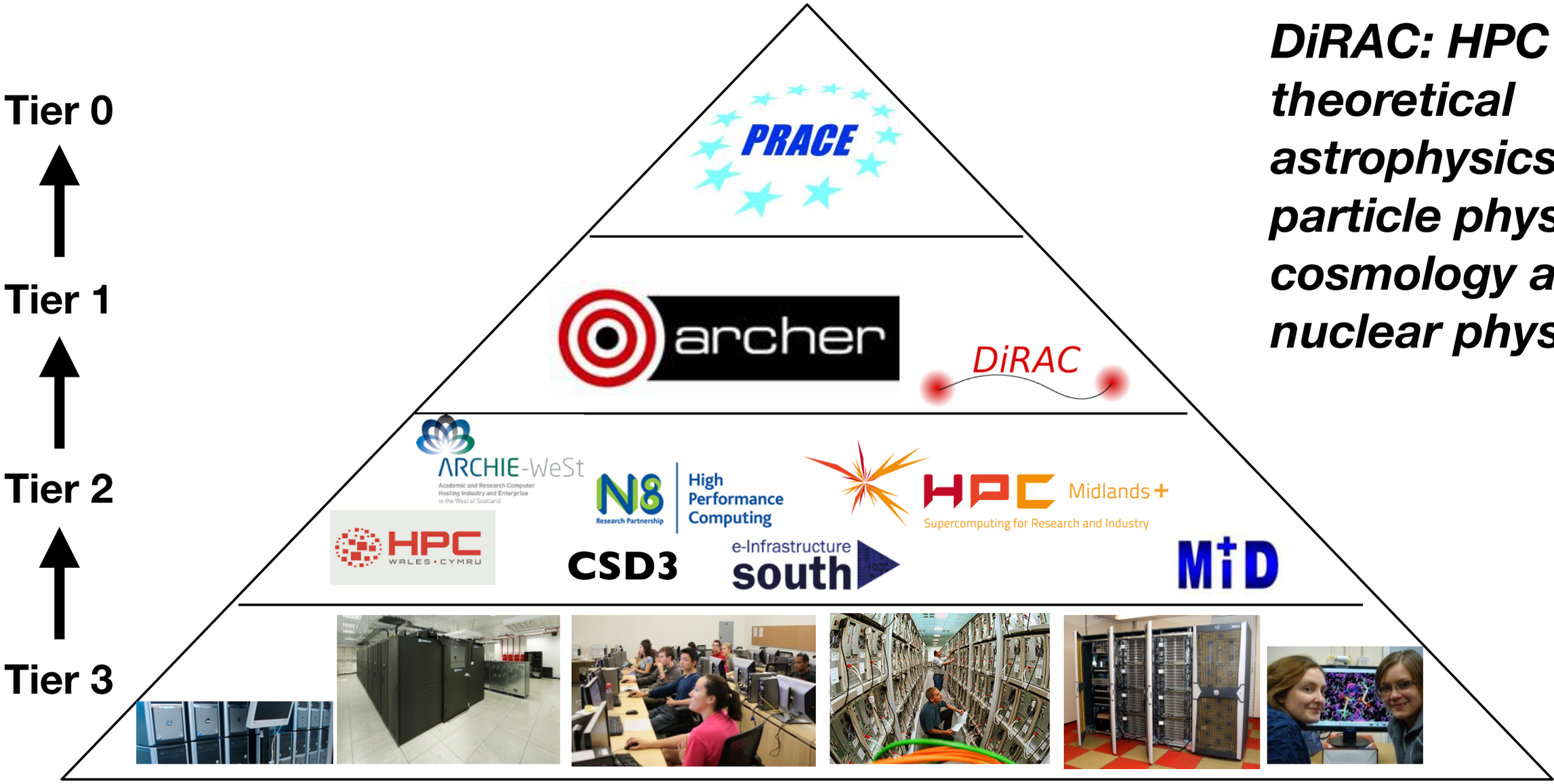
Research Software Engineering: The DiRAC Facility Experience

