Science and Technology Facilities Council

Conference Proceedings
STFC-CONF-2025-001

Computing Insight UK 2024

Manchester Central Convention Centre, UK
5th-6th December 2024

G Lomas, D Jones (editors)

February 2025

Enquiries concerning this report should be addressed to:

RAL Library
STFC Rutherford Appleton Laboratory
Harwell Oxford
Didcot
OX11 0QX

Tel: +44(0)1235 445577
email: library@stfc.ac.uk

Science and Technology Facilities Council reports are available online at:
https://epubs.stfc.ac.uk

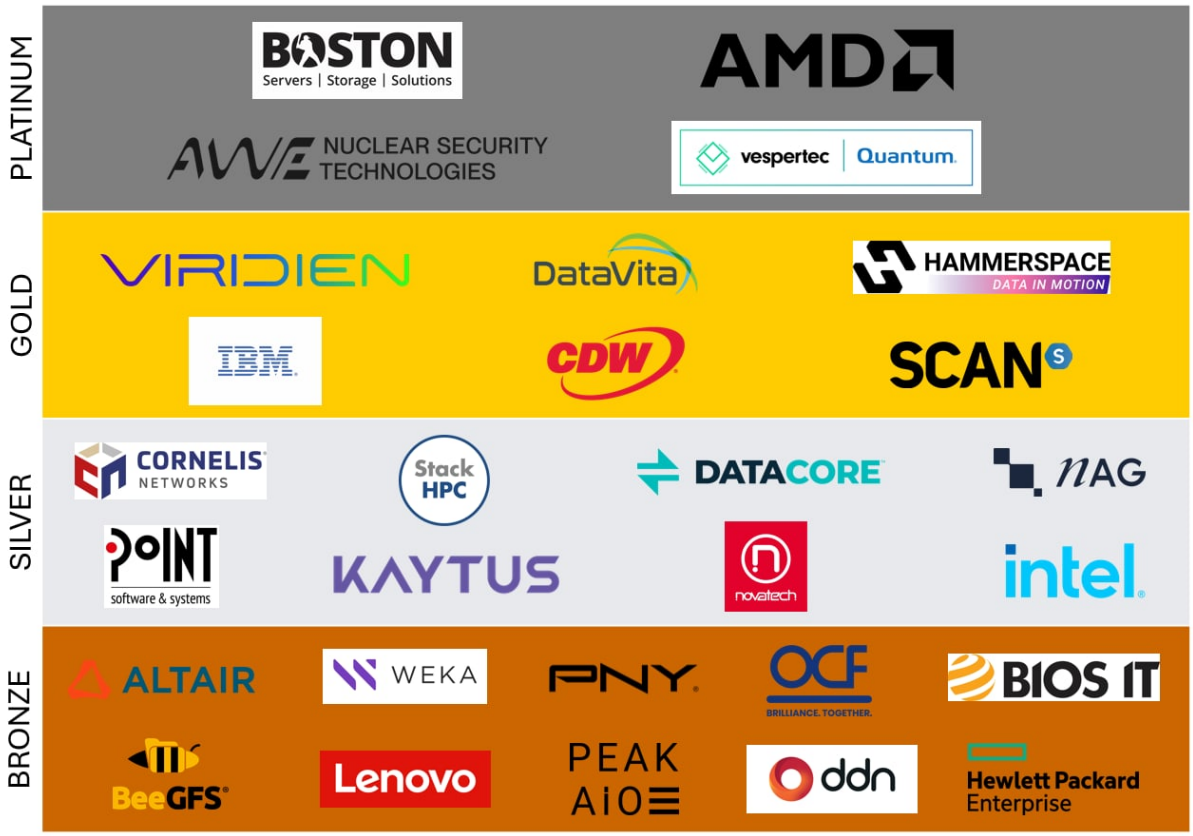
Accessibility: a Microsoft Word version of this document (for use with assistive technology) may be available on request.

DOI: 10.5286/stfcconf.2025001

ISSN 2753-5800

Computing Insight UK (CIUK) 2024 took place on and 5th and 6th of December 2024 at the Manchester Central Convention Centre. These proceedings are a record of the presentations and posters from the Conference.

The CIUK Organising Committee would like to thank the exhibitors, sponsors, presenters and attendees who help to make the Conference a continued success.



Computing Insight UK 2024 Introduction

Computing Insight UK (CIUK) 2024 was the 35th edition of an annual conference organised by the Science and Technology Facilities Council's (STFC) Scientific Computing Department (SCD). The event was held on the 5-6 December at Manchester Central Convention Complex and attracted a record crowd of over eight hundred attendees.

The theme for this year's conference is "Catalysing Research", with sub-themes including "Sustainability", "AI Focus", "UKRI DRI and DSIT Update", "Technology Trends" and "Collaboration".

CIUK 2024 included an exhibition of the latest hardware and software releases plus a full, two day programme of presentations and a series of parallel breakout sessions. There was a poster competition, won by Zeyuan Miao for their poster on " Thermal Modeling in Neutronics Using Physics-Informed Neural Networks: Leveraging HPC for Surrogate Model Training", and we also presented our annual Jacky Pallas Memorial Award, which this year was awarded to Lisa Lampunio (University of Pisa) for her work on "Advanced Modelling and Simulation for the Analysis of Novel Radiation Technology and Thermal Fatigue Phenomena within the Nuclear Energy Sector". Lisa presented her work as part of the main programme during the conference.

The CIUK 2024 Keynote Presentation was given by Luigi Del Debbio from The University of Edinburgh. Luigi's presentation was titled "Visions of Computing".

Following the success of "Day Zero" at CIUK 2023, it was brought back for its second year - a pre-CIUK day that included the CoSeC Annual Conference, the Lustre User Group meeting and two hands-on training courses on " Technical/SysAdmin Meetup" and " STEP-UP - developing HPC technical professionals".

CIUK 2024 also saw the fifth edition of the CIUK Cluster Challenge competition with an incredible fifteen teams entering from University of Bath, Imperial College London, University of Birmingham, Durham University, University of Bristol, University of Plymouth, University College London, University of Warwick, Kings College London, University of York, The Rosalind Franklin Institute and Manchester University. The teams completed four online challenges leading up to the conference, followed by four challenges during the conference in Manchester. Team Decarb from Warwick University and Imperial College London took the title after a closely fought competition and earned their place at the ISC'25 Cluster Challenge competition where they will represent CIUK against the best student teams from around the word.

Computing Insight UK 2024 *"Catalysing Research"*

CIUK DAY ZERO - Wednesday 4 December 2024

TIME	PRE-CIUK ACTIVITIES These activities will take place in the CIUK Breakout Room and the meeting rooms upstairs		
From 08:00	**CIUK 2024 EXHIBITION SET-UP (Exchange Hall)**		
09:30 - 10:00			
10:00 - 10:30	CoSeC Computational Science Centre for Research Communities **Annual Conference 2024** Wednesday 4 December @ CIUK 2024 **10:00 - 17:00** **(CIUK Breakout Room)** **Registration from 09:30** **www.scd.stfc.ac.uk/CoSeC**		
10:30 - 11:00			
11:00 - 11:30			
11:30 - 12:00			
12:00 - 12:30			
12:30 - 13:00			
13:00 - 13:30		**Technical/SysAdmin Meetup** **13:00 - 16:00** **(Exchange Room 1 upstairs at Manchester Central)**	
13:30 - 14:00			
14:00 - 14:30	LUG UK **Lustre User Group 2024** **Exchange Rooms 9 and 10 upstairs at Manchester Central**		COMPUTING INSIGHT UK 2024 Catalysing Research 5 - 6 DECEMBER 2024 Manchester Central, UK www.ukri.org/CIUK **CIUK 2024 Student Cluster Challenge (Exchange Hall)** OCF BRILLIANCE. TOGETHER.
14:30 - 15:00			
15:00 - 15:30			
15:30 - 16:00			
16:00 - 16:30		**STEP-UP - developing HPC technical professionals** **16:00 - 17:00** **(Exchange Room 1 upstairs at Manchester Central)**	
16:30 - 17:00			
17:00 - 17:30			
17:30 - 20:00	**CIUK 2024 Day Zero Networking Event - The Gas Works Brew Bar, 5 Jack Rosenthal St, Manchester M15 4RA** **(Open to all registered CIUK attendees and exhibitors - CIUK lanyard required for entry)**		
Until 20:00	**CIUK 2024 EXHIBITION SET-UP (Exchange Hall)**		

Computing Insight UK 2024 *"Catalysing Research"*

Main Programme Session Themes

	Session 1: Sustainability Chair: Joseph Thacker (UKRI-STFC)
	Session 2: AI Focus
	Session 3: UKRI DRI Update

CIUK DAY ONE - Thursday 5 December 2024

TIME	MAIN PROGRAMME	BREAKOUT SESSIONS	CIUK CLUSTER CHALLENGE
From 08:30	**REGISTRATION OPEN (Exchange Foyer) CIUK 2024 EXHIBITION OPEN (Exchange Hall)**		
09:15 - 09:30	*Welcome and Introduction* **Tom Griffin (Director, Scientific Computing, STFC)**		
09:30 - 10:00	**Carter Quinn (University College London)** *Creating a Staffing Pipeline for the Research Infrastructure Profession with Apprenticeships*	**Portable benchmarking and profiling using ReFrame** **9:30 - 13:00** **(CIUK Breakout Room)**	
10:00 - 10:30	**Sadie Bartholomew (University of Reading/NCAS)** *Reducing the impact of energy consumption from computing with CATS, The Climate Aware Task Scheduler*		
10:30 - 11:00	**Jessica Huntley (UKRI-STFC)** *Tracking the Carbon Cost of Optimization Algorithms*		**CIUK 2024 Student Cluster Challenge (Exchange Hall)**
11:00 - 11:30	**REFRESHMENTS**		LOGICALIS Architects of Change
11:30 - 12:00	**Vassil Alexandrov and Adriano Agnello (STFC-Hartree Centre)** Harnessing AI to Address Grand Challenges in Scientific Innovation		
12:00 - 12:30	**Noel O'Connor (Red Hat)** *Partnering to take LLM to the next level*		
12:30 - 13:00	**Simon McIntosh-Smith, Priya Sharma and Chritsopher Edsall** *The National AI Research Resource (AIRR) – The journey to launch*		
13:00 - 14:00	**LUNCH**		
14:00 - 14:30	**Richard Gunn (UKRI DRI)** *UKRI DRI Update* **Followed by a panel Q&A discussion**	**Cybersecurity and Federation for National DRI, AI and HPC Resources** **14:00 - 16:00** **(CIUK Breakout Room)**	
14:30 - 15:00			**CIUK 2024 Student Cluster Challenge (Exchange Hall)**
15:00 - 15:30			ddn THE AI DATA COMPANY
15:30 - 16:15	**REFRESHMENTS**		
16:15 - 17:00	**Martyn Guest (ARCCA, Cardiff University)** *Community Code Performance on Multi-core Processors. An Analysis of Computational Chemistry and Ocean Modelling Applications.*		
17:00 - 18:00	**CIUK 2024 Keynote Presentation** **Luigi Del Debbio, The University of Edinburgh - *Visions of Computing***		
18:30 - 23:00	**CIUK 2024 Networking Event - Peaky Blinders Bar, Peter Street, M2 5QR** **(*CIUK 2024 lanyard and badge required for entry*)**		

