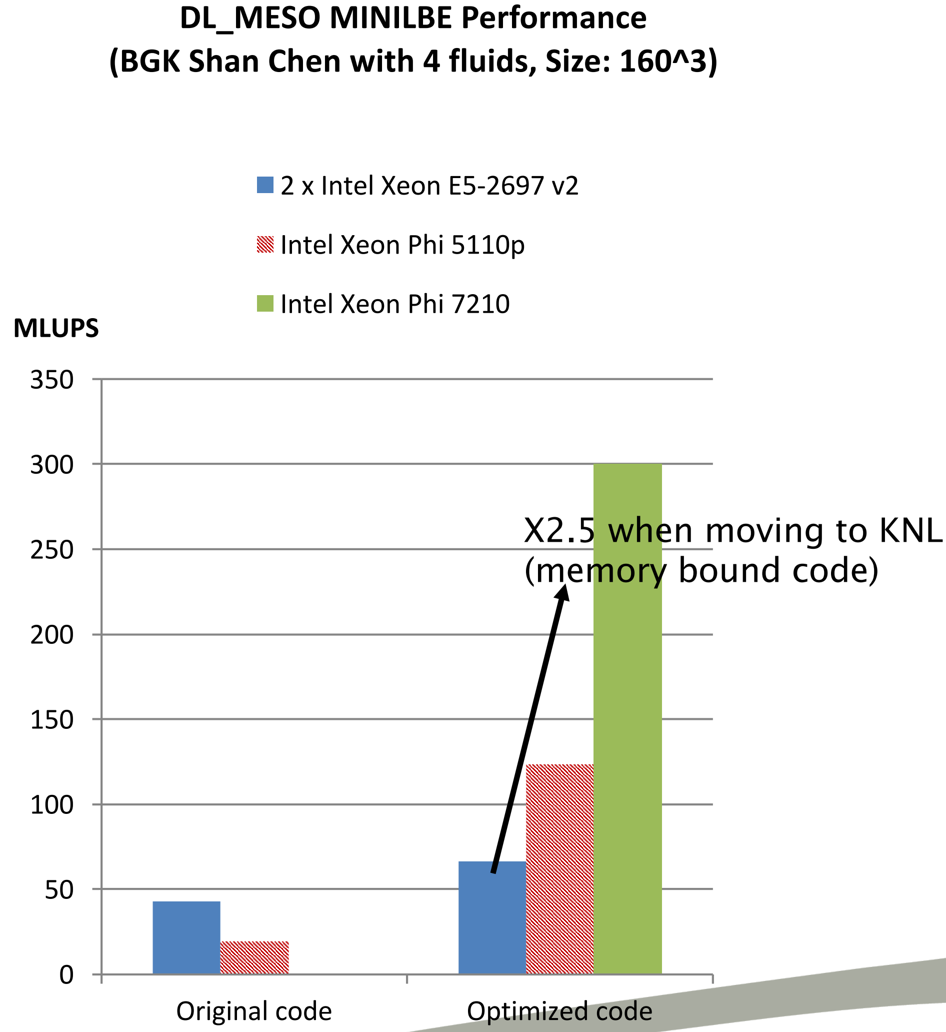
Mesoscale Chemistry Simulations : DL_MESO

