

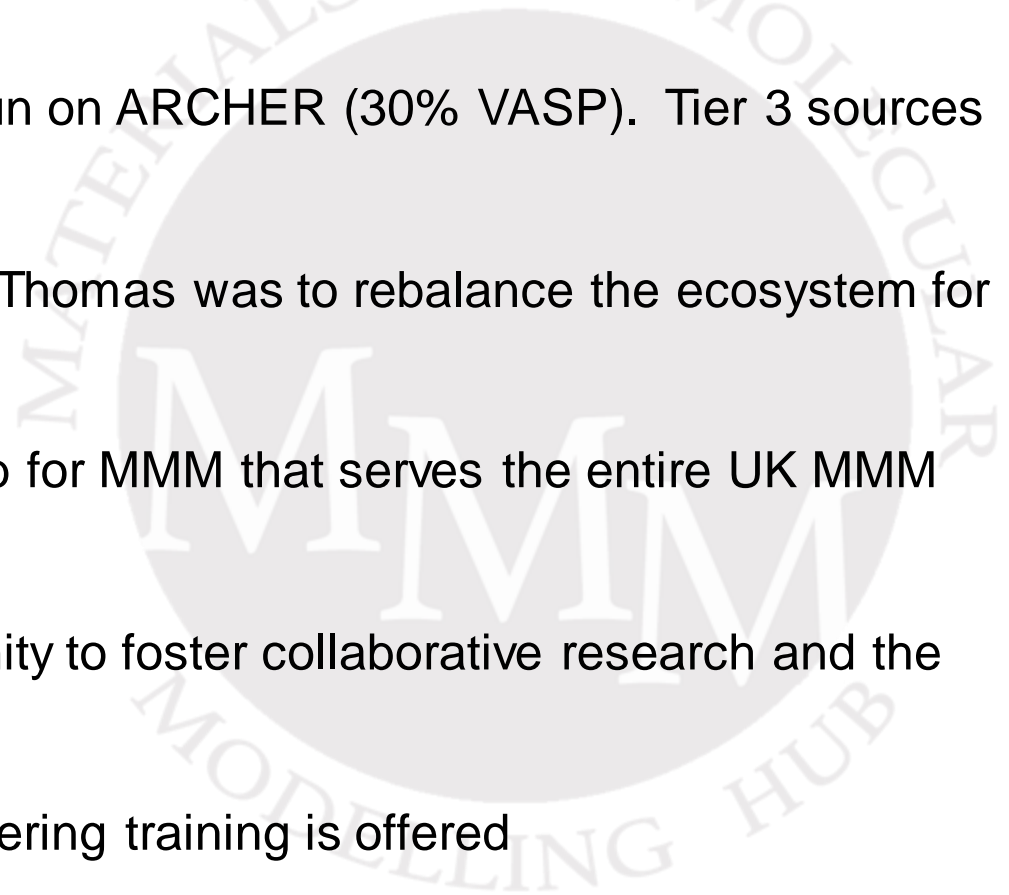
**Introduction to Thomas:  
Materials and Molecular Modelling Hub**



# Motivation for an MMM Hub



- Increased growth in UK MMM research created an unprecedented need for HPC, particularly for medium-sized, high-throughput simulations
- These were predominantly run on ARCHER (30% VASP). Tier 3 sources were too constrained
- The aim of the installation of Thomas was to rebalance the ecosystem for the MMM community
- It has created a UK-wide Hub for MMM that serves the entire UK MMM community
- The Hub will build a community to foster collaborative research and the cross-fertilisation of ideas
- Support and software engineering training is offered



## Partner Institutions



- Initial bid development driven by the Thomas Young Centre: The London Centre for the Theory and Simulation of Materials: **Imperial, King's, QMUL and UCL**
- During development of the full bid team expanded to include: **Belfast, Kent, Oxford, Southampton**
- Funding for operational and hosting costs derived from across the partners: **>£2m over initial term**
- National MMM-HPC Consortia: **Materials Chemistry Consortium** and **UKCP**
- Thomas is hosted at the Virtus Data Centre in Slough
- **UCL** is providing first-line support
- **OCF-Lenovo** contributing to Hub Coordinator post