Computing Insight UK 2024 *"Catalysing Research"*

Main Programme Session Themes

	Session 4: Technology Trends Chair: Omar Ahmed Mahfoze (UKRI-STFC)
	Session 5: Collaboration
	Session 6: Test Bed Panel Chair: Stephen Longshaw

CIUK DAY TWO - Friday 6 December 2024

TIME	MAIN PROGRAMME	BREAKOUT SESSIONS	CIUK CLUSTER CHALLENGE
From 08:30	**REGISTRATION OPEN (Exchange Foyer)** **CIUK 2024 EXHIBITION OPEN (Exchange Hall)**	WHPC WOMEN IN HIGH PERFORMANCE COMPUTING **Women in HPC Breakfast** **08:30 - 10:30** **(CIUK Breakout Room)**	
09:30 - 10:00	**Mayank Kumar (UKRI-STFC)** *Integration of SYCL with MPI in Multi-scale Universal Interface (MUI) library for Multi-Physics Coupling at Exascale*		
10:00 - 10:30	**Gregory Tourte (University of Oxford)** *Use of WEKA FS for fast ephemeral storage powering academic HPC*		 **CIUK 2024 Student Cluster Challenge (Exchange Hall)** alcesflight
10:30 - 11:00	**Oliver Brown (EPCC, University of Edinburgh)** *The Scalability of Quantum Air Traffic Control*	**Storage Scale (GPFS) User Group / New Users Session** **10:30 - 13:30** **(CIUK Breakout Room)**	
11:00 - 11:30	**REFRESHMENTS**		
11:30 - 12:00	**Duncan McBain (Codeplay Software)** *Portable SYCL code using oneMKL on AMD, Intel and Nvidia GPUs*		
12:00 - 12:30	**Dimitrios Bellos (The Rosalind Franklin Institute)** *Flowcron: Improving access to HPC via a Function-as-a-Service which utilizes a Globus Flow and a cron service*	**ExCALIBUR Hardware and enabling software: Exploring next generation technologies for HPC** **11:00 - 13:00** **(Exchange Room 1 upstairs at Manchester Central)**	
12:30 - 13:00	**Dave Bond, Andy Herdman, Neil Gaspar (AWE National Security Technologies)** *AWE - Centres of Excellence, an outline of academic engagement*		
13:00 - 14:15	**LUNCH**		
14:15 - 14:30	**Awards Presentation** **The CIUK 2024 Student Cluster Challenge and CIUK 2024 Poster Competition**	**UKRI National Federated Compute Services NetworkPlus** **14:00 - 15:00** **(CIUK Breakout Room)**	
14:30 - 15:00	**The Jacky Pallas Memorial Presentation - Lisa Lampunio** *Advanced Modelling and Simulation for the Analysis of Novel Radiation Detection Technology and Thermal Fatigue Phenomena within the Nuclear Energy Sector*		
15:00 - 15:30	**Alastair Basden (DiRAC/Durham University)** *The Durham HPC Hardware Laboratory* **Nick Brown (EPCC, University of Edinburgh)** *RISC-V: The most exciting hardware innovation of recent times*		
15:30 - 16:00	**Tina Friedrich (University of Oxford)** *A first look at JADE 2.5*		
16:00	**CIUK 2024 CLOSES - See you in 2025!!!**		

CIUK 2024 Presentations

Carter Quinn (university College London)

Creating a Staffing Pipeline for the Research Infrastructure Professions with Apprenticeships

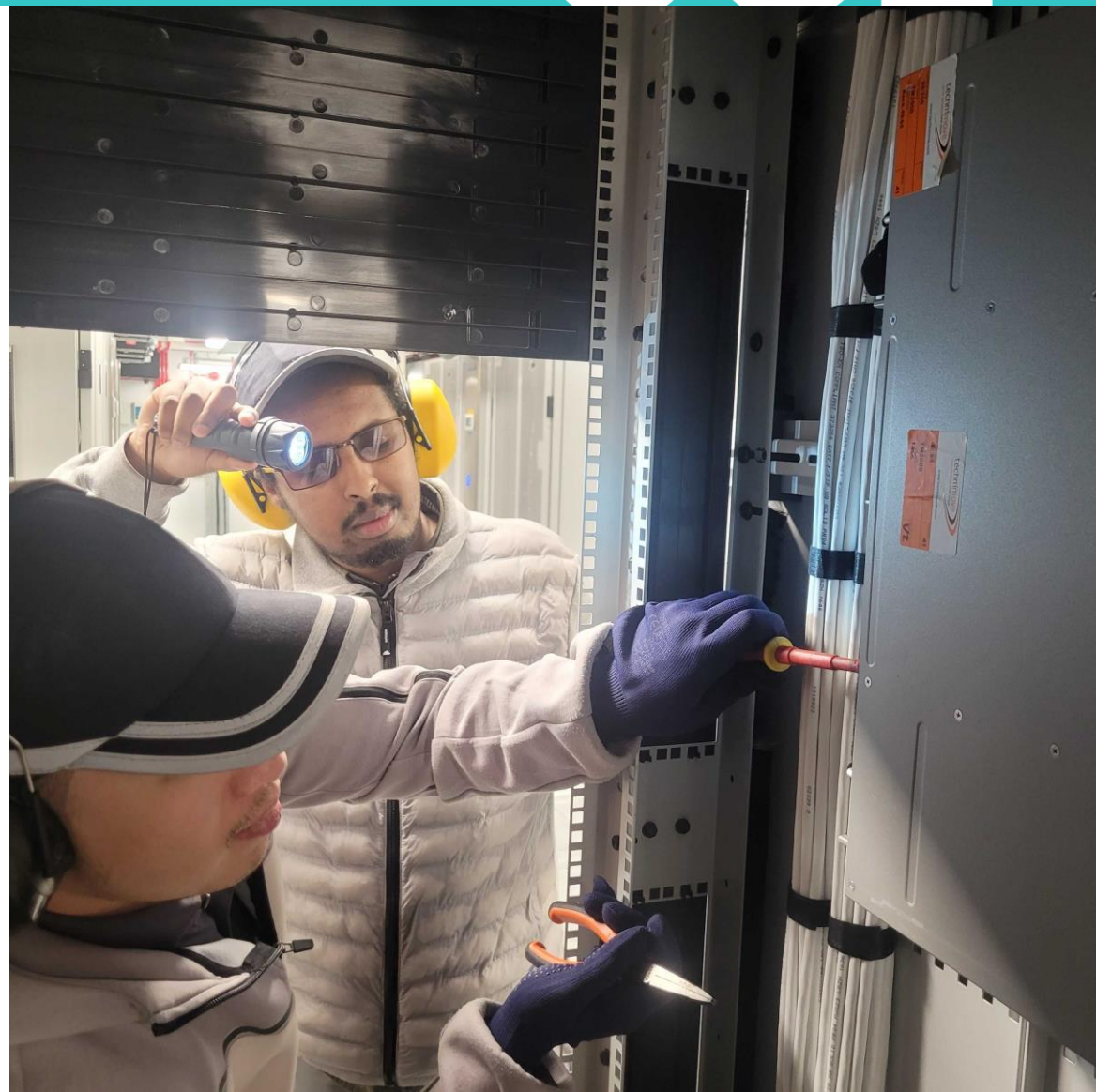
Abstract: Within the research computing infrastructure community, staffing shortages are becoming more pronounced, producing top-heavy structures, where insufficient people are filling the entry level roles. This imbalance places increasing pressure in the short-term on teams and a future staffing crisis in the long-term. Partially responsible is the lack of a direct pipeline to create junior staff. In this presentation, we explore UCL's experience piloting an apprenticeship scheme in this field and how it may be applied community wide.

Bio: Carter initially began a career in IT through self-employment during sixth form and moved into salaried positions for various employers. In 2021, he was exposed to HPC in research through a contact at UCL and interviewed for a position in UCL – ARC. Having been successful, about a year later, was invited to become involved in the department's pilot apprenticeship scheme due to his unique experience.

University College London – Centre for Advanced Research Computing

Creating a Staffing Pipeline for the Research Infrastructure Professions with Apprenticeships

Carter Vernon Quinn



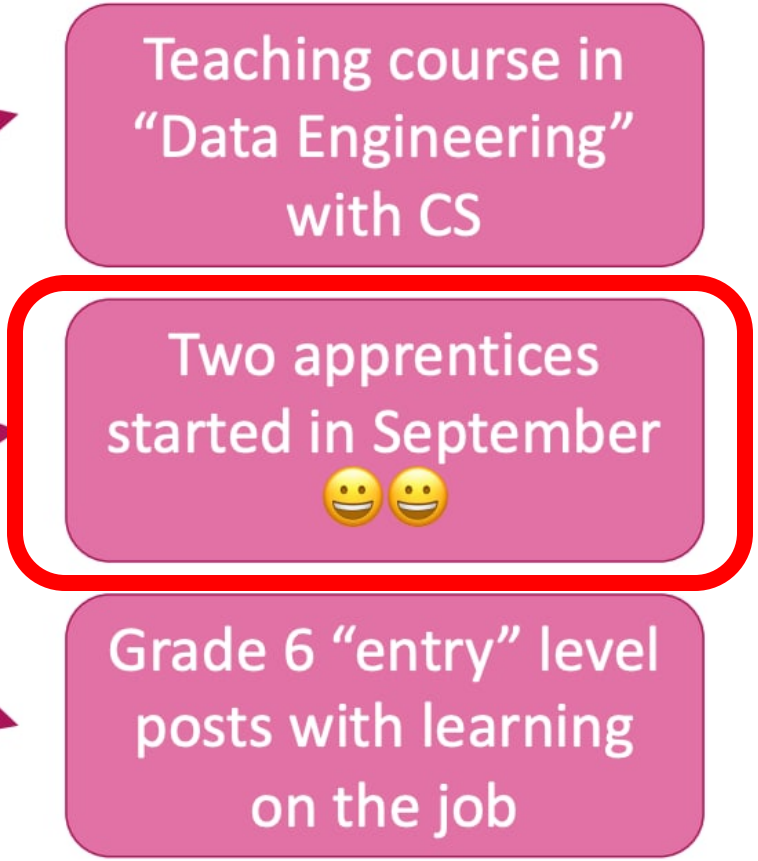
A Follow On…


UCL
Bridging the Professional Services/Academic gap
Creating a hybrid Academic/Professional structure for research IT professionals at UCL (and where you are!)
Dr Owain Kenway, (@owainkenway)
UCL/ARC/Head of Research Computing
@CompInsightUK #CIUK2023
/in/computing-insight-uk
COMPUTING INSIGHT UK

