Extreme-scaling on Omni-Path fabric:

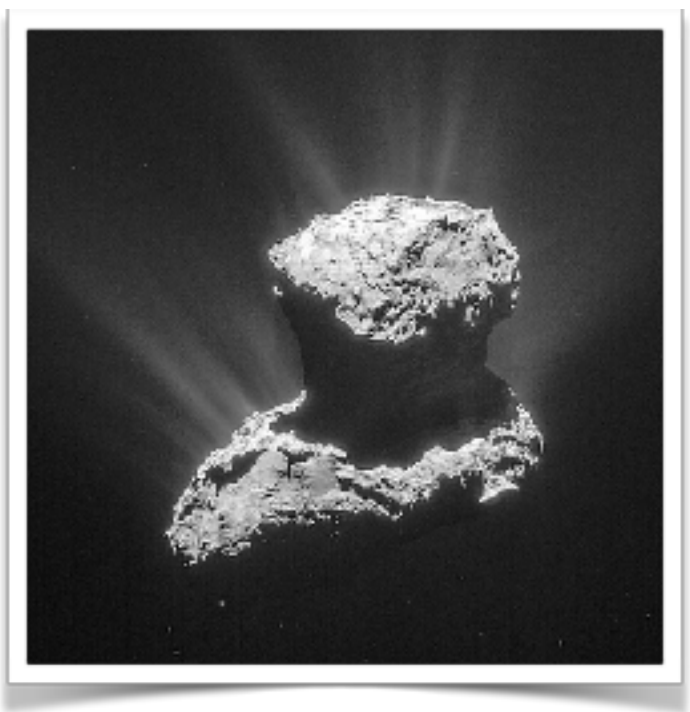
performance for computational astrochemistry

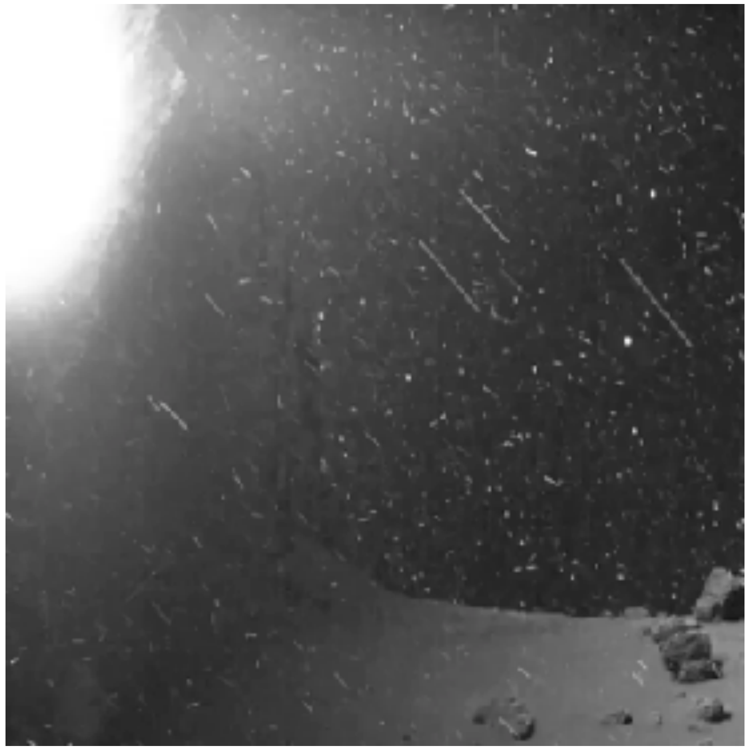
David M. Benoit
E.A. Milne Centre for Astrophysics
Department of Physics and Mathematics
University of Hull, Cottingham Road,
Kingston upon Hull HU6 7RX, UK
d.benoit@hull.ac.uk

@dbenoit1

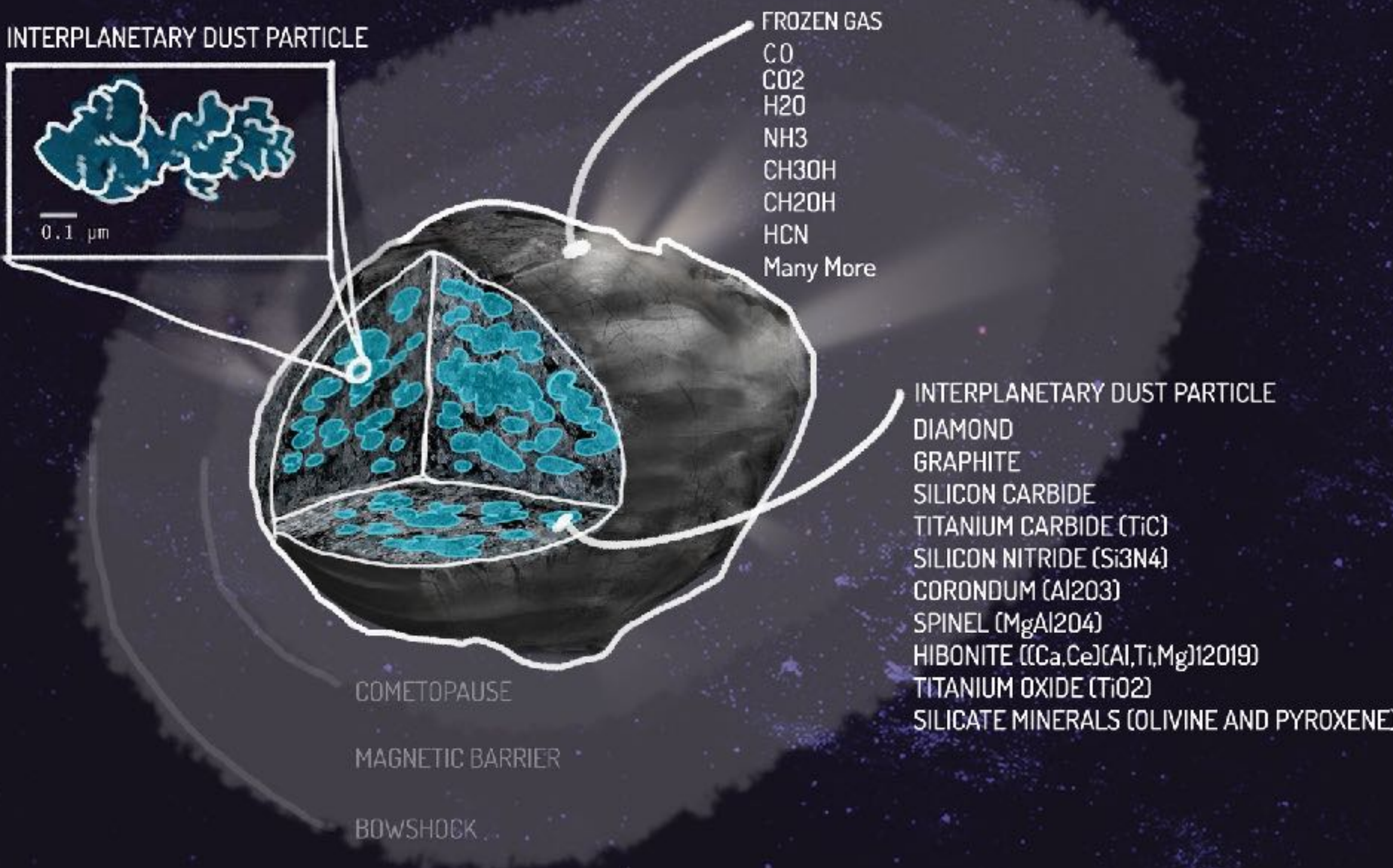
Landing on a comet – ESA Rosetta Mission

- ESA space probe launched in 2004
- Lander module Philae explored comet 67/P Churyumov-Gerasimenko in 2014