# Thomas Architecture

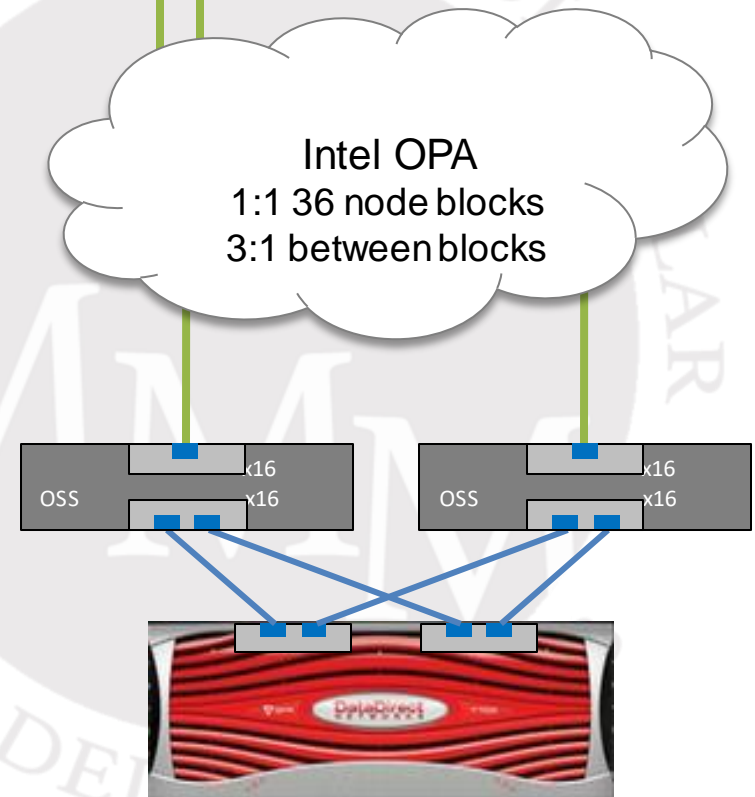
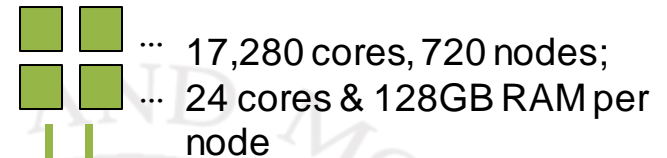
- 720 Lenovo Intel x86-64 Broadwell nodes
- Intel OmniPath interconnect
- DDN Lustre filesystem

## Technical performance

- 523 Tflop/s Rmax:  
[www.top500.org/system/178941](http://www.top500.org/system/178941)
- 5.5 GiB/s IO bandwidth

## Real world performance

- LAMMPS Rhodopsin benchmark, ONETEP 864 core jobs
- $2.2 \pm 0.3$  speedup over UCL Grace system
- Improvements down to OmniPath interconnect