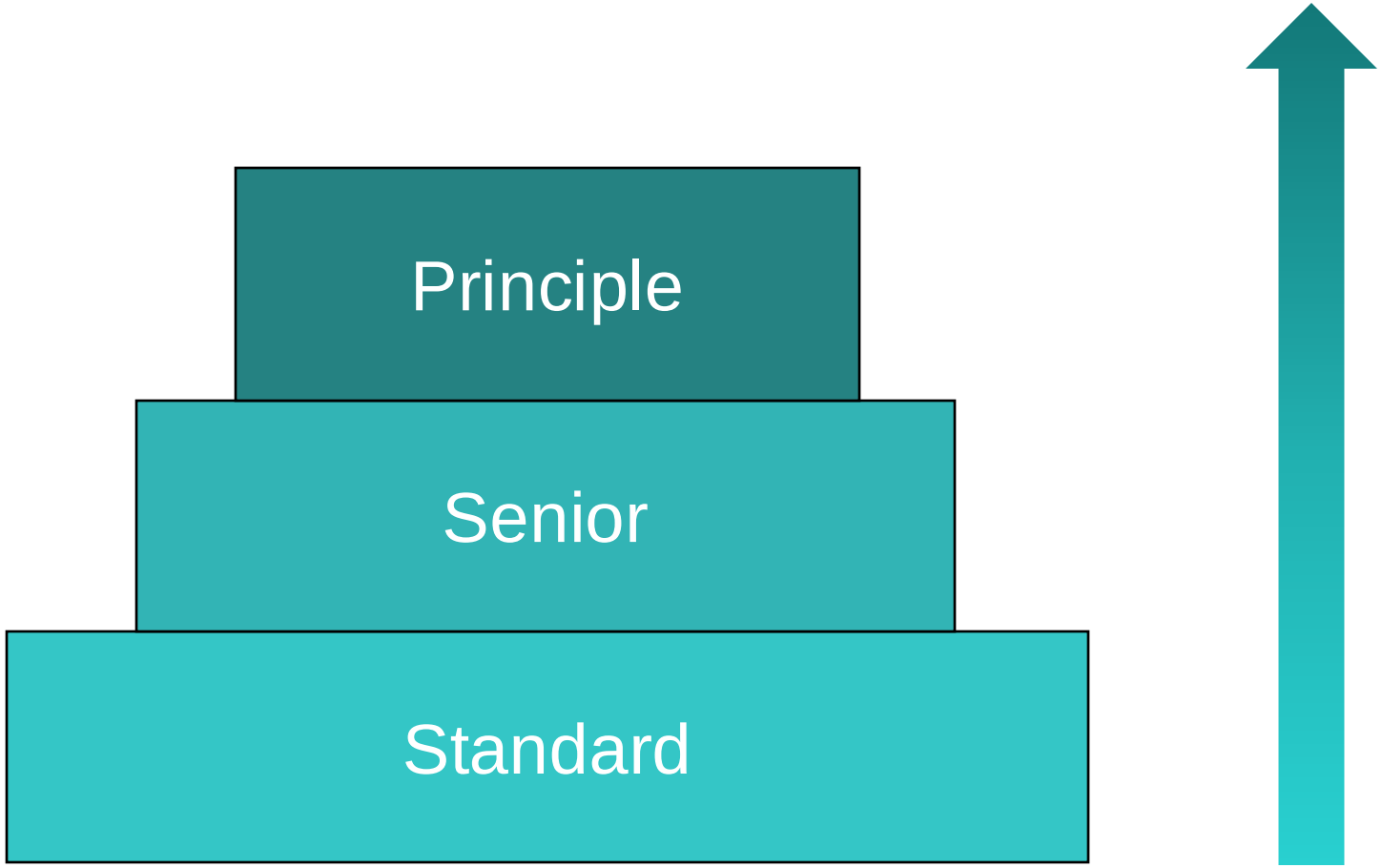
A Follow On…

Creating future employees

- Multi-track approach:
  - Get involved in undergrad teaching -> build skills academically
  - Run apprenticeships -> academia is not the only route into becoming a "research IT professional"
  - Recruit skilled people who have the right attitude and can learn

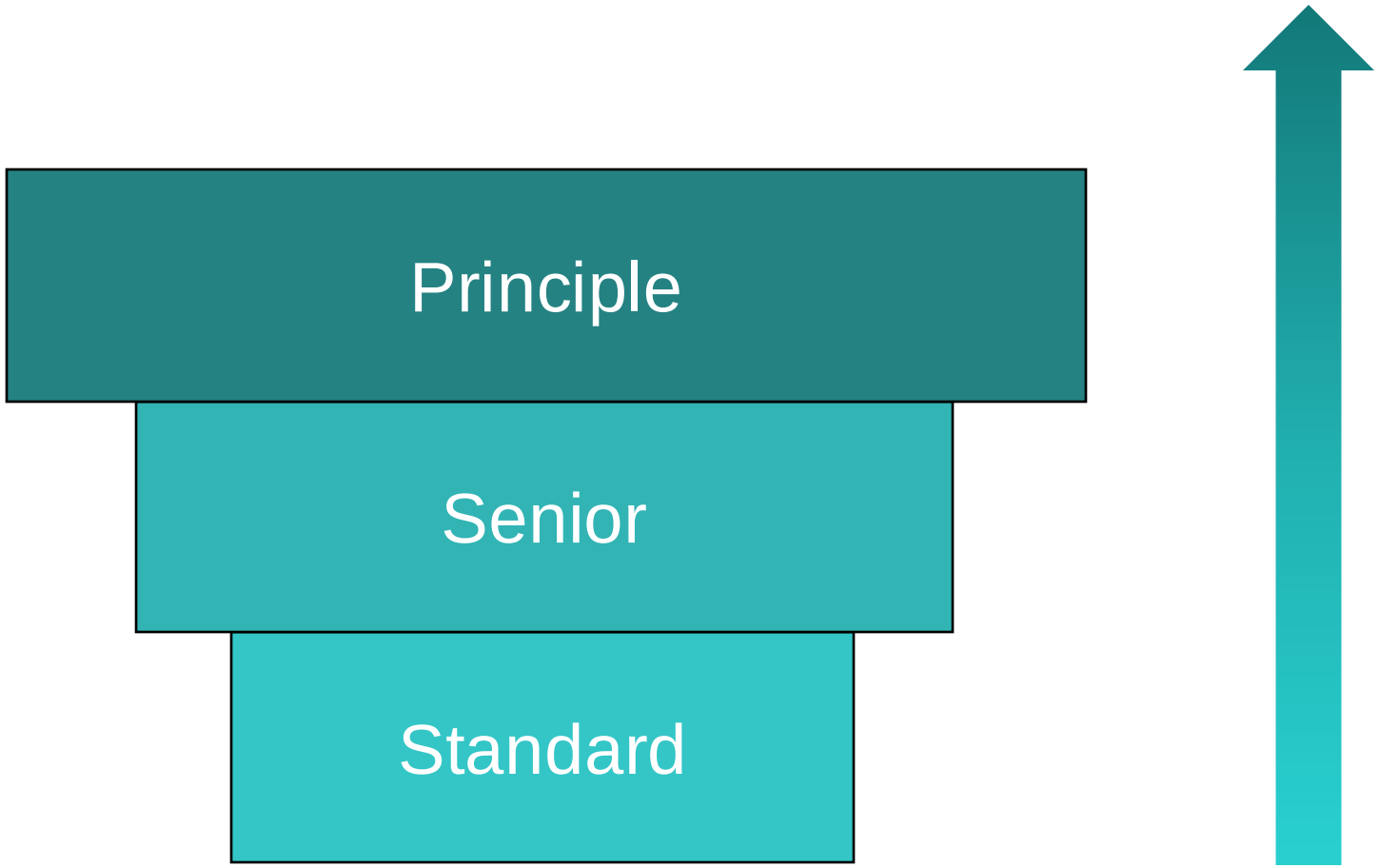


The Ideal Situation



Seniority of Staff
Within a
Profession

The Current Situation



Seniority of Staff Within a Profession

Causes

Intake of Staff:

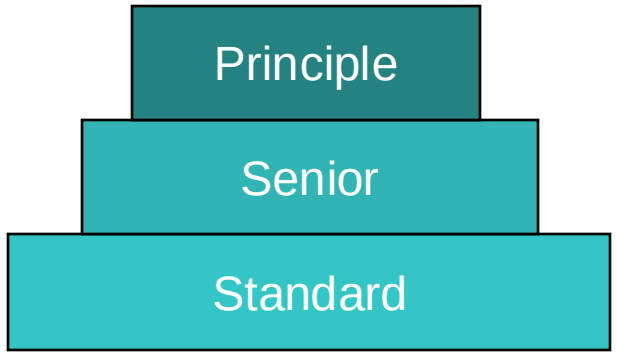
- *No dedicated pipeline for sysadmins and infrastructure professionals.*
- *Drying up of traditional sources of staff.*

Retention of Staff:

- *Competitiveness of research/academia in relative terms to wider industry.*
- *Institutional relations with employees.*

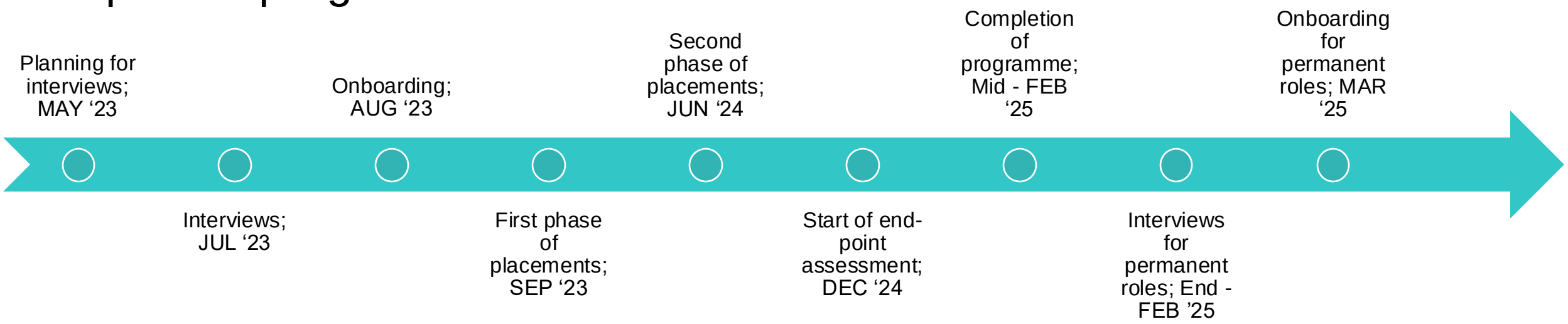
The Place Apprenticeship Programmes Fill

- Relieving pressure at senior levels of staff by allowing for more efficient use of their time.
- Filling in the gaps created by the drying up of traditional staffing sources at the junior end.
- Widen the conventional recruiting pools.
- Offset the pull that industry has.



Considerations for the Programme

- How long to spin up the programme/deliver results?
- What scale is the programme expected to operate?
- What support is required?
- Expected programme costs?



Recruitment and Selection

- Joint process with our central IT.
- Standard two-round interview process with aptitude assessment.
- The aptitude assessment was written without a pass mark in mind.
- Deliberately open to a broad backgrounds/disciplines.