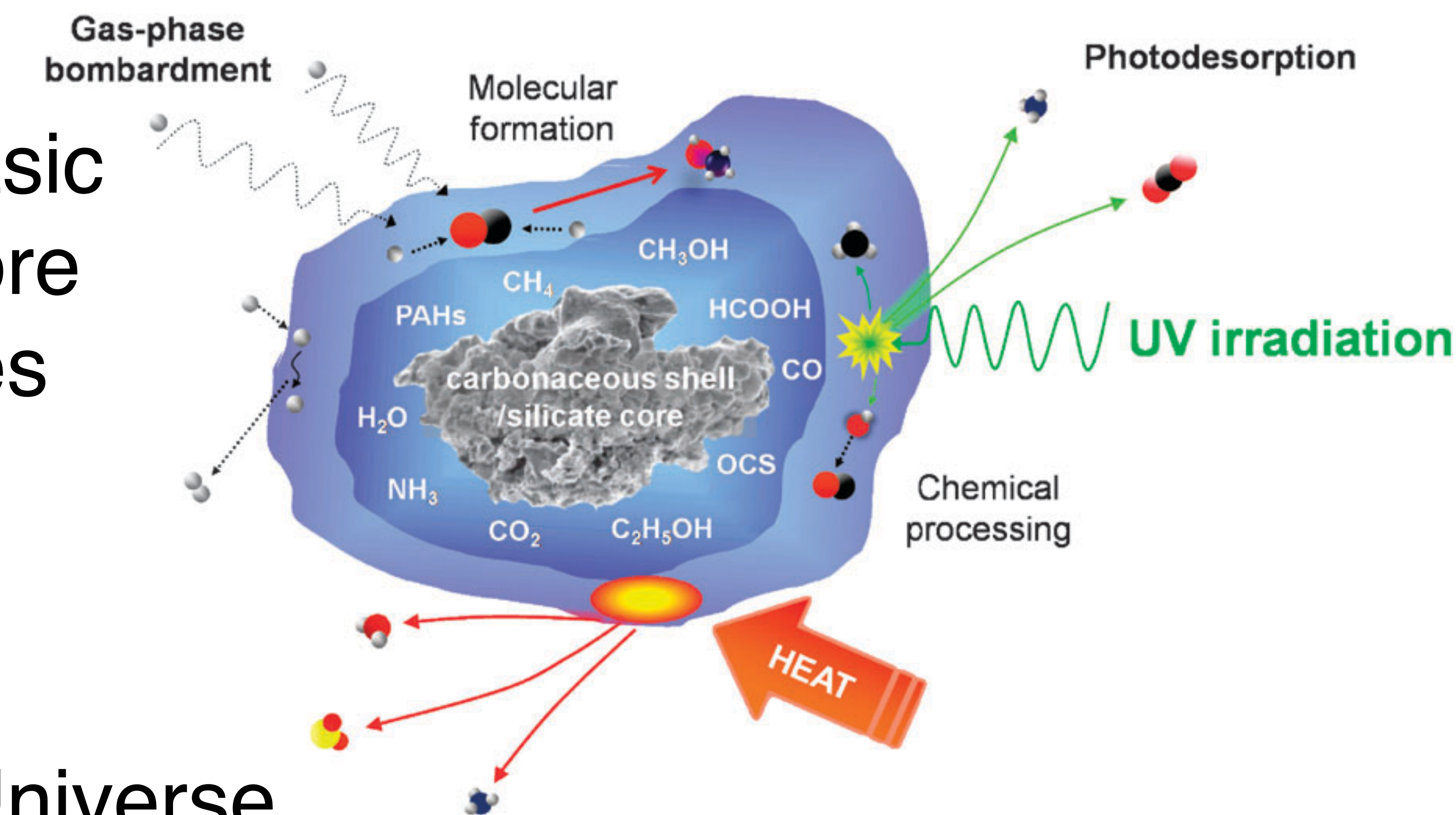


WHAT IS A COMET MADE OF?



Surface chemistry on dust grains

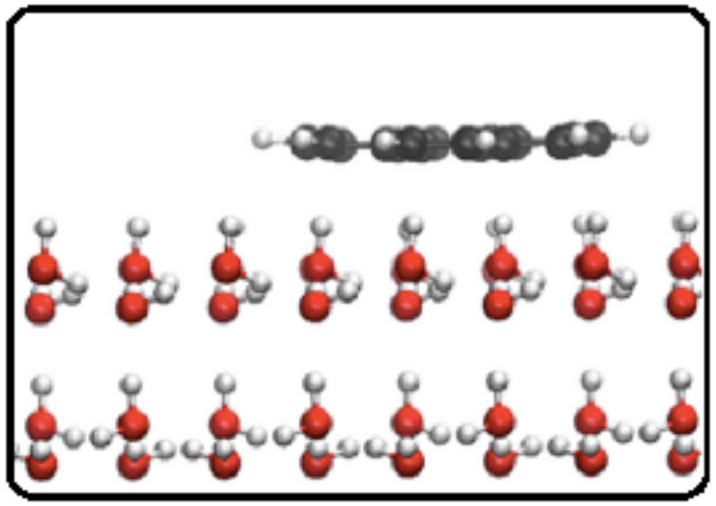
- Dust grains acts as cosmic “bench tops” for chemical reactions
- Help transform basic chemicals into more complex molecules
- Their study is key to understanding low-temperature chemistry in the Universe



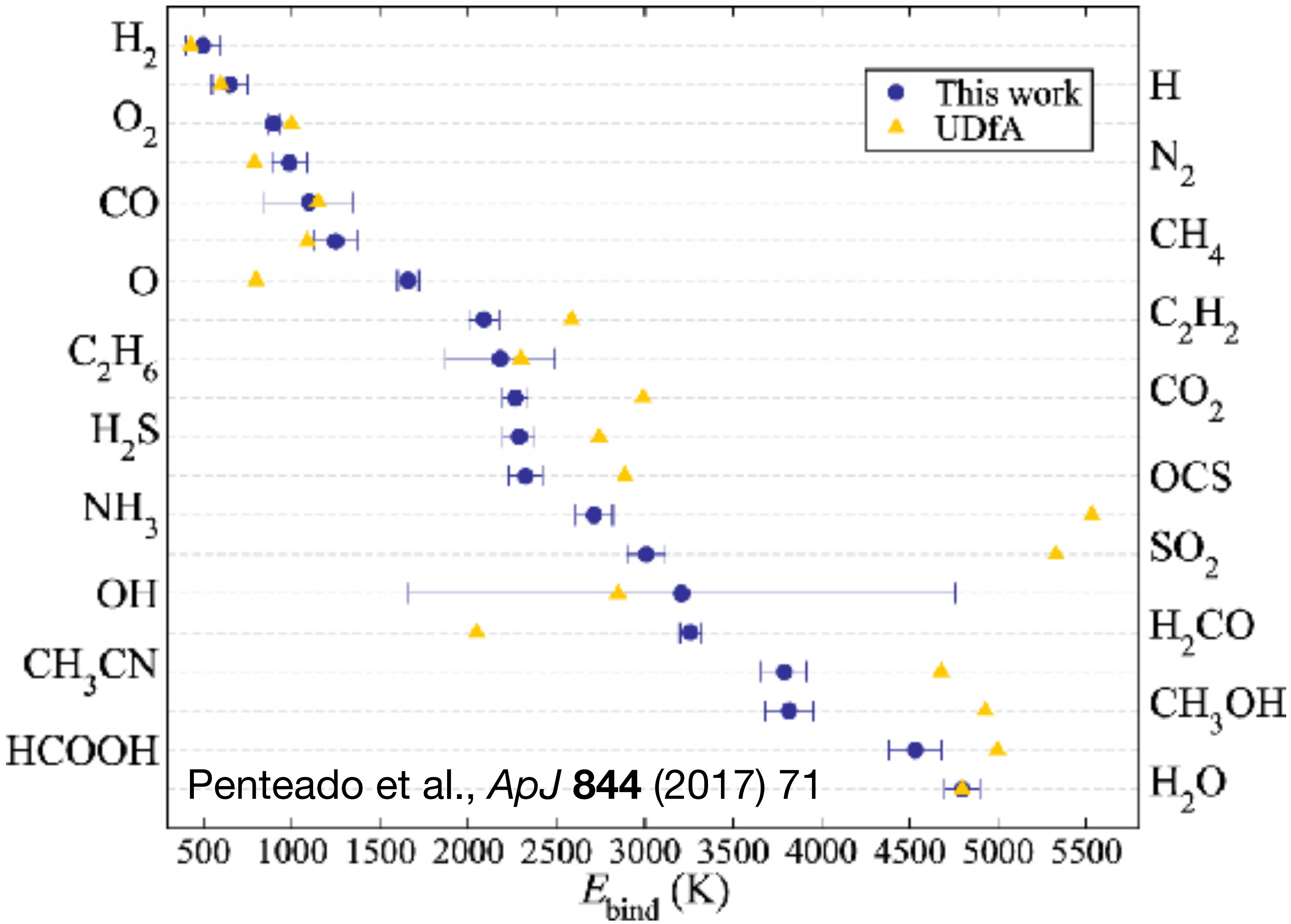
From: D.J. Burke et al., PCCP **12** (2010) 5947

- But how strongly do molecules stick to ice-covered grains/comets?