Mark Wilkinson, Director

DiRAC

Science & Technology Facilities Council

Research Computing in the UK



DiRAC: HPC for theoretical astrophysics, particle physics, cosmology and nuclear physics

- **Research software engineering effort facilitates movement between tiers and across services within tiers**
- **Optimises use of all available resources**

DiRAC 2.5x

Diverse science cases require heterogenous architectures

Service	Description	Available
Extreme Scaling (Edinburgh)	20256 cores; 0.97 PF	1st May 2018
Memory Intensive (Durham)	Max RAM footprint 114TB; 15428 cores; 0.38 PF	Partial until 1st May 2018
Data Intensive (Cambridge, Leicester)	67 TF Xeon Phi; Max job size 500 TF	✓
	130 TFlop/s GPU; Max job size 1 PF	✓
	17000 Xeon cores; 0.5 PFlop/s	Partial until 1st May 2018
	3x 1.5TB RAM nodes; 1x 6TB (144 core) Superdome shared-memory server	1st May 2018

- Total ~2 PFlop/s across all services

Research Software Engineers

computational algorithmic

"With great power comes great responsibility"

David Keyes

- Science requirements for DiRAC-3 demand 10-40x increases in computing power to stay competitive
  - hardware alone cannot deliver this
- We can no longer rely on "free lunch" from the Xeon era
- Vectorisation and code efficiency now vital
- Current and next generation hardware more difficult to program
  - Vendors are putting more demands on users
- Research Software Engineers are increasingly important
  - RSEs can help with code profiling, optimisation, porting, etc