The Two Strands of the Apprenticeship

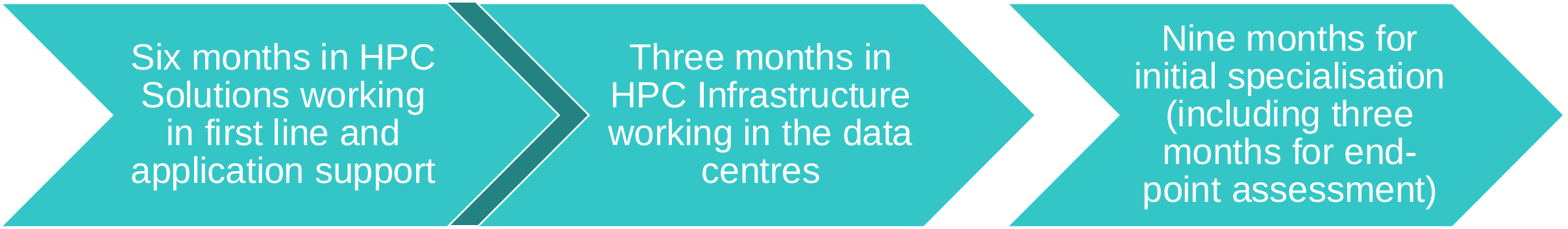
On-the-Job:

The training we provide to the apprentices to be successful in their specific future roles and create a groundwork for future professional development.

Off-the-Job:

The training and structure implemented by the training provider to deliver training apprentices may not receive on the job, to develop soft skills required in the workplace and provide progress monitoring.

On-the-Job



- First nine months represent initial phase of core training
- Second nine months are there to prepare for future career development.

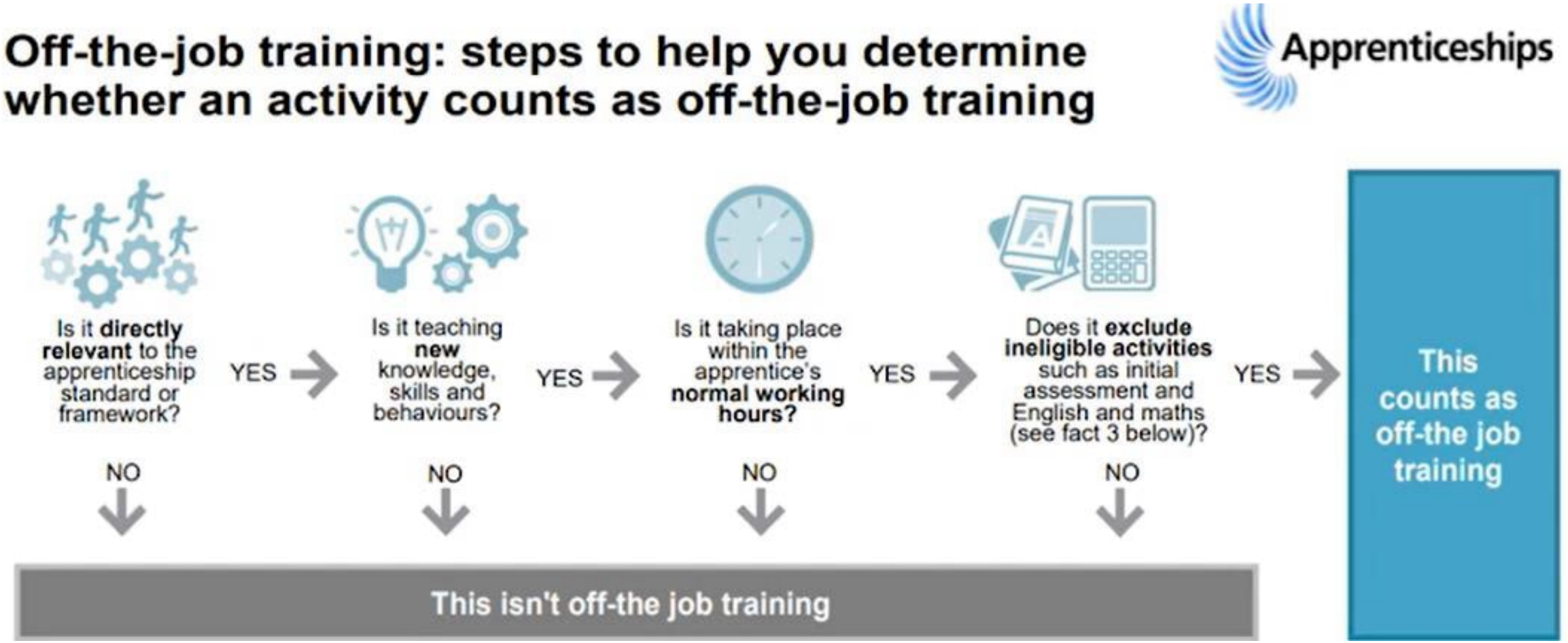
Off-the-Job

15 months (one day a week) of theoretical work and portfolio development.

Three-month end-point assessment including a project, exams and discussion about the portfolio.

- Theoretical training to complement the on-the-job training and help provide the basis for future development.
- Time to turn project work done on the job into evidence for assessors to demonstrate learnt skills and behaviours from experience with their employers.

Off-the-Job


Off-the-job training: steps to help you determine whether an activity counts as off-the-job training
Apprenticeships
Is it directly relevant to the apprenticeship standard or framework?
YES
Is it teaching new knowledge, skills and behaviours?
YES
Is it taking place within the apprentice's normal working hours?
YES
Does it exclude ineligible activities such as initial assessment and English and maths (see fact 3 below)?
YES
This counts as off-the job training
NO
NO
NO
NO
This isn't off-the job training

Take Aways from the Programme

Lessons Learnt:

- *Having two apprentices instead of one was better for mutual support.*
- *We ended up training to a standard where hiring is planned at a higher grade than was originally planned for.*
- *When seconding apprentices between professions, it served as an opportunity for community building in our department.*

Improvements for Next Time:

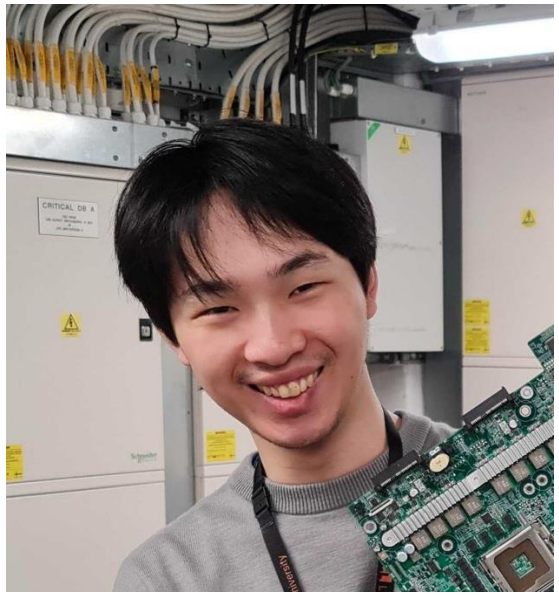
- *Project work has been denser than in other IT apprenticeships and we should plan to accommodate the added pressure/burden to learn.*
- *Get a better start on handling the administrative side from the outset.*

Conclusions

- The scheme has demonstrated potential as a part of the long-term solution to staffing shortages.
- We were able to find work accessible to apprentices that allow them to positively contribute to the team.
- The existing standard we used off-the-shelf works for the role that they are being trained to speeding up the process of creating apprenticeships.

Questions

Carter Quinn
carter.quinn@ucl.ac.uk



Daniel Chau
daniel.chau@ucl.ac.uk

Mohamed Guled
m.guled@ucl.ac.uk

CIUK 2024 Presentations

Sadie Bartholomew (University of Reading and National Centre for Atmospheric Science (NCAS))

Reducing the impact of energy consumption from computing with CATS, The Climate Aware Task Scheduler

Abstract: As well as reducing computational energy consumption, it's important to consider its carbon cost. Performing compute when electricity has lower carbon intensity, due to more renewable generation, will reduce environmental impact. This idea led to the inception of CATS, the Climate-Aware Task Scheduler, at the Software Sustainability Institute (SSI) Collaborations Workshop 2023. CATS is a Python package that schedules tasks based on the estimated carbon intensity of the electricity grid, using real-time data from the UK's National Grid ESO API. Our small team continued developing CATS after the workshop, with support from SSI, and released Version 1.0, designed for use with the 'at' command hence targeting smaller-scale tasks on local machines, this July. Major work is being finalised for Version 2.0, which will integrate with Slurm for HPC applications. Ethical HPC must promote efforts to minimise carbon impact and CATS can contribute to this through intelligent time shifting of jobs.

Bio: Sadie is a Computational Scientist working for the Computational Modelling Services (CMS) group within the National Centre for Atmospheric Science (NCAS) and the Dept. of Meteorology at the University of Reading. Her work largely involves developing, optimising and maintaining open-source tools that support research and collaboration in earth science and aligned domains. Prior to this role which she has held since 2020, she spent two years as a Scientific Software Engineer at the Met Office, following a master's degree in (particle) physics. In 2022 she was awarded a fellowship by the Software Sustainability Institute.

Reducing the impact of energy consumption from computing with CATS, The **C**limate **A**ware **T**ask **S**cheduler

Sadie Bartholomew
University of Reading (Dept. Meteorology), National Centre for Atmospheric Science (NCAS)

On behalf of the full CATS team:
Colin Sauzé, Abhishek Dasgupta, Andrew Walker, Loïc Lannelongue, Thibault Lestang, Tony Greenberg, Lincoln Colling, Adam Ward and Carlos Martinez

For Computing Insight UK 2024 (Sustainability Strand), 2024-12-05

SOFTWARE SUSTAINABILITY INSTITUTE

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Border image credits: 'Climate Stripes' infographic designed by Prof. Ed Hawkins (University of Reading), see showyourstripes.info

Motivating question

Image credits: https://i.imgflip.com/208mpa.jpg, from IT Crowd (Channel 4)

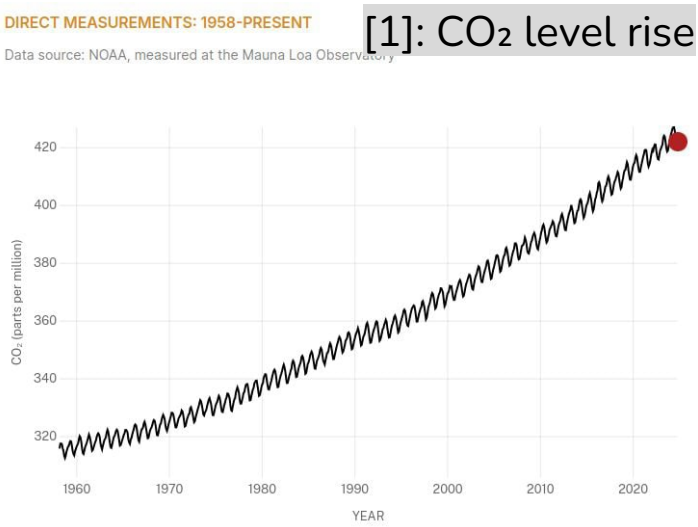
Computing *always requires energy* (electricity etc.) - how can we do it in a sustainable way to not exacerbate the climate crisis?