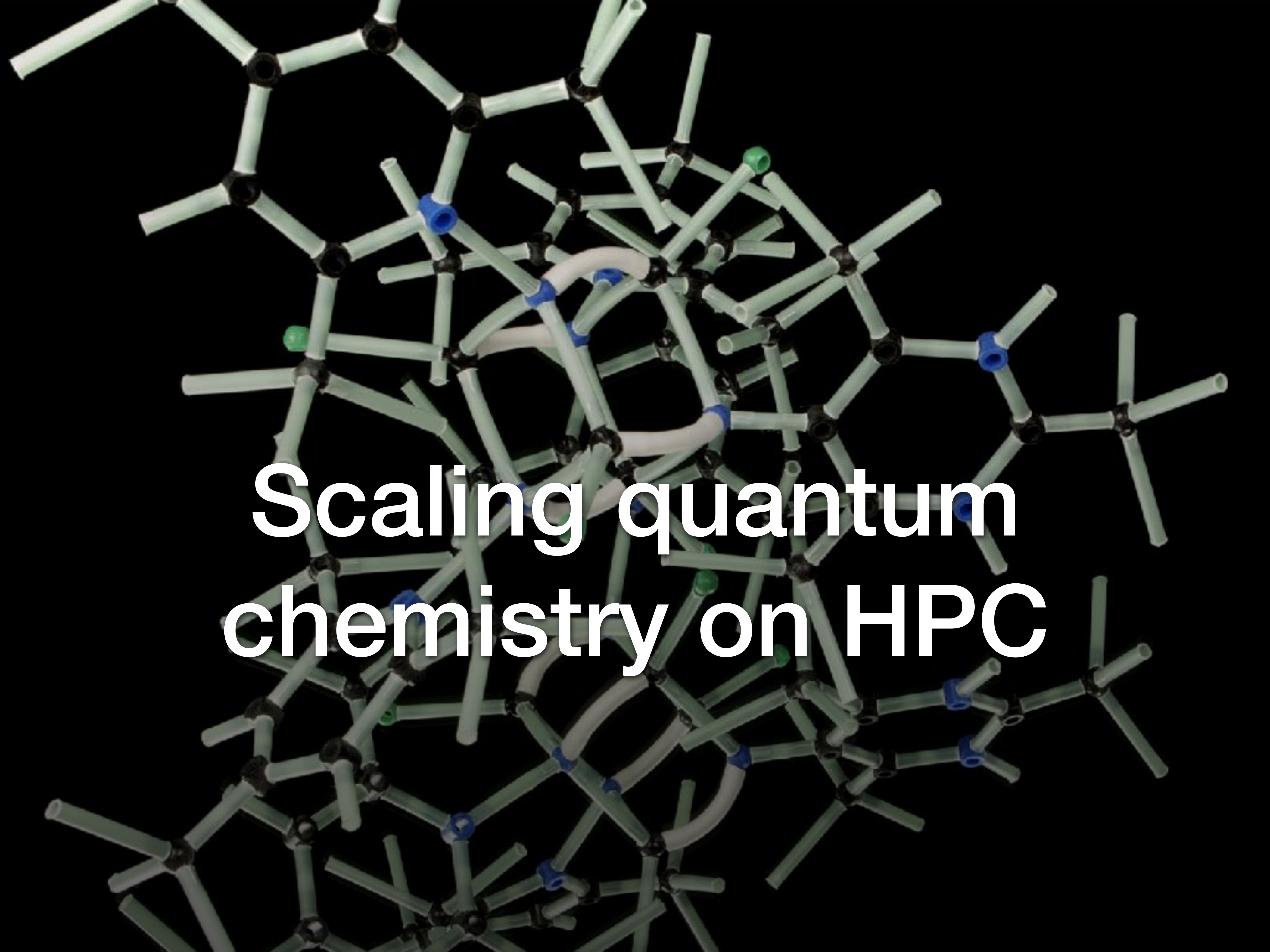
Adsorption / sticking energies?



- Adsorption energies are hard to measure accurately
- Experimental database estimates are from '90s with little experimental/ theoretical validation
- Quantum chemical calculations can provide a reliable estimate



Penteado et al., *ApJ* **844** (2017) 71



Scaling quantum chemistry on HPC

VIPER technical profile

Red Oak Consulting

- 5040 Intel Broadwell E5-2860v4 (2.4 GHz) cores in 180 compute nodes
- Intel X16 100Gb/s Omni-Path interconnect
- 4 x 1 TB high memory nodes
- 2 x visualisation nodes (2 x Nvidia GeForce GTX 980 Ti per node)
- 4 x accelerator nodes (4 x Nvidia Tesla K40M GPUs per node)
- 500 TB of user storage running BeeGFS
- Each node runs a Docker container

clustervision docker BeeGFS

TESSERACT technical profile (SGI 8600)

- 20,256 Intel Skylake silver 4116 (2.1 GHz) cores in 844 compute nodes
- 96 GB NUMA per node (48GB/processor)
- Intel X16 100Gb/s Omni-Path interconnect
- 3 PB of Lustre file system
- DiRAC's Extreme Scaling facility
- Hosted at EPPC/Edinburgh

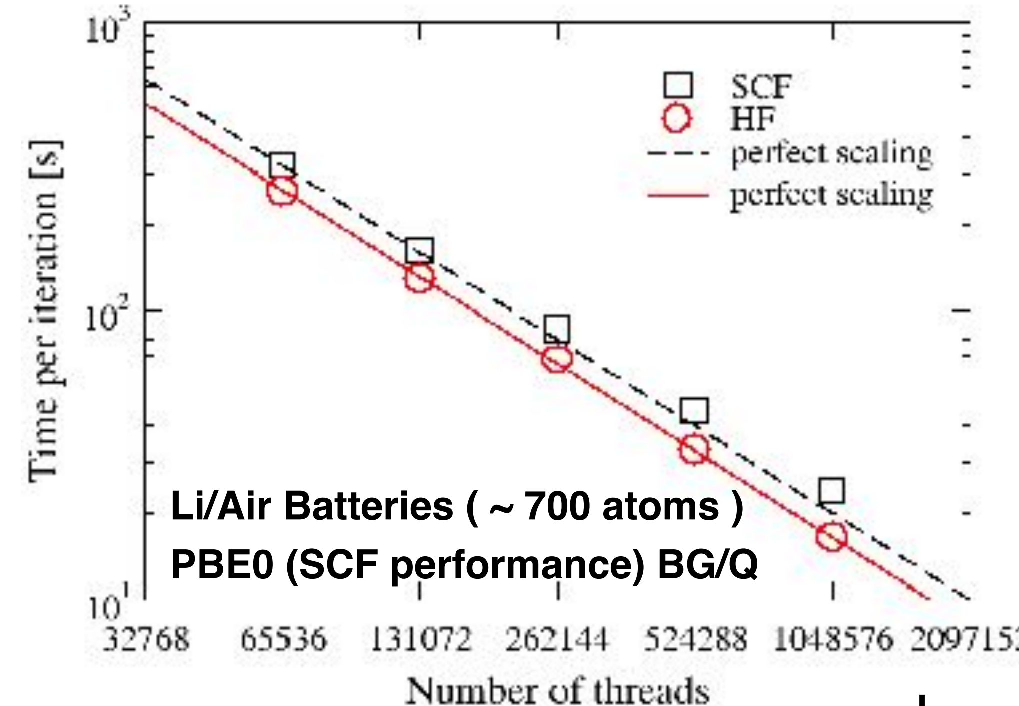
epcc

DiRAC Distributed Research utilizing Advanced Computing

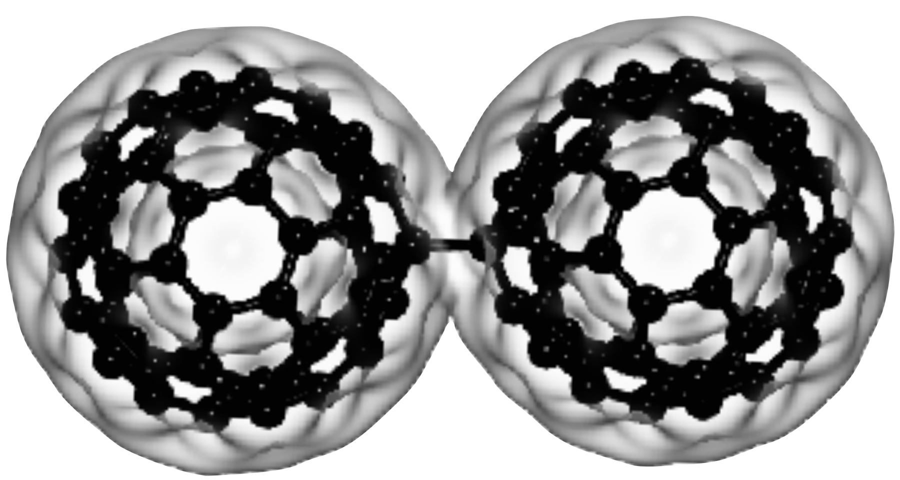
Science & Technology Facilities Council

Application: CPMD – Large-scale electronic structure code

- Well established, freely available, density functional theory (DFT) code
- Developed by MPI-FKF Stuttgart&IBM
- Demonstrated scaling on large HPC
- Hybrid MPI/OpenMP parallelisation
- Bottlenecks: FFT, memory footprint and all-to-all comms
- Our system: fullerene dimer (C_{120}), 480 electrons, triplet state, B3LYP functional and a 500 Ry cutoff