- Lattice Boltzmann



- Dissipative Particle Dynamics

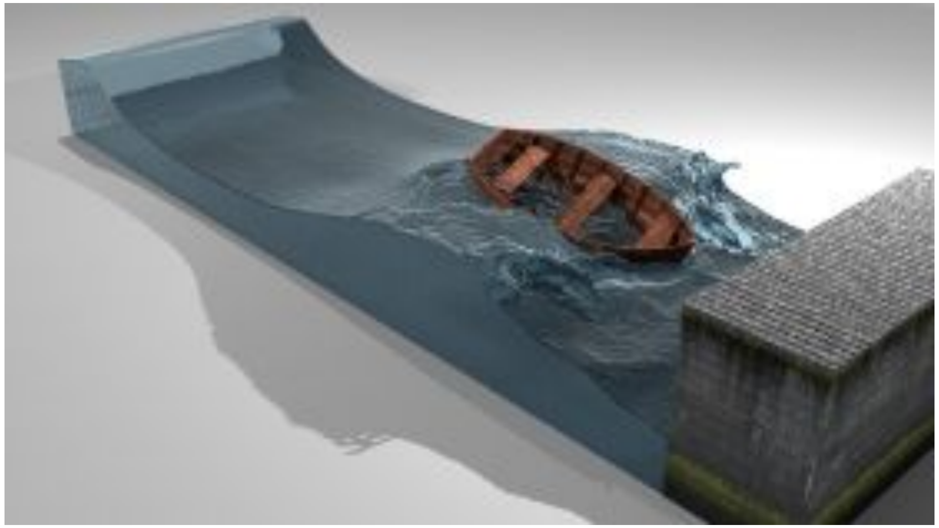


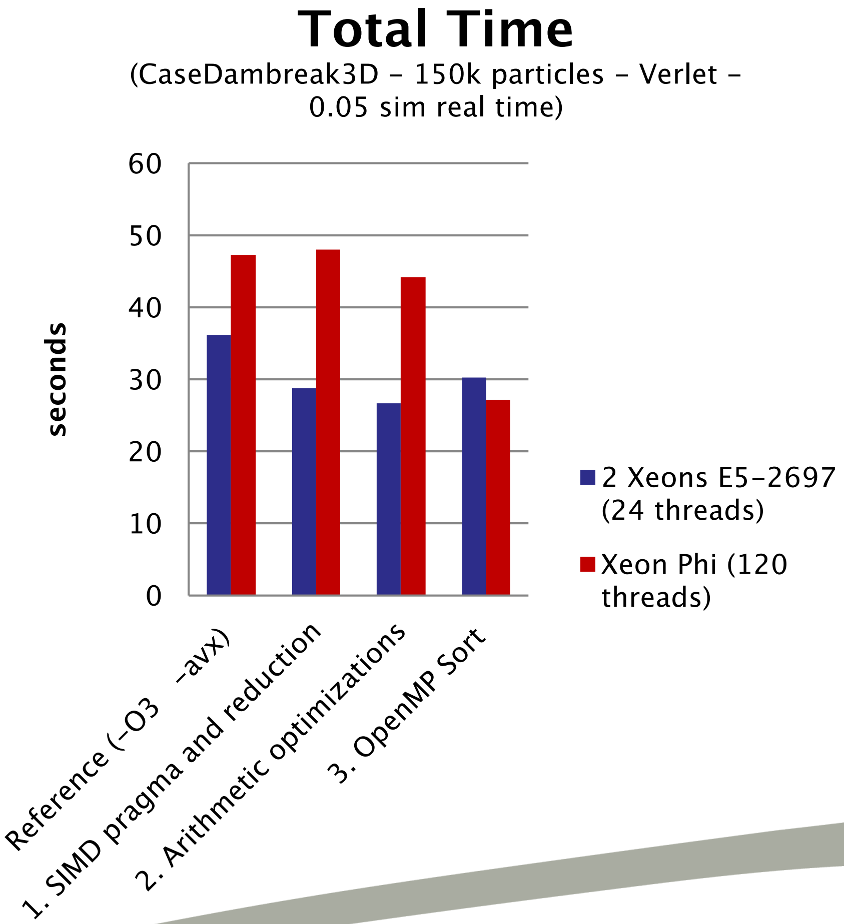


Work from Hartree IPCC

Hartree Centre
Science & Technology Facilities Council

Hydrodynamics simulation : DualSPHysics





Work from Hartree IPCC

Hartree Centre
Science & Technology Facilities Council

Conclusions

- The future of HPC is going to be:
  - More parallel, more heterogeneous, more dynamic
  - More work for the programmer

- We need to start preparing codes now
  - Practical benefit = more places to run on
  - Funding available
    - IPCC, ARCHER eCSE, EPSRC Tier-2 Support, PRACE…
- Get in touch with your local RSE team!

Image: Jorge Cham, www.phdcomics.com

Hartree Centre
Science & Technology Facilities Council

Thanks for listening!

iain.bethune@stfc.ac.uk

@ibethune

@HartreeCentre