Motivating issue: the climate crisis

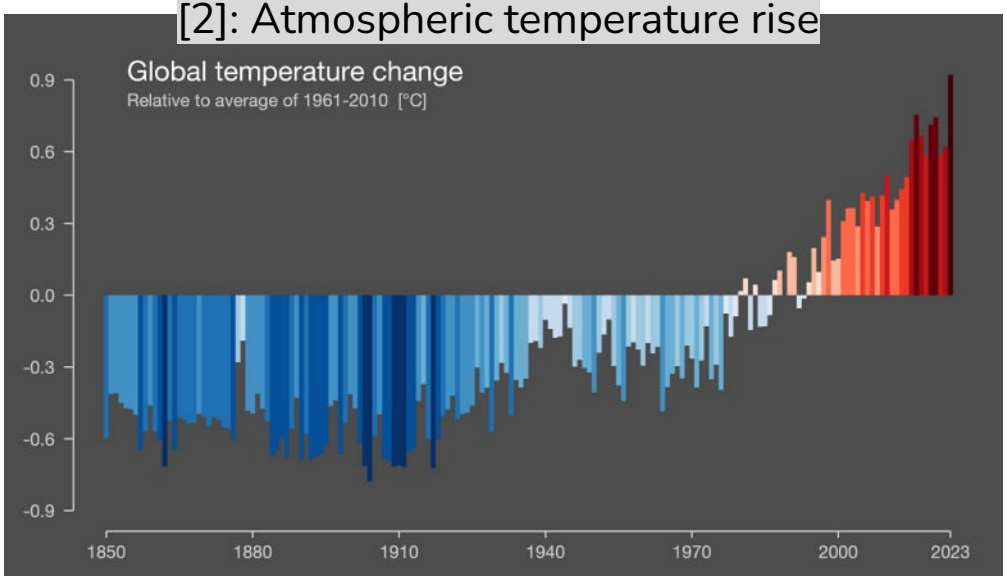
Human activities, notably fossil fuel burning to generate energy, are (largely) responsible for accumulation of greenhouse gases (e.g. CO_2 [1]) in the Earth's atmosphere causing rise in global temperatures [2] and sea levels [3] etc. - **activities including computing**

Sources of plots: showyourstripes.info, https://climate.nasa.gov/vital-signs/

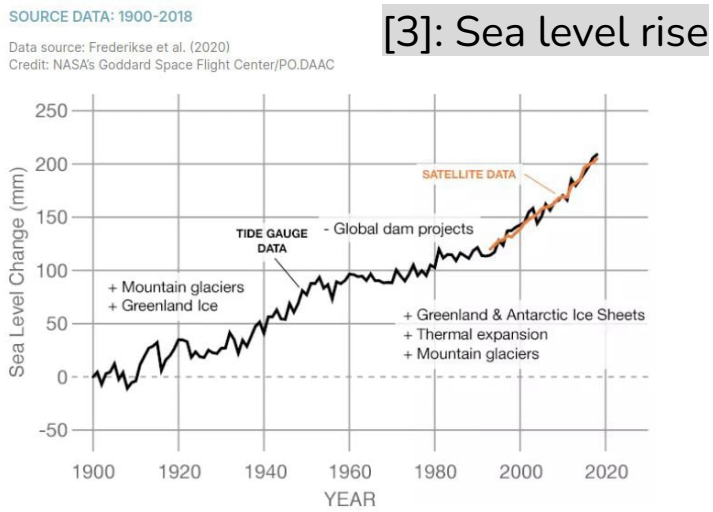
[1]: CO_2 level rise



[2]: Atmospheric temperature rise

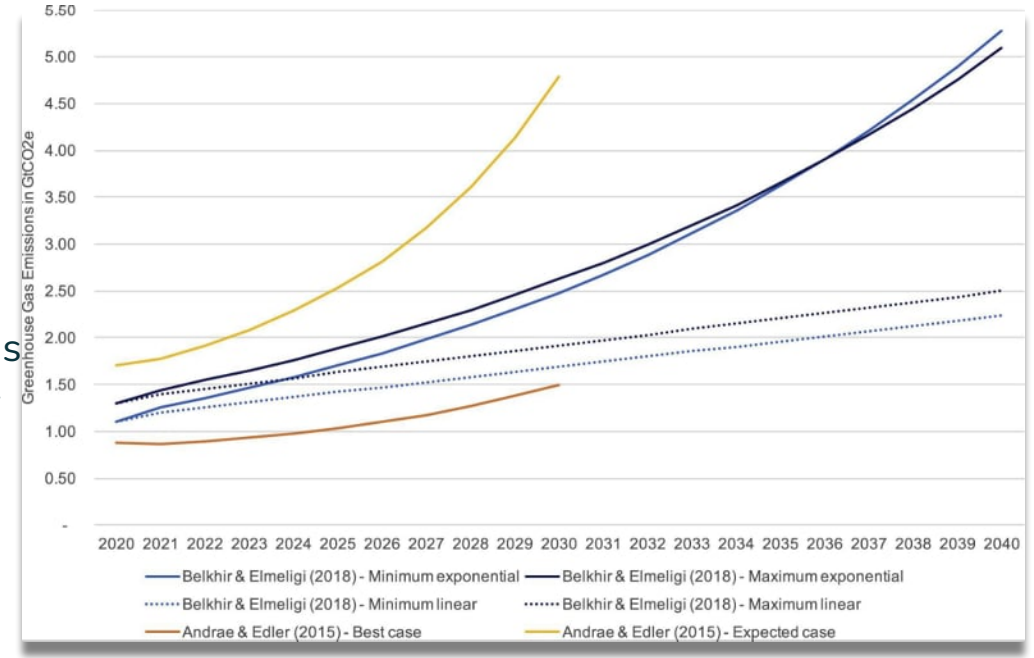


[3]: Sea level rise



The consumption of computing: significant & increasing!

- The ICT sector uses 4-10% of the world's electricity and generates 1.5-5% of its greenhouse gas emissions*
- Computing's global share is modest but growing at a much faster rate than many other energy-consuming sectors. Energy demands of data centers and HPC systems are expected to increase significantly over the next few decades, driven partly by growing use of cloud services, AI, and machine learning models plus more need for data centers which are power hungry



- "ICT's footprint has likely grown faster than global emissions, with a very uncertain best estimate of twice as fast" †, though it is hard to estimate accurately (see plot †)

**Source*: The EU climate strategy for the ICT sector

† Source: 'The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations', Freitag et al., https://doi.org/10.1016/j.patter.2021.100340

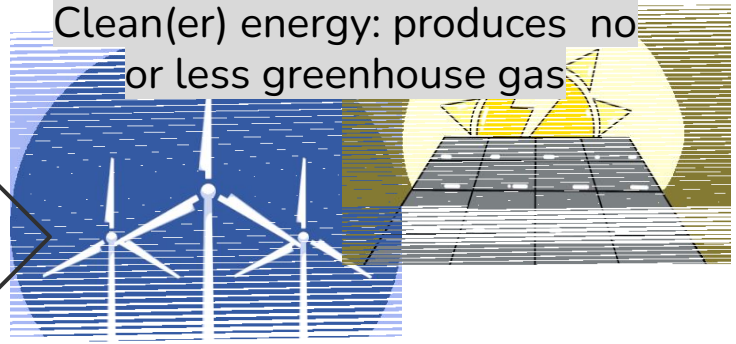
The underlying idea

We can (and should) work to reduce our energy consumption from the computing we do. But we can *also* reduce our climate impact by being *more clever* with the *set energy we do use* so that we end up using **more energy from renewables (clean) than fossil fuels (dirty)**

Dirty energy: produces (more) greenhouse gas

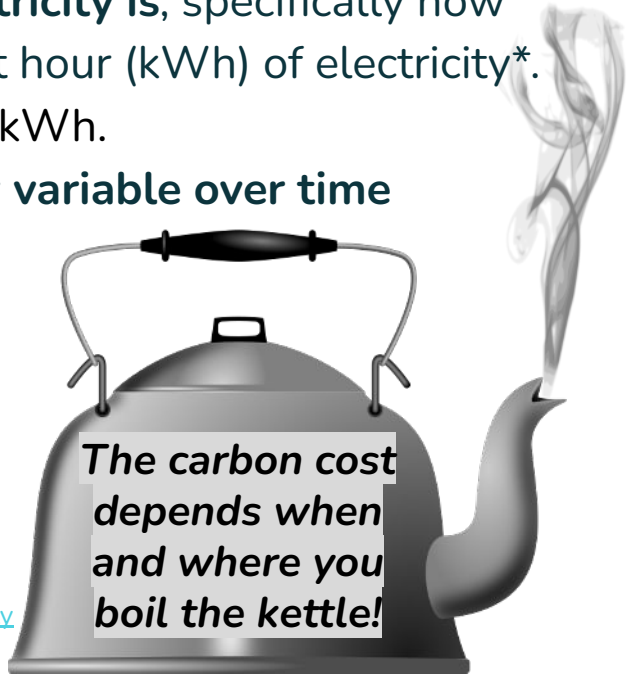
Aim to shift our energy consumption this way

Clean(er) energy: produces no or less greenhouse gas



How do we measure the ‘cleanliness’ of energy we use?

- Because renewable sources aren’t available in a steady manner, and demand on a given electricity grid varies, CO_2 emissions from some task requiring a set amount of electricity *depend on the datetime and location in which the task is done*
- **Carbon intensity is a measure of how clean our electricity is**, specifically how many grams of CO_2 are released to produce a kilowatt hour (kWh) of electricity*. This becomes our metric of interest. Units are gCO_2e/kWh.
- The **carbon intensity of electricity (in the UK) is very variable over time**
  - Windy and/or sunny weather ⇒ lower carbon
  - Generally between 0 and 400 gCO_2e/kWh
  - EU average 251 gCO_2e/kWh in 2022 †

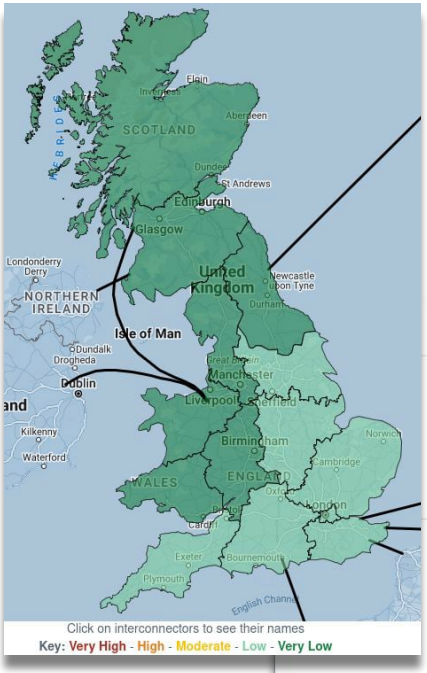


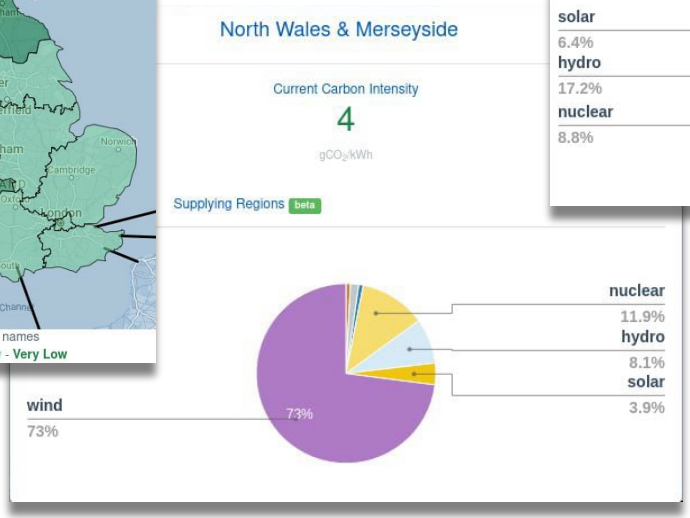
* *Source of definition*: https://www.nationalgrid.com/stories/energy-explained/what-is-carbon-intensity