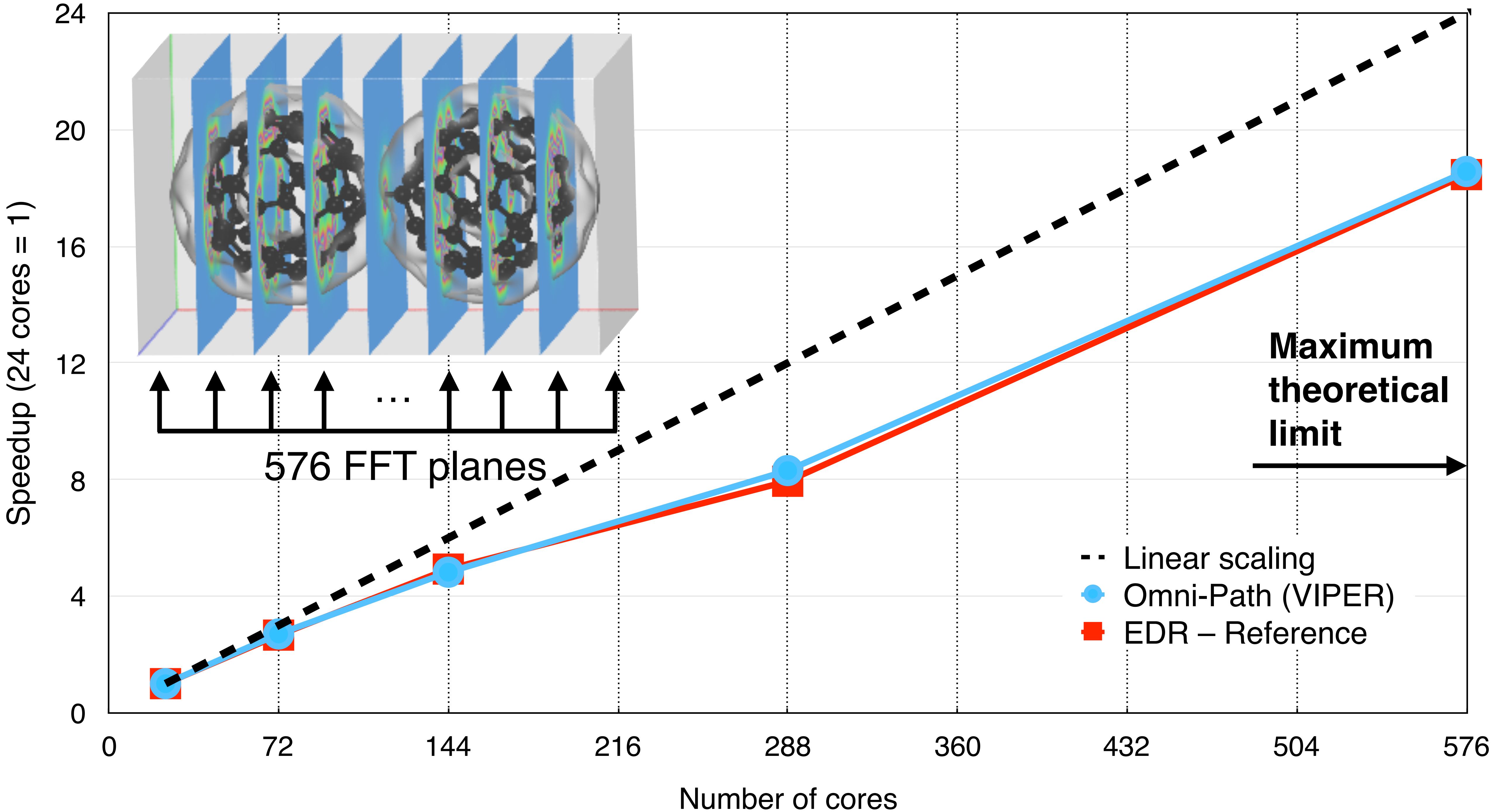


cpmd.org



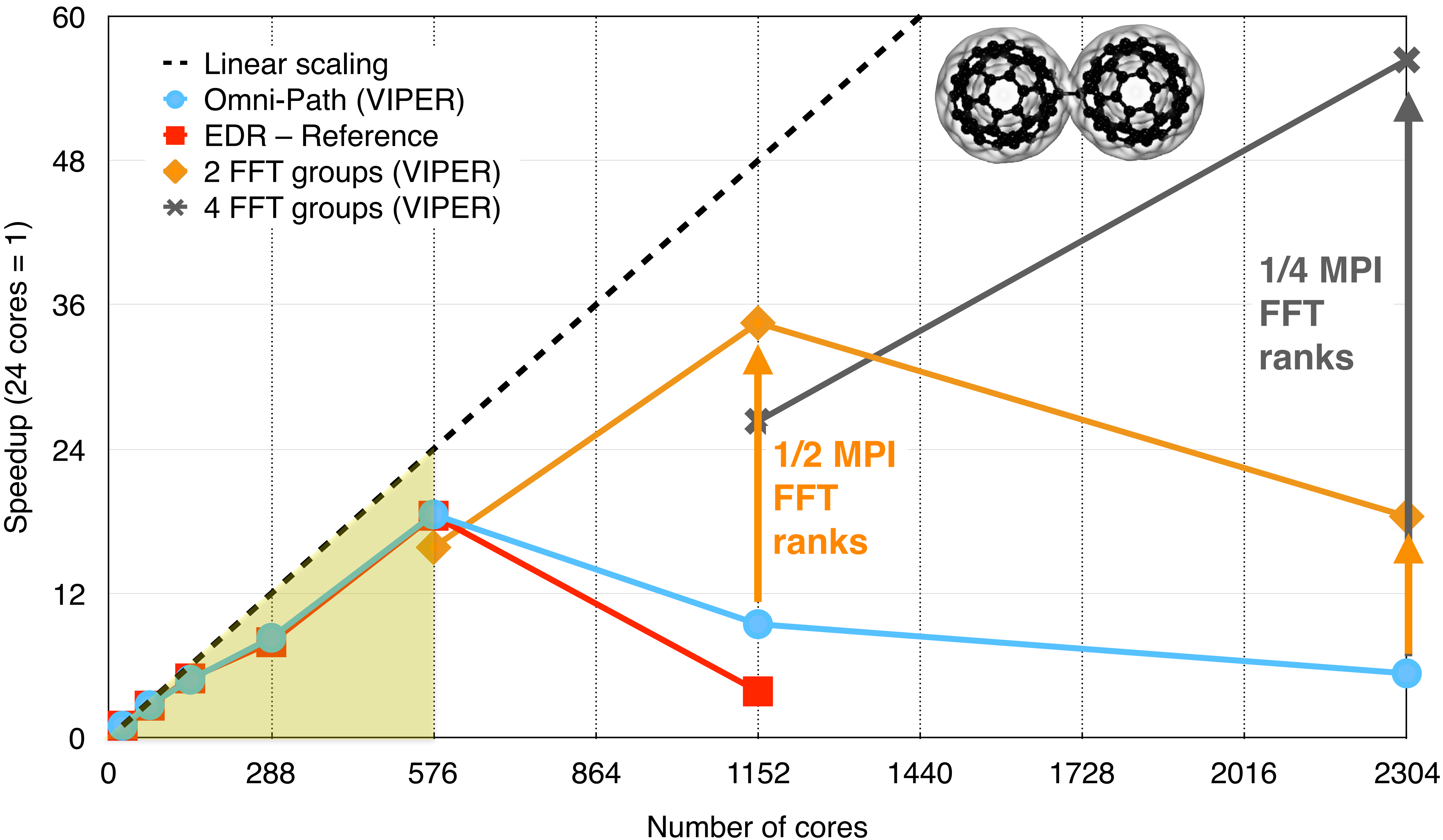
Fullerene dimer, C_{120}

Large-scale density functional theory on OPA

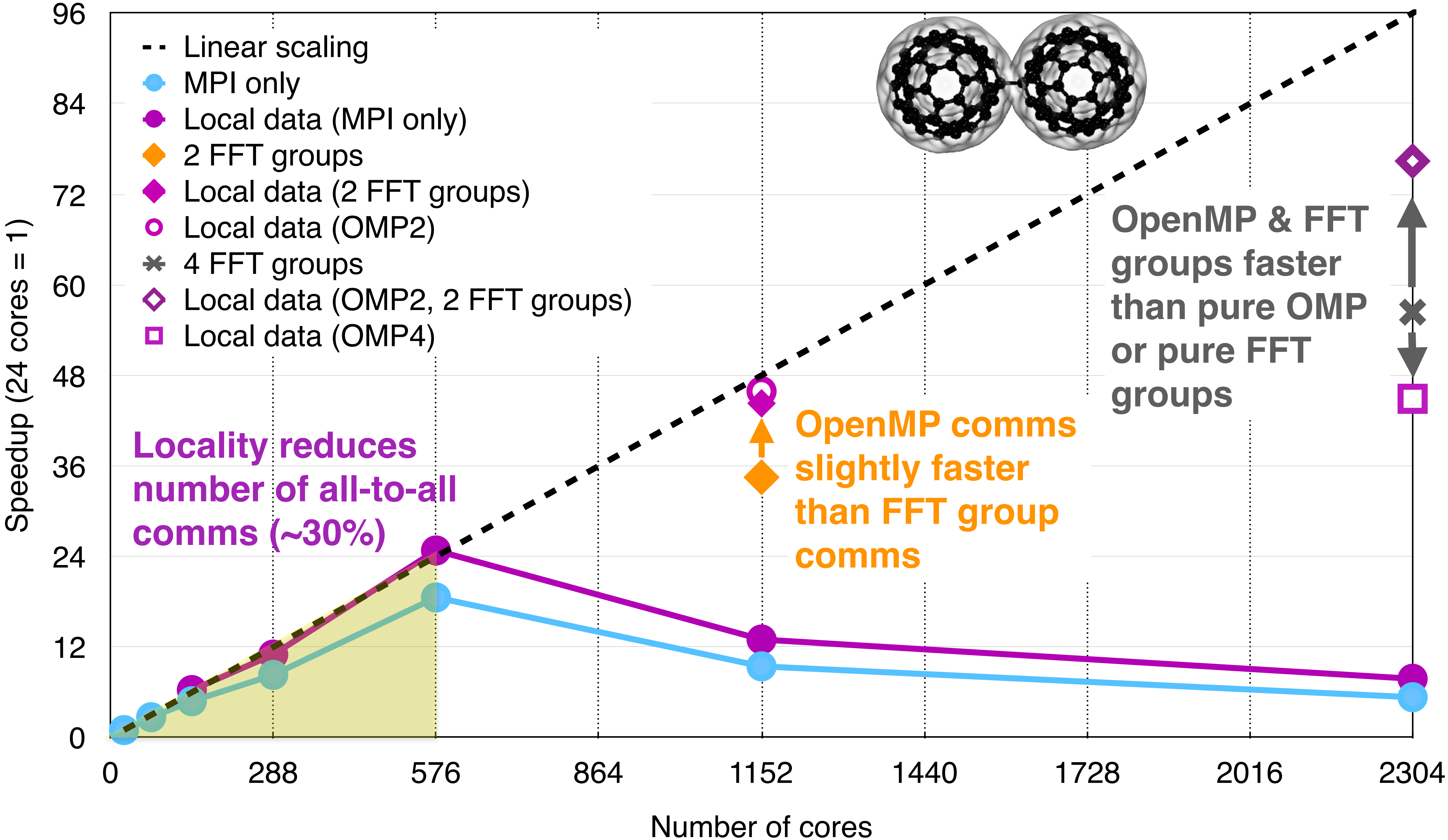


Conditions: Omni-Path HFI Silicon 100 Series; Intel 