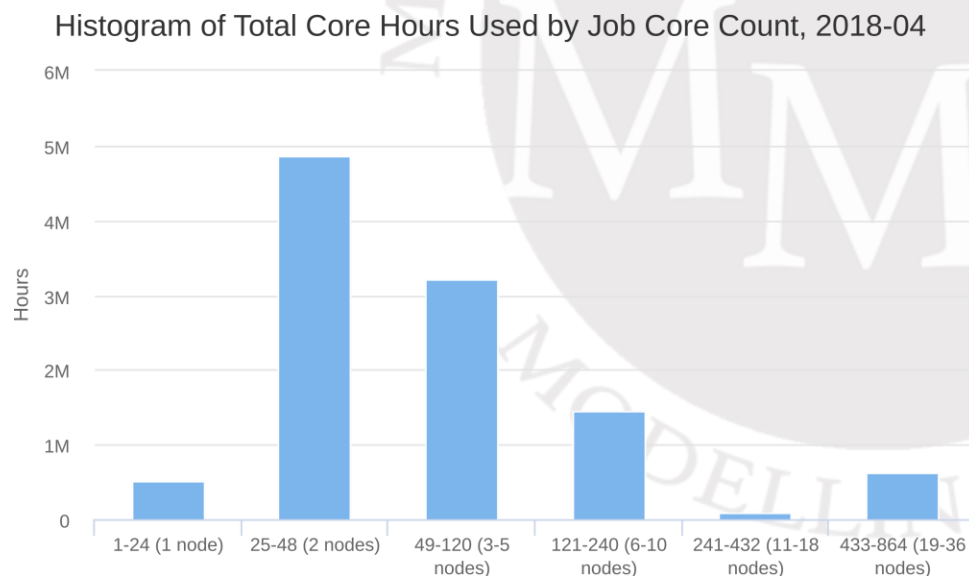


- Thomas has access to (most) of UCL Research Computing's standard software stack
  - Caveat: licensed software may not all be available
- Buildscript repository:
  - <https://github.com/UCL-RITS/rcps-buildscripts>
  - Allows us to rebuild software easily on all our clusters
  - We do nothing as root – users can use our scripts for local installs
  - Install request for one cluster benefits users on the others
- Before Thomas went into service, surveyed our partners and built requested software available on ARCHER but not in our stack yet
  - Helped us get up and running quickly!