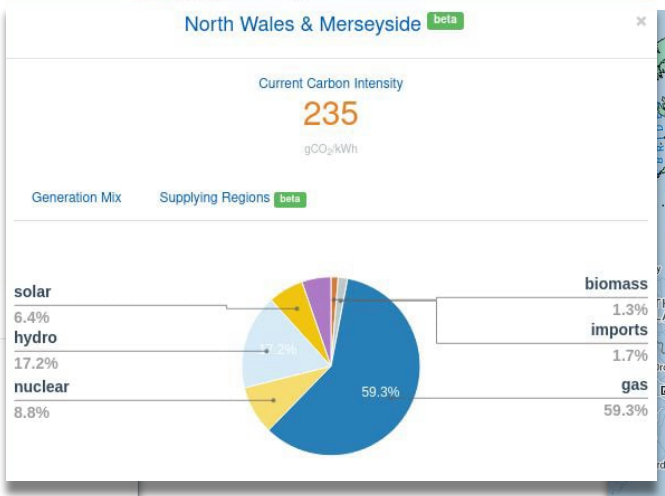
† https://www.eea.europa.eu/en/analysis/maps-and-charts/co2-emission-intensity-15

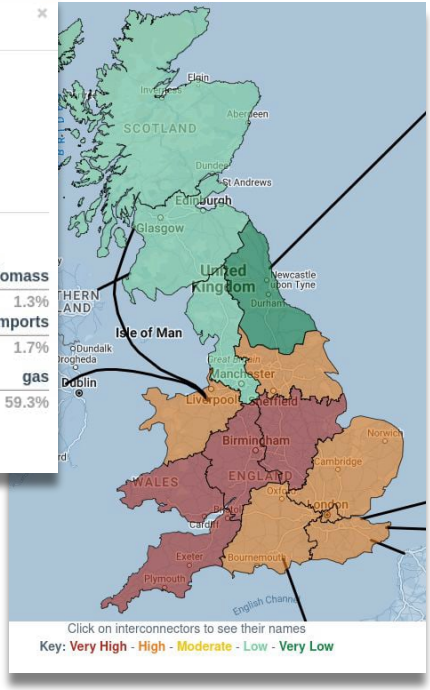
Regional & weather-based influence on carbon intensity

- Left: windy & quite sunny day across UK, right: neither windy nor sunny across UK
- For example showing regional carbon intensity factors for North Wales & Merseyside





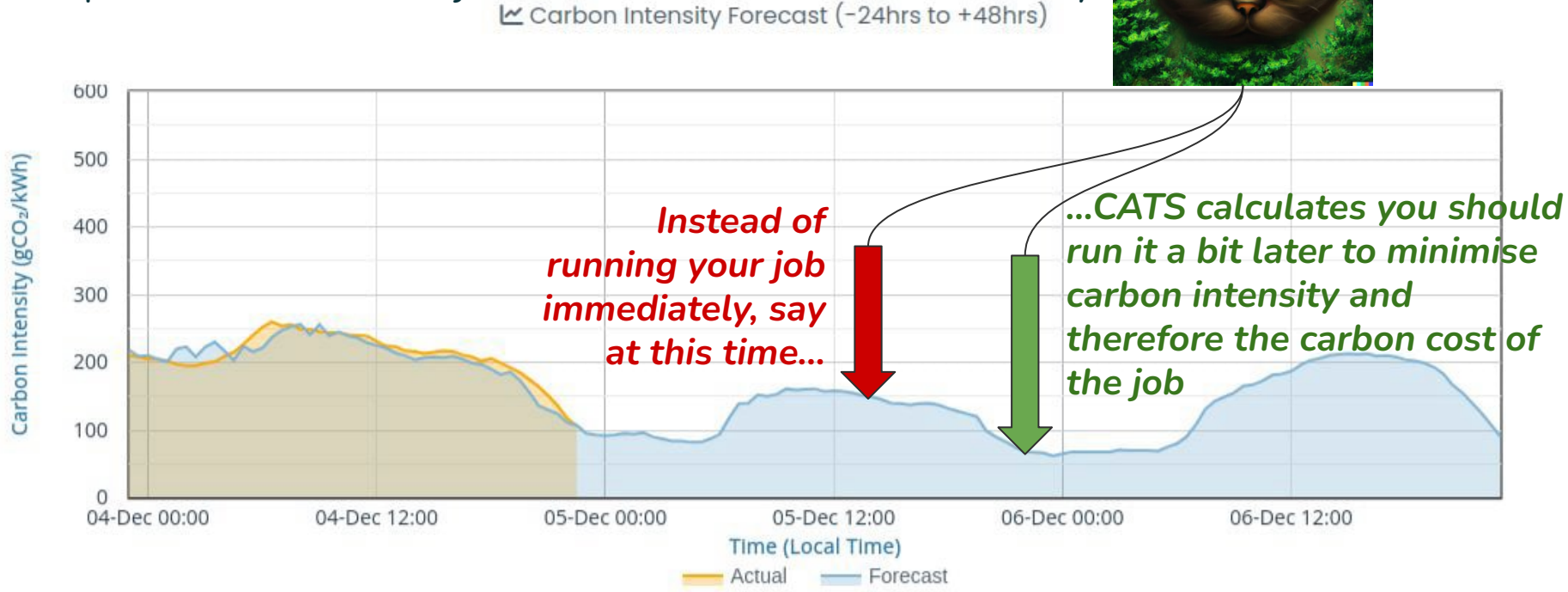




Diagrams and pie charts from: https://carbonintensity.org.uk/

Introducing our tool CATS to manage time-shifting of jobs

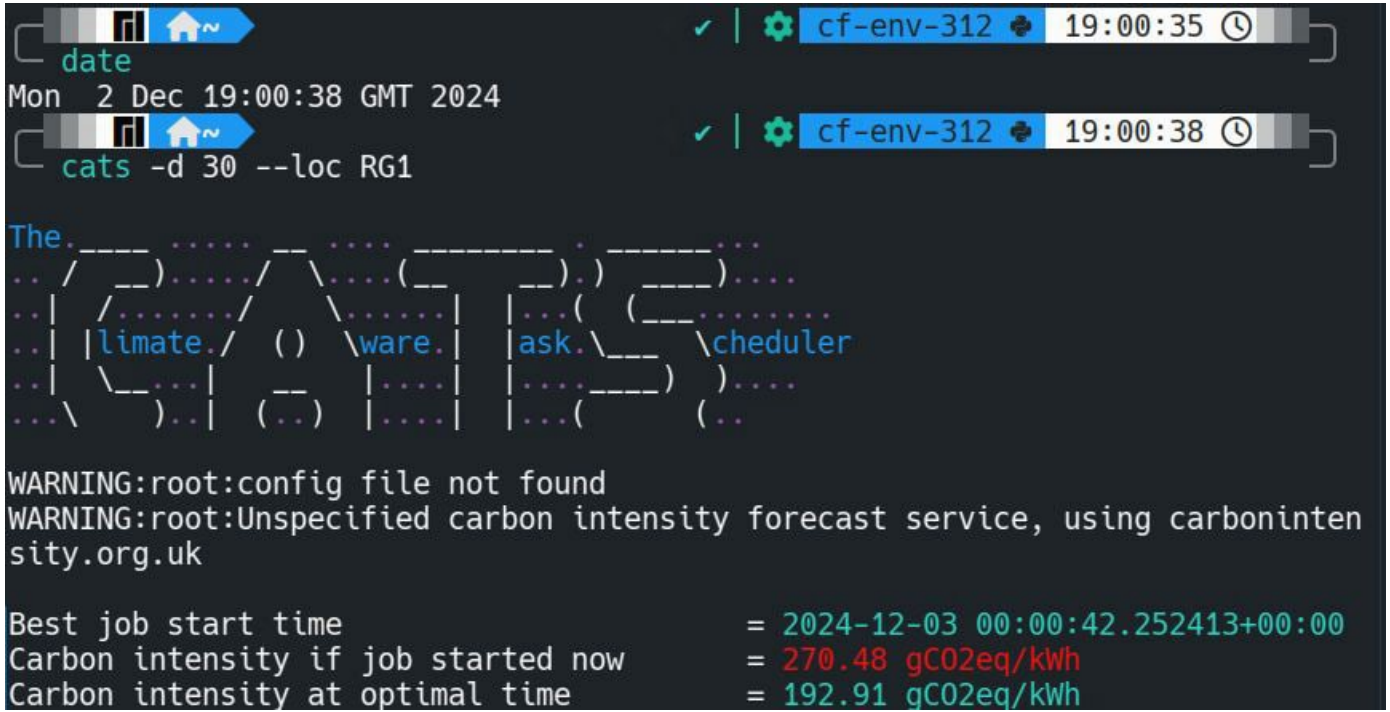
The **C**limate **A**ware **T**ask **S**cheduler (https://github.com/GreenScheduler/cats) calculates the optimal time to run a job to minimise its carbon intensity



Plot from: https://carbonintensity.org.uk/ (more on this resource in later slides!)