100 Series 48 port unmanaged switches; CentOS 7.2.1511; dockerised nodes; OPA 10.4.1.0-1; CPMD V4.5 compiled with ifort 2017 (-O2 -ipo -xHOST) ATLAS (single thread) and scaLAPACK

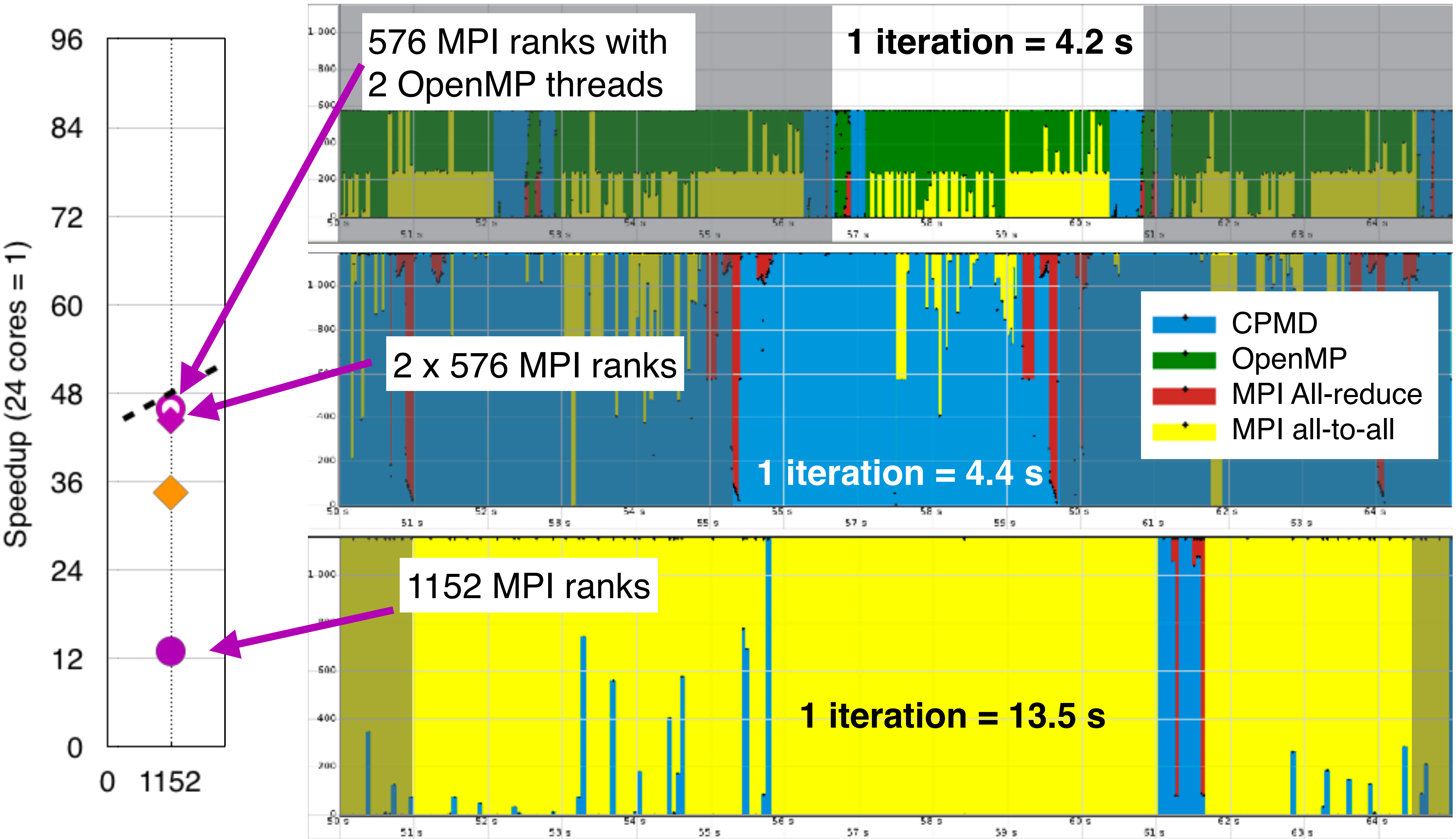
Scaling past the maximum theoretical limit