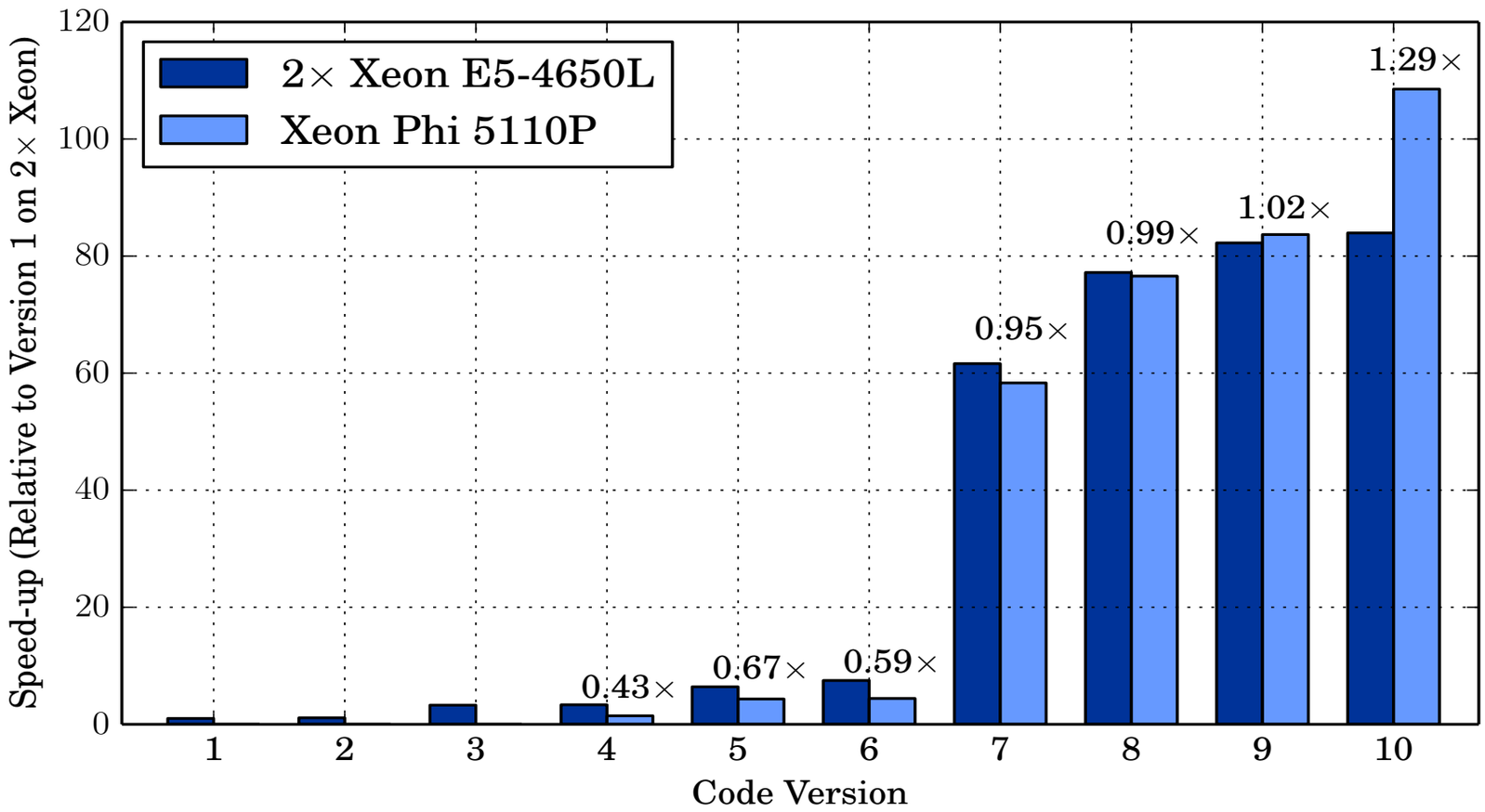
DiRAC Software engineering activities

- RSE effort:
  - 3 FTE available - embedded within University teams
  - Allocated via peer review process: details at dirac.ac.uk
- Training workshops: Many-Core programming; Software Design & Optimisation; MPI programming
- Three Intel Parallel Computing Centres:
  - COSMOS: MODAL, GRChombo, OSPRay
  - Edinburgh: GRID
  - Durham: Swift (see Matthieu Schaller's talk)
- Hackathons:
  - NVidia hackathon before DiRAC Day, Sept 2018
- Support for STFC CDTs in Data Intensive Science
- Focus on library development to maximise impact
- Enables use of new hardware and provides greater flexibility in future procurements

DiRAC Software Innovation

MODAL XEON PHI MODERNIZATION (COSMOS IPCC)

- **Multi-/many-core optimisation: 100-1000x speed-ups**
  Enabling new science - required significant RSE effort
- HPC publications; Xeon Phi course (EPCC); ISC'15 Lecture;
- Winner of HPCWire Readers' Choice award 2015