Basic usage of CATS: via a command-line interface

- Terminal use, with configuration via YAML file and/or CLI options & arguments
- Minimal use: `cats -d <job duration in mins> --loc <postcode>`
- Example: shows savings of >75 gCO2e/kWh by waiting ~6h to run a job



Further usage: direct scheduling & estimating carbon footprint

- **To directly schedule a job with the CATS calculation**, use the argument `--scheduler`. We currently support the UNIX `at` command, for example to run a Python script `work.py` expected to take an hour or so: `cats -d 60 --loc RG1 --scheduler at --command 'python work.py'`

- **You can go further than carbon *intensity* information and extract the estimated carbon *footprint* reduction** from delaying the compute if you provide memory consumption and a hardware profile for the relevant machine: `cats --duration 480 --location "EH8" --footprint --memory 16 --profile my_gpu_profile --gpu 4 --cpu 1`

Example YAML config file, profiles section (only)

```yaml
profiles:
  my_cpu_only_profile:
    cpu:
      model: "Xeon Gold 6142"
      power: 9.4 # in W, per core
      nunits: 2
  my_gpu_profile:
    gpu:
      model: "NVIDIA A100-SXM-80GB GPUs"
      power: 300
      nunits: 2
    cpu:
      model: "AMD EPYC 7763"
      power: 4.4
      nunits: 1
```

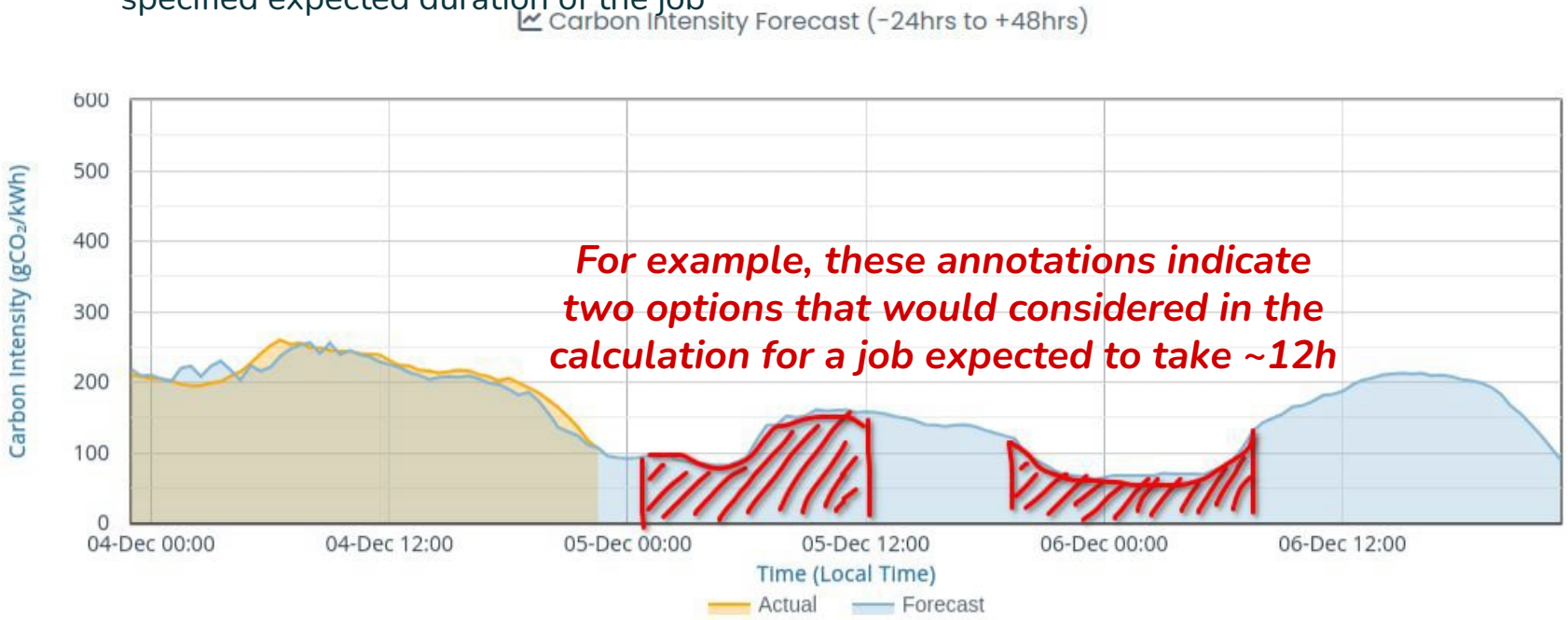
A brief history of CATS

May 2023 (SSI CW23) → *The future*

- Devised & prototyped at the Software Sustainability Institute's Collaborations Workshop 2023 Hack Day (winning first prize!), proof of concept intended for small-scale compute
- Original hackathon team took the project forward together to continue developing CATS
- Version 1.0 released in July this year, marking the first release of a stable tool (full documentation, improved CLI, test coverage & output formats for humans & machines)
- Work in progress with further support from the SSI, including integration with SLURM, testing on real HPCs & submitting a publication to the Journal of Open Source Software

How does CATS work?

- To run software when renewable sources of energy are most plentiful, CATS:
  - uses National Grid ESO's Carbon Intensity API (carbonintensity.org.uk) for carbon intensity forecast
  - takes such data appropriate to the local region (found from a given postcode as proxy for location)
  - calculates to effectively minimise the area under the curve (as illustrated on the plot here) for the specified expected duration of the job



Plot from https://carbonintensity.org.uk/, with SB annotations added (drawn lines in red)

Use cases for CATS: from small- to large-scale compute

- **Version 1.0** (July 2024): first stable/mature release, designed for ‘small-scale’ computing e.g. a few hours on a workstation/desktop or laptop overnight
- Work towards **Version 2.0** is in progress, which aims to target the more pressing source of carbon emissions, HPC and HTC
    - Includes work to test CATS on a ‘mini HPC’ (Raspberry Pi cluster, funded by the SSI and built by CATS team members Sadie and Colin)

Further work and upcoming version 2!

- Work underway for integration with the batch scheduler SLURM (https://slurm.schedmd.com/) which will be in CATS version 2
  - Simplest approach: using `sbatch` to offset start time
  - Our ideal result: HPC systems can implement ‘green’ queues to use CATS to delay jobs that users are happy to in return for reduced carbon footprint (and/or incentives)
  - Integrating carbon accounting as a Slurm plugin (will need rewrite in C)
  - SSI funding provided a few months of developer time, coming to the end of this and approaching completion of work

Example of v2 carbon footprint saving for a fictional HPC

- For an example of a fictional HPC, with hardware as follows:
  - 64 core AMD EPYC 7773X (Milan) CPUs
    - 10 nodes, 2 CPUs per node, 20 CPUs total, 1280 cores
  - Fully loaded CPU = 255 W, Idle CPU = 37.5 W (from https://www.phoronix.com/review/amd-epyc-7773x-linux/9)
  - Idle saving = 217.5 W per CPU
  - Cluster idle vs peak = 4.35 kW
- Time shifting reduces grid intensity from 200 to 50 g/kWh = 150 g/kWh reduction
- The calculation:
  - 12 hour job using all cores
  - 12h * 4.35 kW = 52.2 kWh
  - 52.2 kWh * 0.15 kg = 7.83 kg
- Comparable to driving an average car (150 g/km) 50 km (7.5 kg)!

Limitations of CATS and notes regarding value

- Only works for the UK (at present) due to lack of APIs like the National Grid ESO's Carbon Intensity one used, for other countries/regions (open Issue https://github.com/GreenScheduler/cats/issues/22)
- Relies on user specifying the job length correctly - and this can be hard to estimate and might require pre-run(s) to estimate well (enough)
- Won't be able to do much on systems at/near 100% load
- Can't handle jobs expected to take more than 2 days due to forecast cutoff of the National Grid ESO API
- *Note*: the UK electricity grid is planned to be net-zero by 2035, but that's quite optimistic and besides, if we can do something now, then why wait?
- *Note*: not the only thing you can/should do to reduce the climate impact of your computing! You can look to also reduce emissions from scope 3 (manufacturing), cooling, storage & networks (e.g. see blog post 'Tracking the environmental impact of research computing' SSI blog post covering useful background: https://www.software.ac.uk/blog/tracking-environmental-impact-research-computing)

Summary of CATS, The Climate Aware Task Scheduler

- Computing uses (a lot of!) energy - HPC, HTC, data centers and AI in particular
- One approach to reduce our impact on the climate crisis from greenhouse gas emissions resulting from our energy consumption is to shift to **using more of the 'clean' renewable sources over 'dirty' sources like fossil fuels**
- We can do this by **using the local electricity when it is lower in carbon intensity**
- By **intelligently time shifting compute jobs** to run them at the time that minimises carbon footprint across their expected duration, using real-time carbon intensity data from the National Grid ESO API, CATS can contribute to more sustainable computing
- CATS was **initially developed for small-scale compute jobs, but work is underway to support HPC/HTC** (better targets for reducing carbon impact!) via SLURM integration
- For now, **try out CATS Version 1**! See https://github.com/GreenScheduler/cats

SOFTWARE SUSTAINABILITY INSTITUTE

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Thanks for listening.

For more info. about CATS and/or other aspects from this talk, please ask me anything now or you can explore such resources as:

- the CATS codebase, OSS on Github: https://github.com/GreenScheduler/cats
- the CATS package documentation: https://greenscheduler.github.io/cats/
- a recent episode of the 'Code for Thought' podcast in which myself and Colin talk about CATS: https://www.buzzsprout.com/1326658/episodes/15766448-en-bonus-green-computing-at-the-rse-conference-2024-in-newcastle?t=0
- 'Tracking the environmental impact of research computing' SSI blog post covering useful background: https://www.software.ac.uk/blog/tracking-environmental-impact-research-computing

Border image credits: 'Climate Stripes' infographic designed by Prof. Ed Hawkins (University of Reading), see showyourstripes.info

CIUK 2024 Presentations

Jessica Huntley (Science and Technology Facilities Council)

Tracking the carbon cost of numerical optimization algorithms

Abstract: One of the biggest global challenges is tackling climate change and building a sustainable future. Whilst software on its own doesn't emit CO2, understanding the cost of running it is important so that the way it is developed and used can be optimised for minimal carbon emissions. There are several carbon tracking tools available for estimating hardware power consumption, allowing users and developers to quantify the carbon cost of running a program and identify areas for improvement. FitBenchmarking (https://fitbenchmarking.github.io/) is an open source Python package that allows users to compare the performance of different optimization algorithms and their implementations across various fitting problems, through metrics such as runtime and accuracy. In this talk, I will outline how a new carbon emissions metric has been added to FitBenchmarking, some of the challenges faced along the way, and the insight that reporting emissions offers developers and users.

Bio: Jessica works as a Mathematical Research Software Engineer in the Scientific Computing Department's Computational Mathematics Theme. Alongside her day-to-day role, she also leads the department's Energy Efficiency Team, a working group which aims to help staff reduce their carbon footprint.

Tracking the carbon cost of numerical optimization algorithms

Jess Huntley
Mathematical Research Software Engineer, STFC
jessica.huntley@stfc.ac.uk

About me

$$\frac{D\underline{v}}{Dt} = -\nabla p + \mu\nabla^2\underline{v} + \frac{1}{3}\mu\nabla(\nabla\cdot\underline{v}) + \rho$$

$$\frac{D\rho}{Dt} + \rho\nabla\cdot\underline{v} = 0 \qquad -\overline{v_i v_s} = \nu_t\left(\frac{\partial u_i}{\partial x_s} + \frac{\partial u_s}{\partial x_i}\right.$$

$$p = \rho RT \qquad \frac{\partial(\rho v_i)}{\partial t} + \nabla\cdot(\rho v_i\underline{v}) = -p\nabla\cdot\underline{v} + \ldots$$

Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

Session Outline

- Why should we track carbon cost of software?
- How can we track carbon?
- Adding an energy usage table to FitBenchmarking
- Insights and next steps

Why track software carbon emissions?