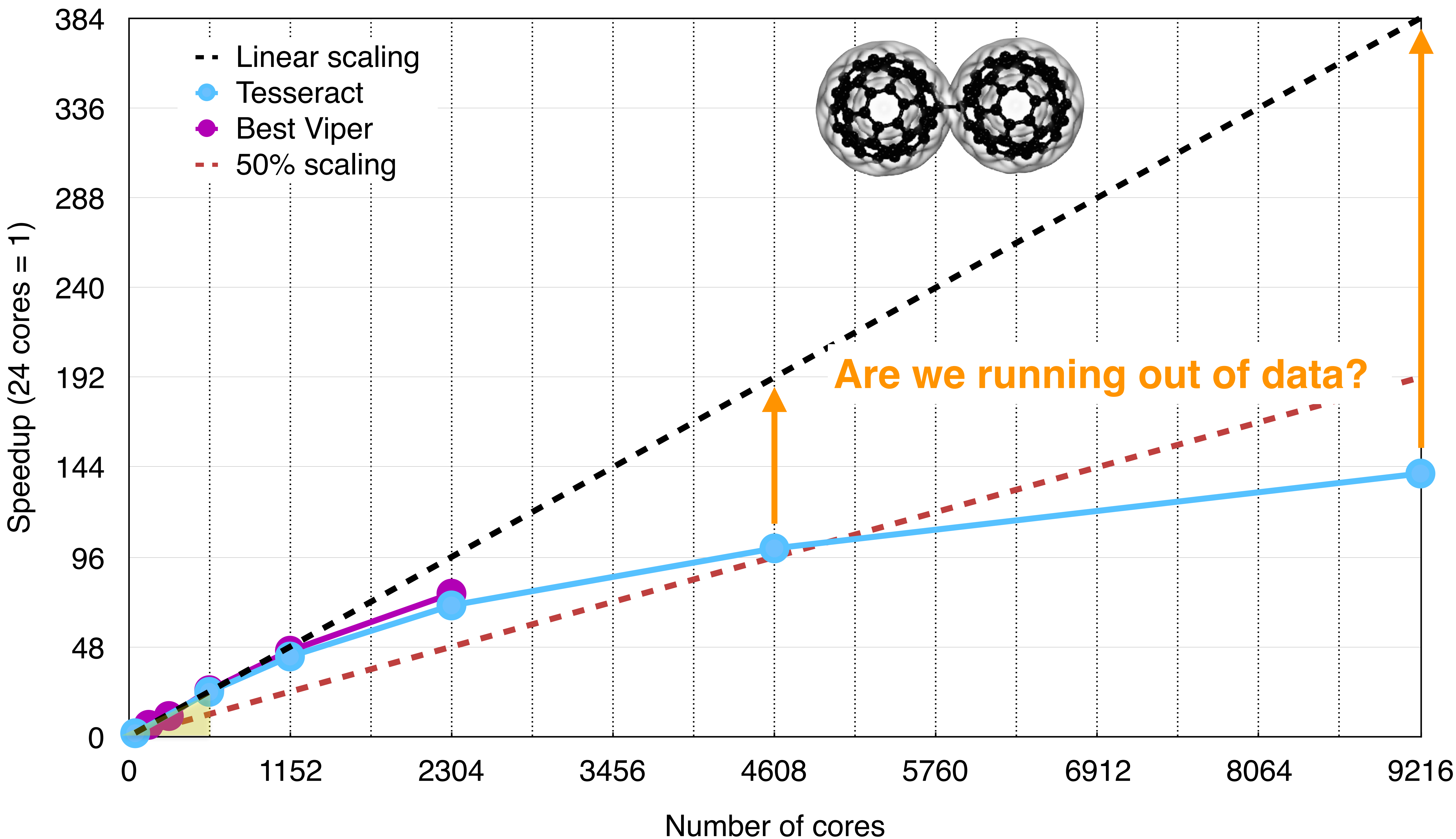
Locality and OpenMP (kitchen sink approach…)

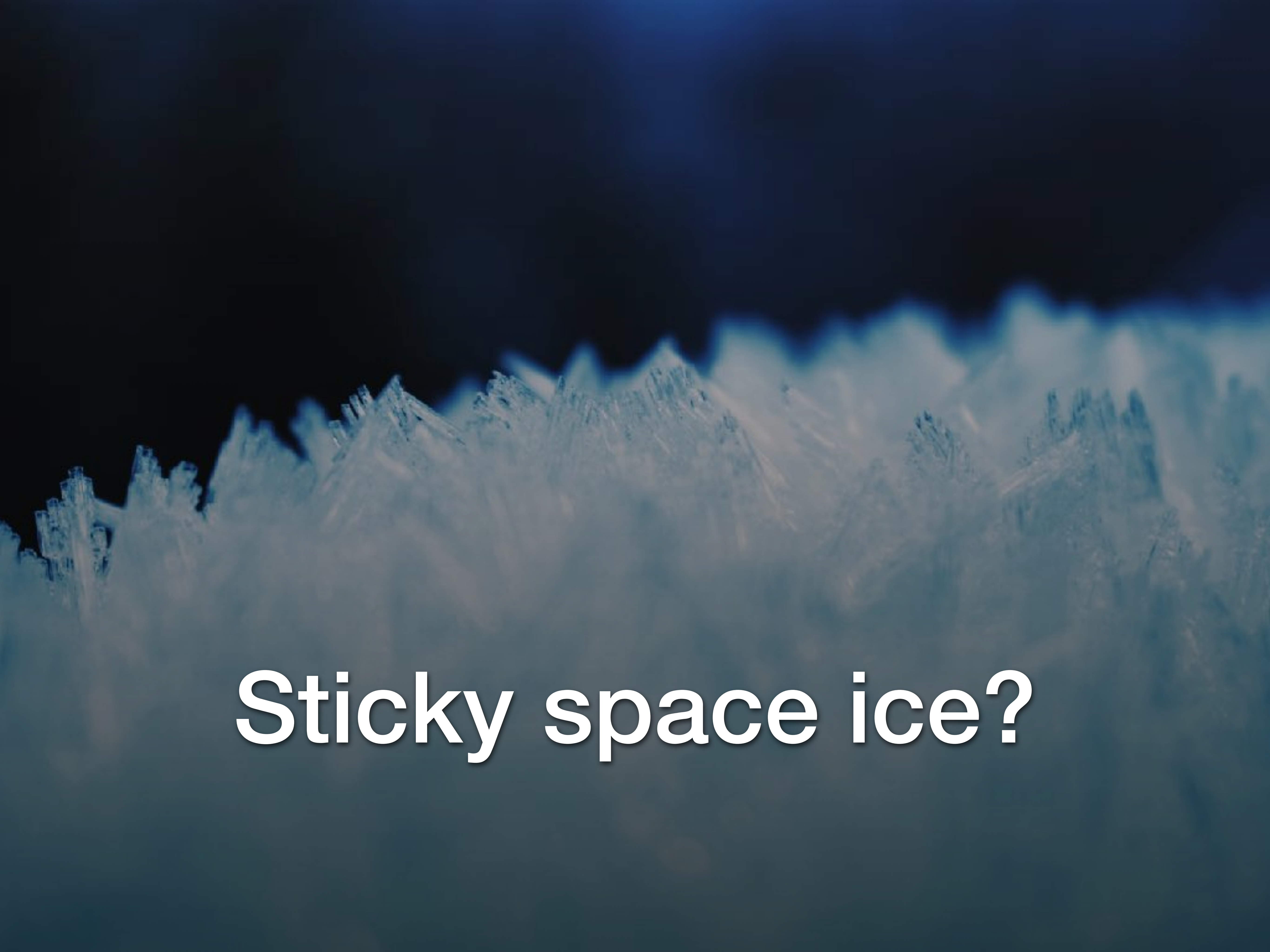


ITAC MPI trace – What happens at 1152 cores?