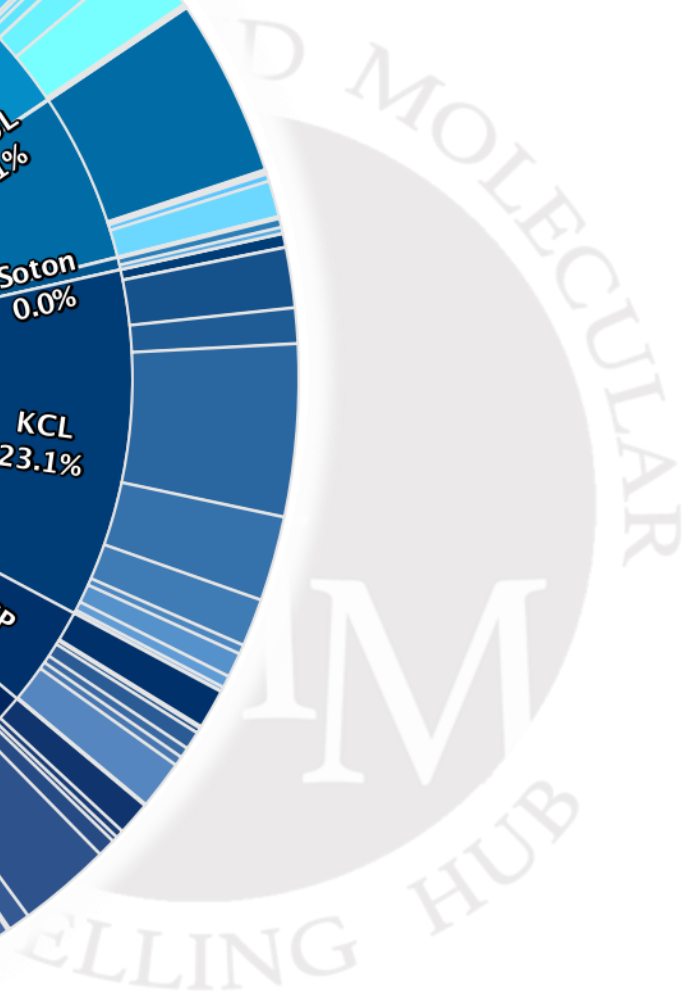
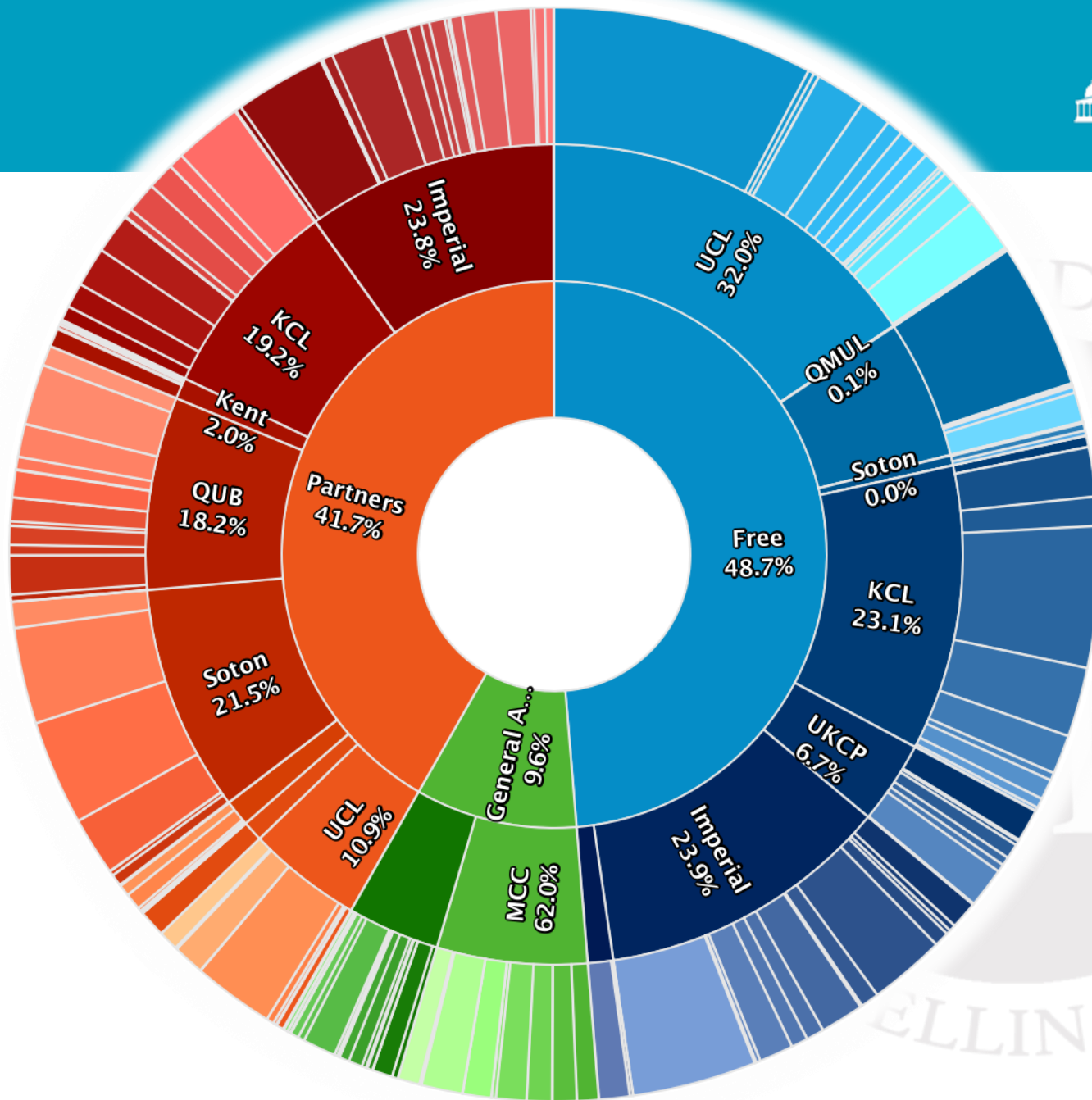
- 75% of machine cycles are available to the university partners providing funding for Thomas' hosting and operations costs
- Each partner has a point of contact in charge of approving accounts and distributing their allocation according to their own rules
- 25% of cycles are available to the wider UK MMM Community
- Access to this 25% managed by two consortia MCC and UKCP (not T2 RAC)
- If you want an account:
  - [mmmhub.ac.uk](http://mmmhub.ac.uk)
- Please acknowledge use of Thomas in papers and presentations!

- Partners have made commitments for RSE time
  - Per-institute basis
- UCL Research Software Development, led by Dr James Hetherington:
  - Programming & tuning support for UCL researchers targeting Tier 2 platforms (not just Thomas)
  - One FTE provided for all UCL Tier 2 users, from across the RSD group
  - Following existing model for UCL:
    - Specific programming projects costed and funded from grants
    - General advice, support, training for free

- Thomas went in to service on 5 June 2017 after a successful pilot
- Almost a year in: ~440 users, 160 active in any given month
- SAFE integration progressing
- Target job size: 2-5 nodes (48-120 cores).
  - Prioritised over other job sizes
- Maximum job size: 864 cores







## Questions?

- Thanks to EPSRC for the funding and the partners and points of contact for all their hard work!