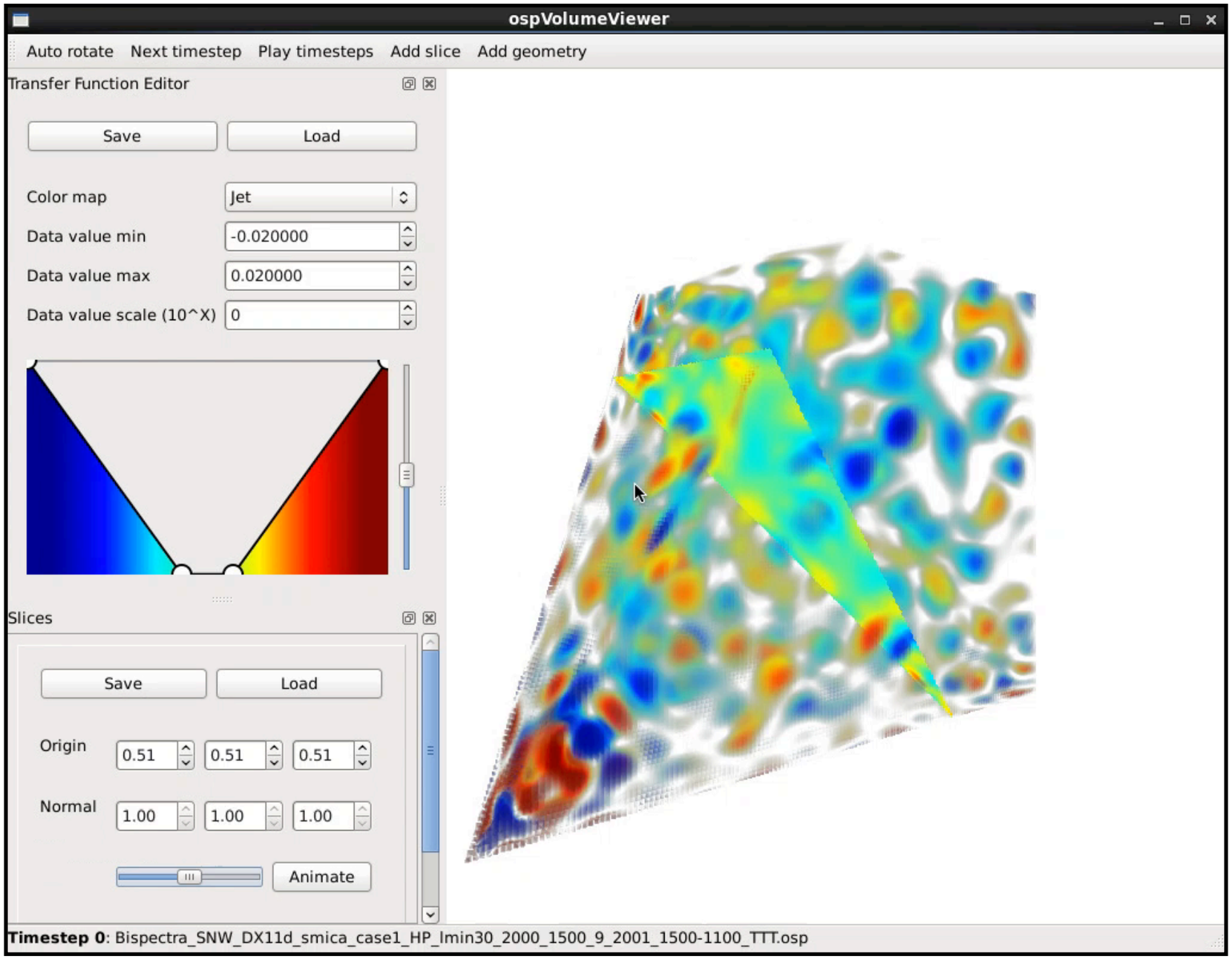


Modal3D optimizations

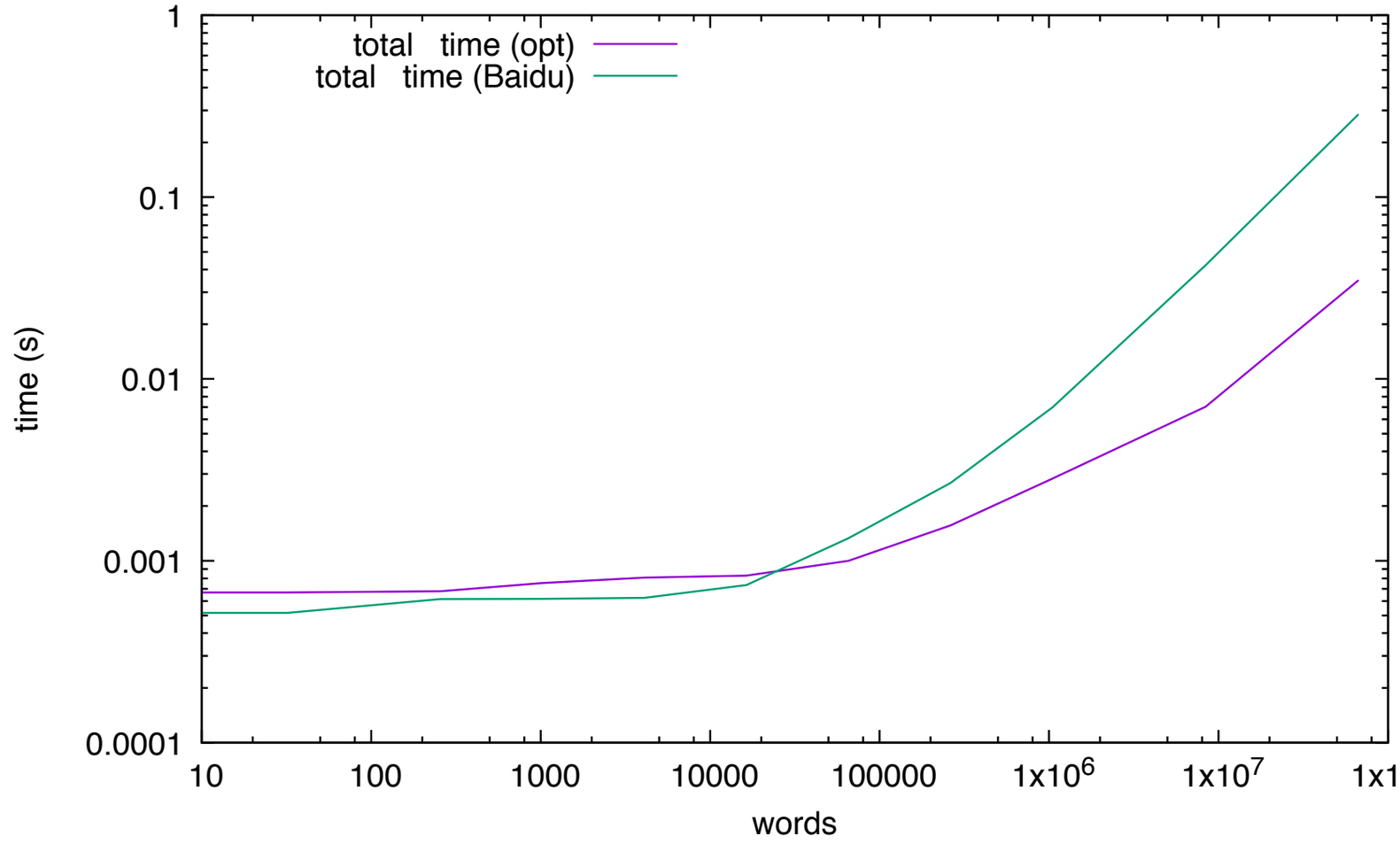
DiRAC Software Innovation

- **OSPRay XEON PHI Viz - collaboration between COSMOS IPCC and Intel**

Demonstrator of remote visualisation on Xeon Phi



Software Innovation - AI on HPC



Boyle et al. 2017

- Demonstration of factor 10 speed-up in the Baidu Research optimised reduction code
  - a publicly available code designed to optimise the performance limiting steps in distributed machine learning
- potentially disruptive implications for design of cloud systems
  - shows that ML workflows can achieve 10x performance improvement when fully optimised and implemented on traditional HPC architectures

DiRAC Training

- DiRAC provides ***access*** to training from wide pool of providers
- Currently offering:
  - DiRAC Driving Test: now available online (and compulsory!)
  - Workshops: Many-Core programming; Software Design & Optimisation; MPI programming
- Under development:
  - Domain-specific workshops
  - Online individual training portal
  - Industry-focussed training pathways

Why do we do this?

- maximise DiRAC science output
- flexibility to adopt most cost-effective technologies
- future-proofing our software and skills
- contributes to increasing skills of wider UK economy

Conclusions

- Goal is to maximise the science you can do on DiRAC
- Research software engineering is increasingly important
- DiRAC provides access to training and RSE effort from wide pool of providers
- Do engage with these opportunities:
  - Greater flexibility in future procurements means DiRAC-3 can be more powerful and productive for your science
- Tell us if we have missed something:

  dirac.ac.uk

  mark.wilkinson@leicester.ac.uk