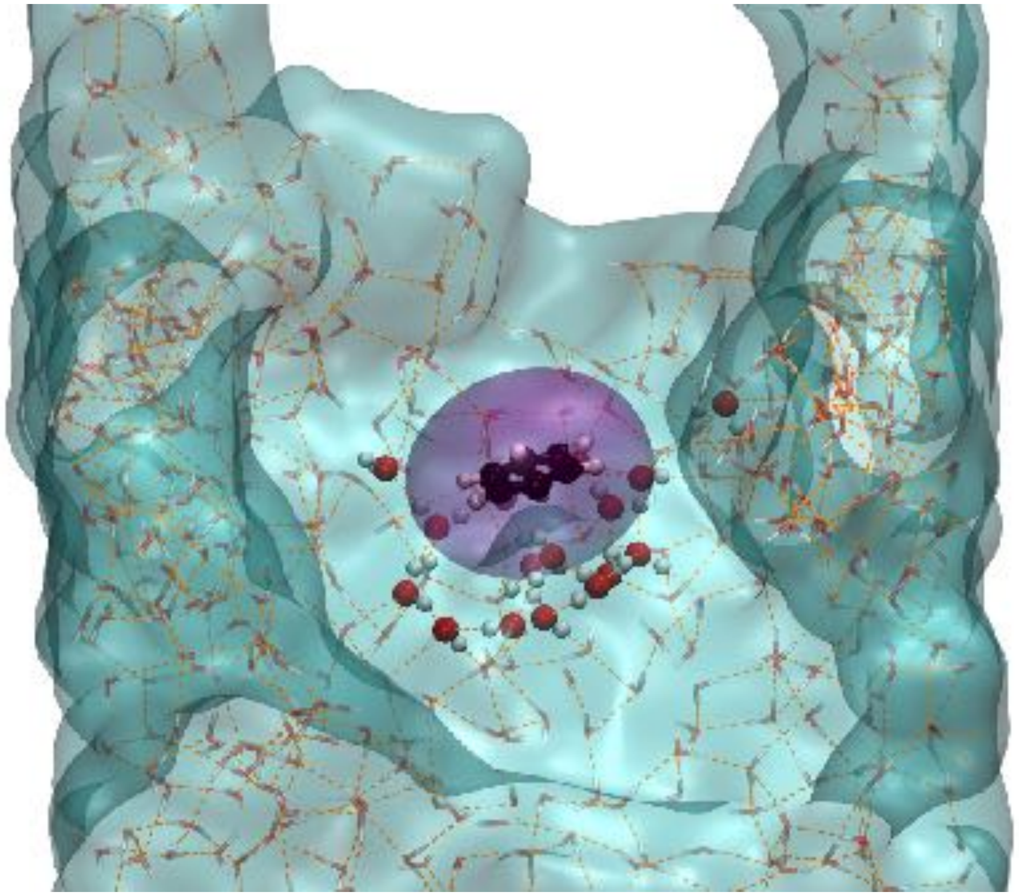
More cores better? – Tesseract scaling

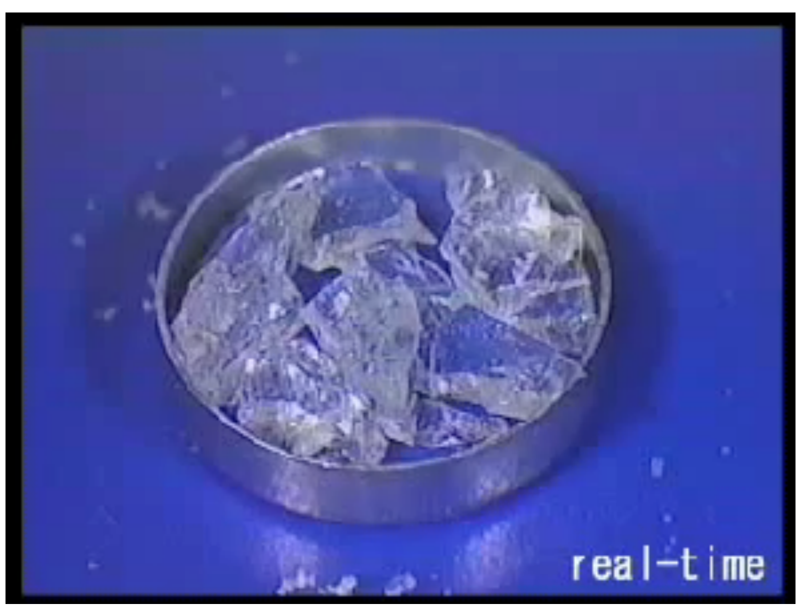



Sticky space ice?

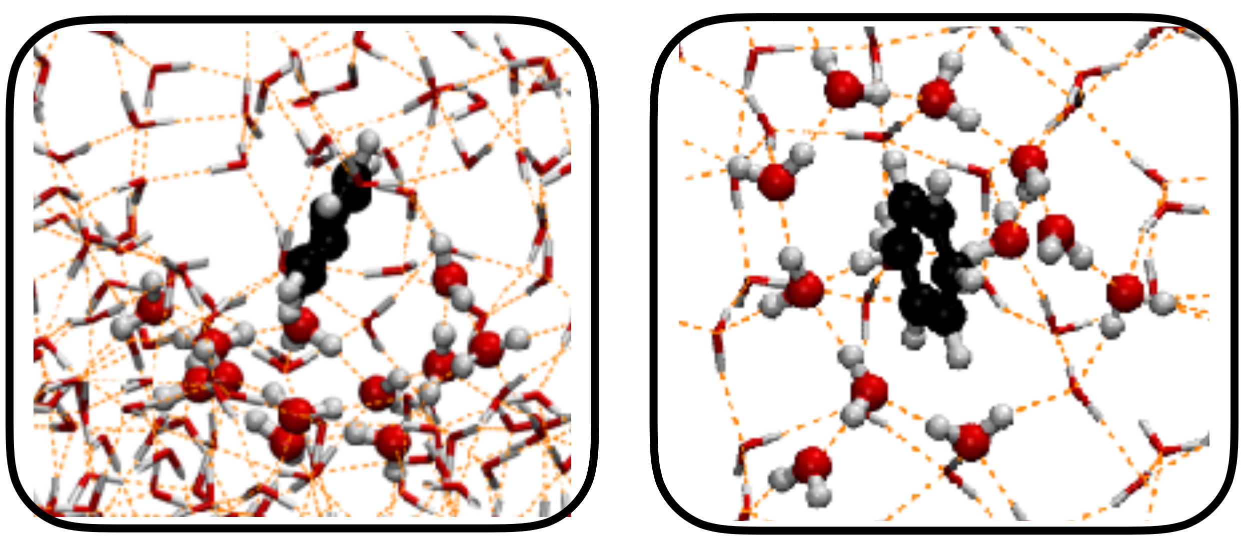
Space ice model: Low density ice

- Amorphous low density ice model with a single benzene molecule, 1512 atoms, 4030 electrons, PBE functional and a 200 Ry cutoff
- 10-fold size increase from C_{120}

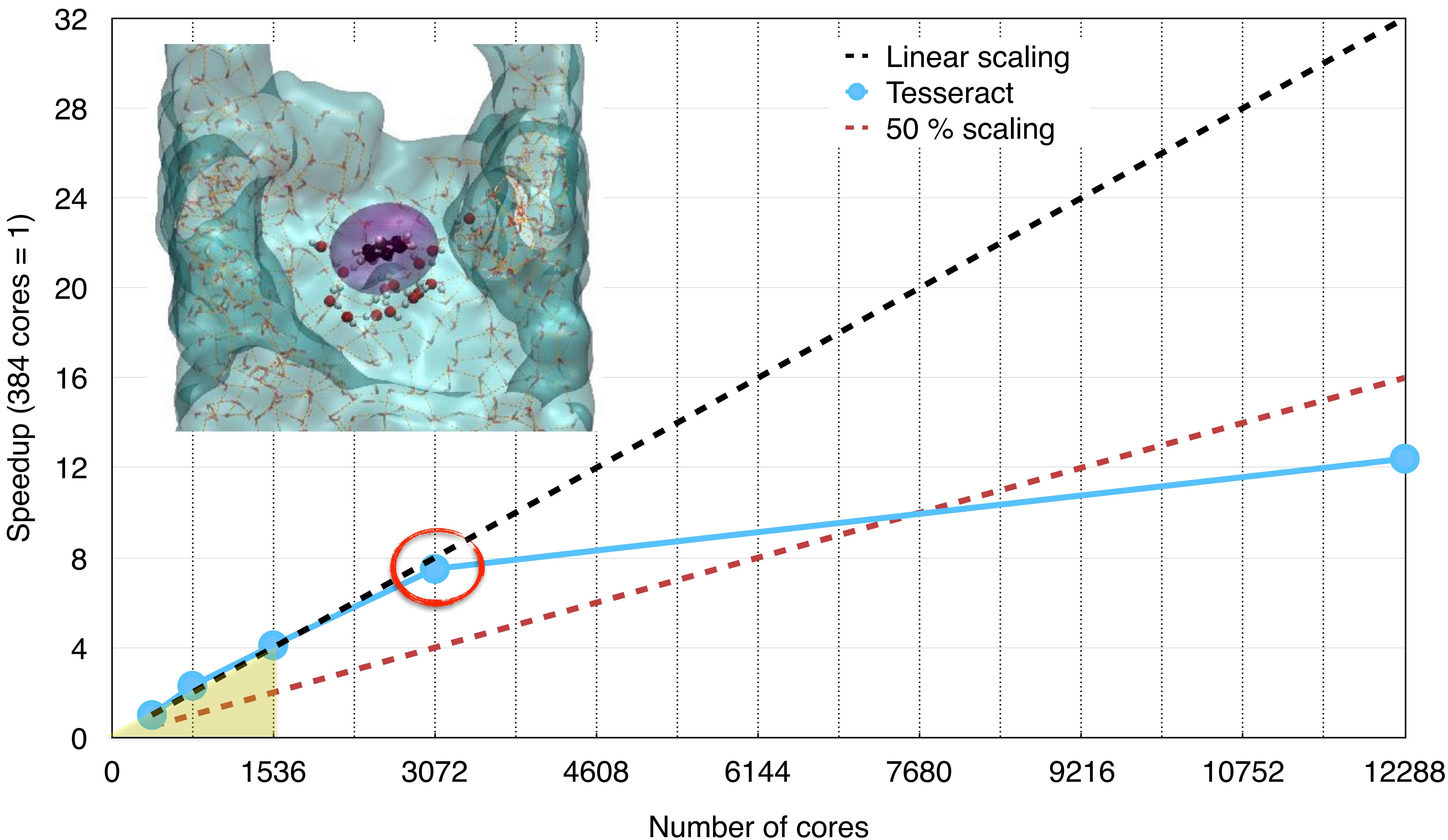




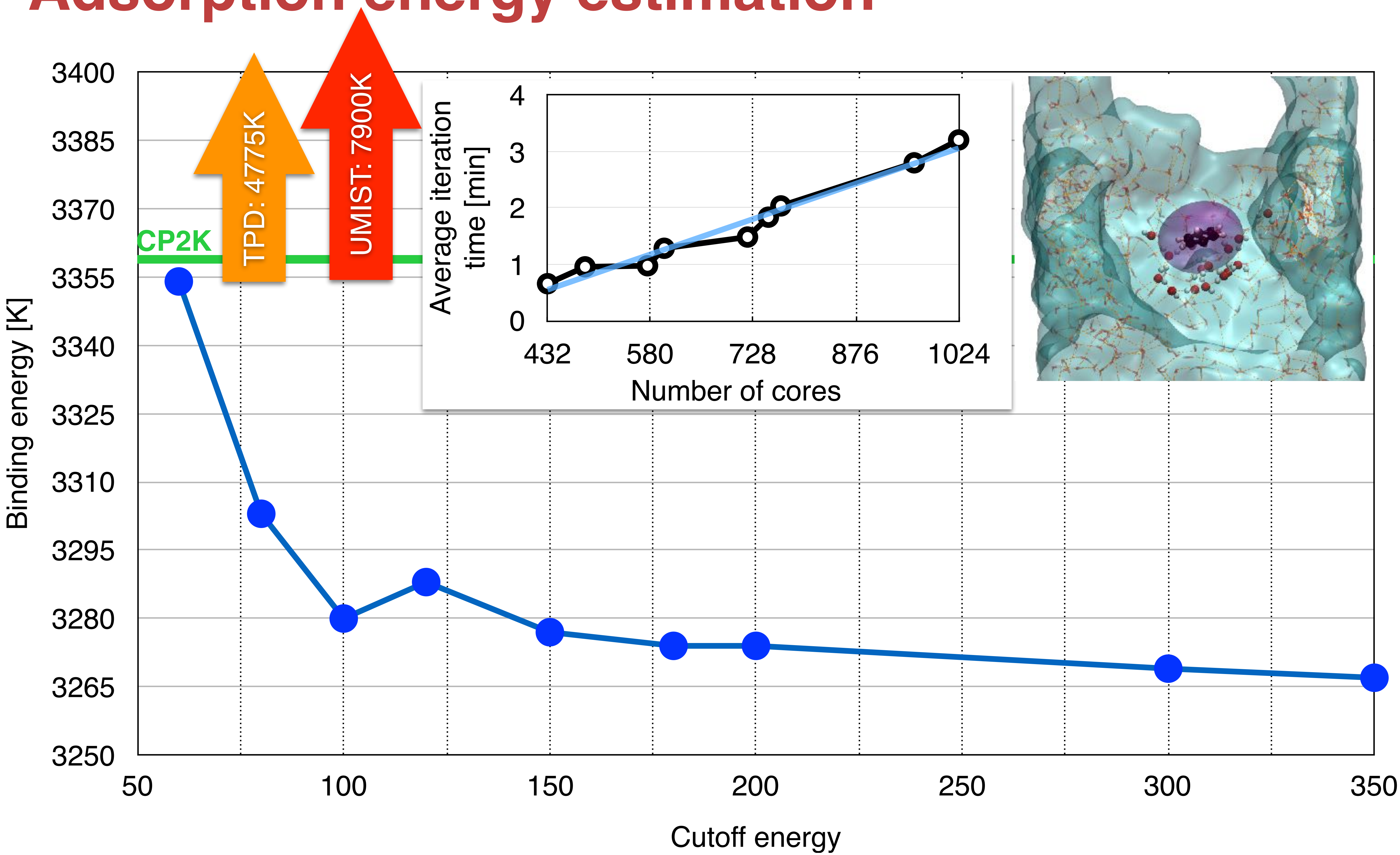
http://www.nims.go.jp/water/hda_lda_tr.html



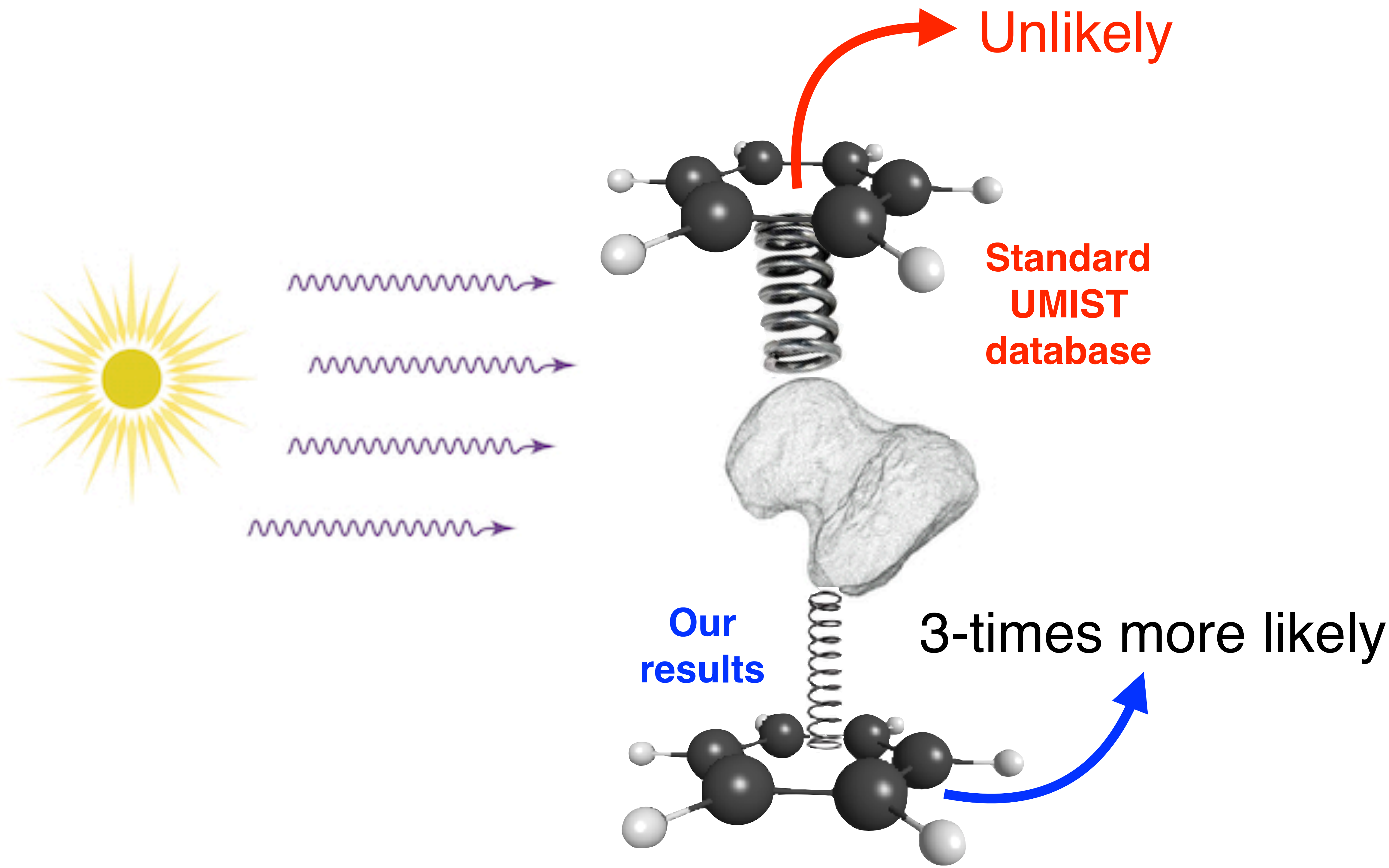
Tesseract scaling



Adsorption energy estimation



Why is it important?



Modified from: ESA/Rosetta/RPC-ICA

Conclusions

- Omni-Path performance are similar to EDR in conventional scaling region but better in non-scalable regime

- Combination of enhanced locality, reduced MPI ranks and FFT grouping leads to predictable scalability

- However, understanding application and problem size are key to scaling at large core counts

Acknowledgements

» VIPER HPC support team

» The University of Hull for funding

» DiRAC / STFC / EPCC for CPU time