We are facing a climate emergency

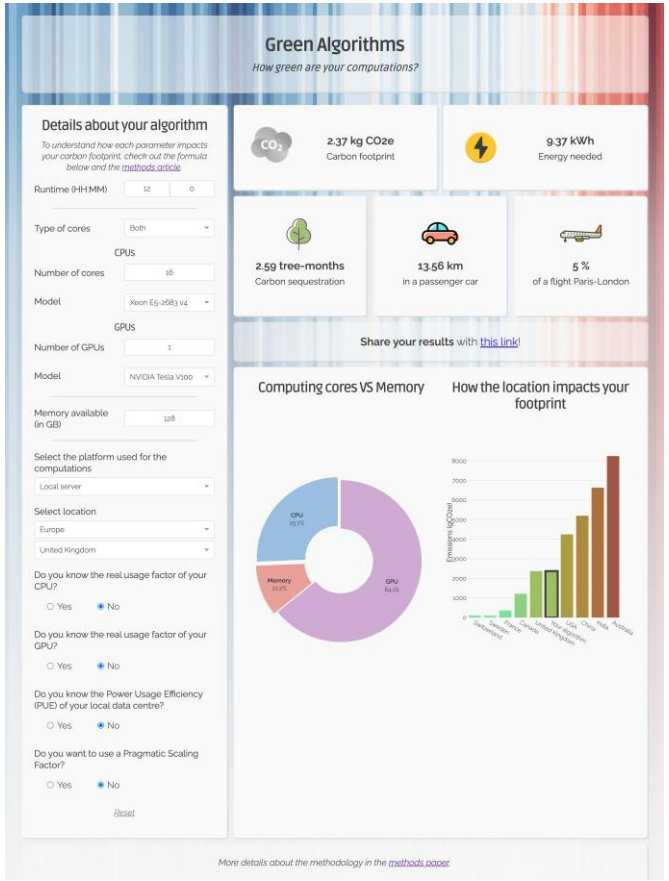
IT sector was estimated to account for 2-6% of global carbon emissions in 2020, could increase to 20% by 2030 [1].

Currently limited information on the carbon cost of computational research

Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

How to track carbon emissions?



[2]

CODE CARBON

[3]

Green Algorithms 4 HPC

Version v0.2.2 | Maintained? yes | Open Source ? Yes!

[4]

Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

FitBenchmarking

- Open-source Python package
- Compares how different nonlinear least squares solvers perform on various fitting problems.
- Supported by ALC funding
  - Centre of expertise in scientific software with primary objective of maximising the scientific impact of STFC's facilities.

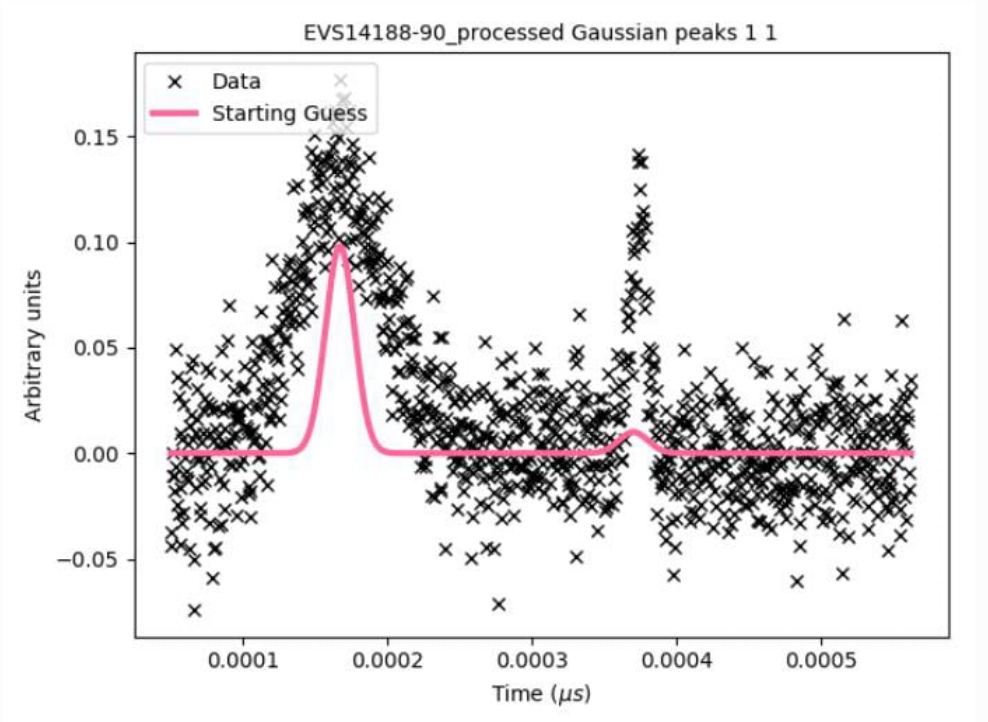
FitBenchmarking

ada lovelace centre

Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

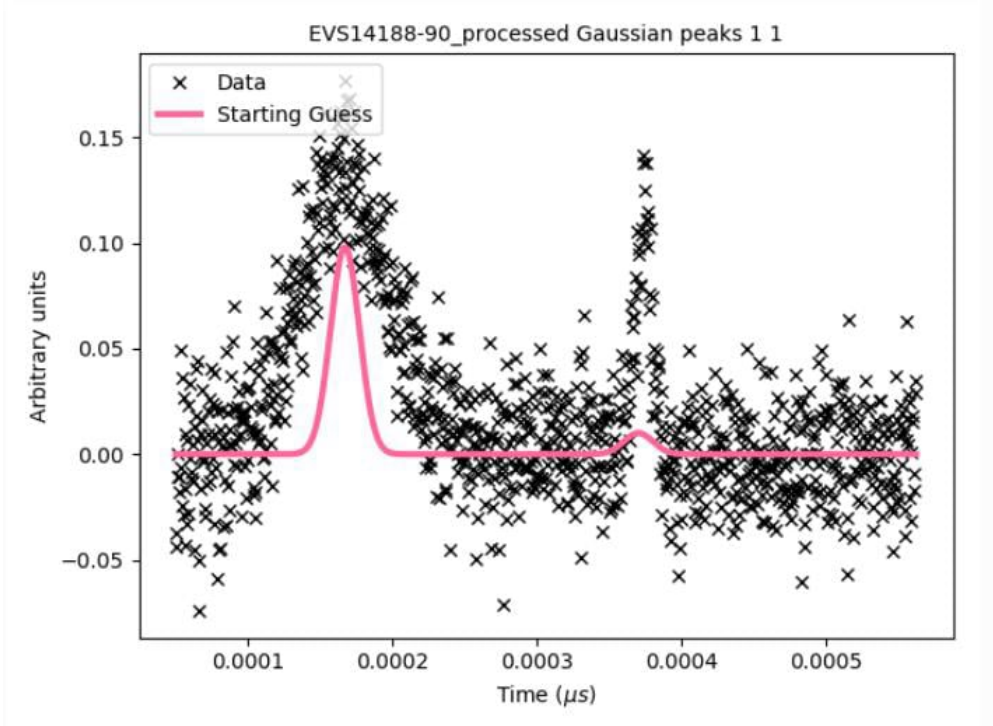
Example workflow

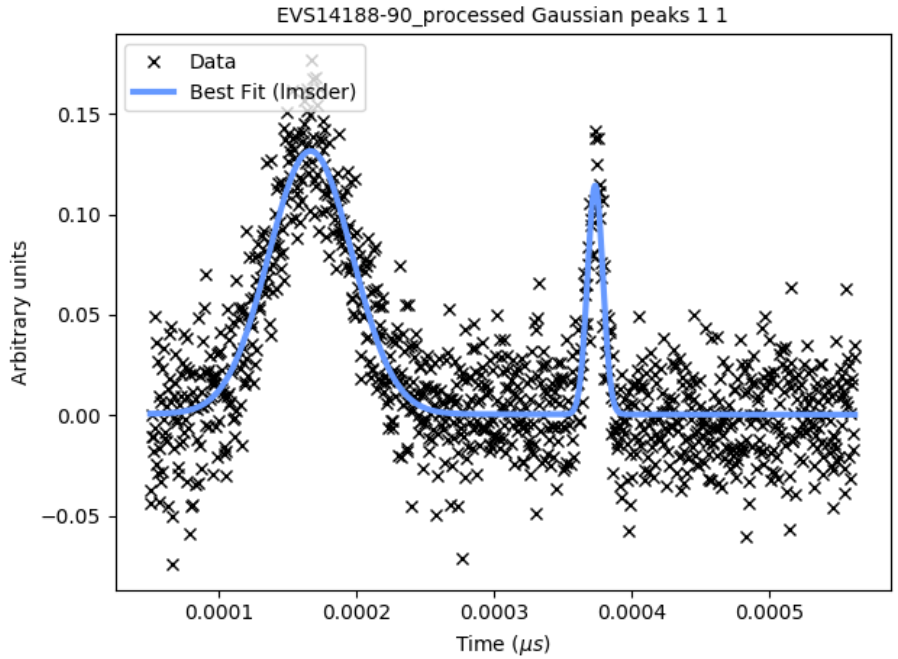


Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

Example workflow

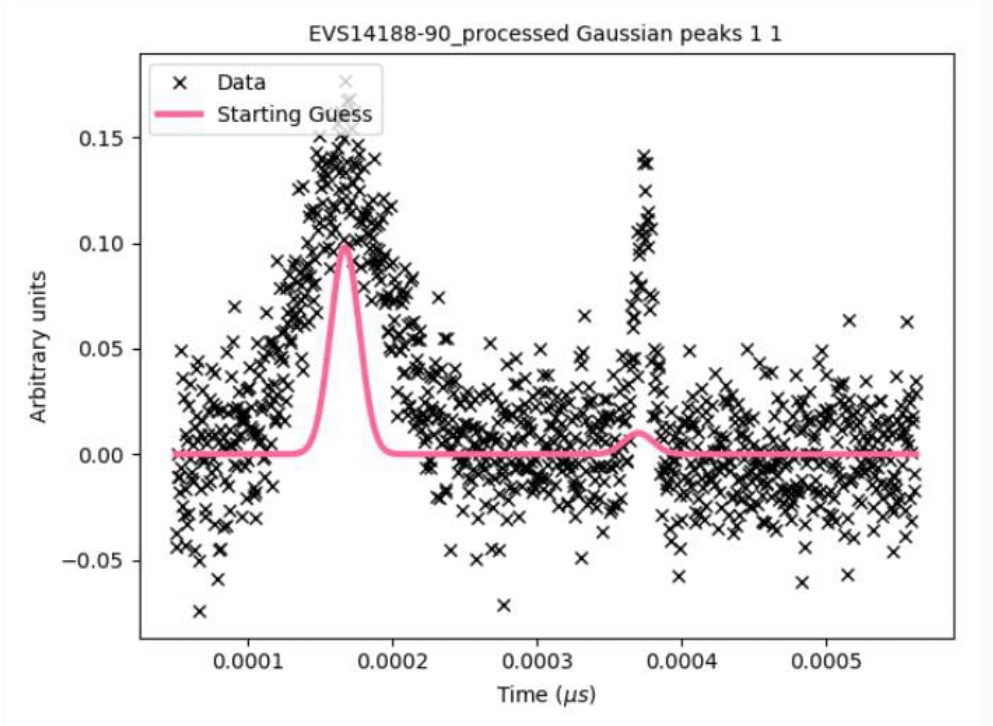


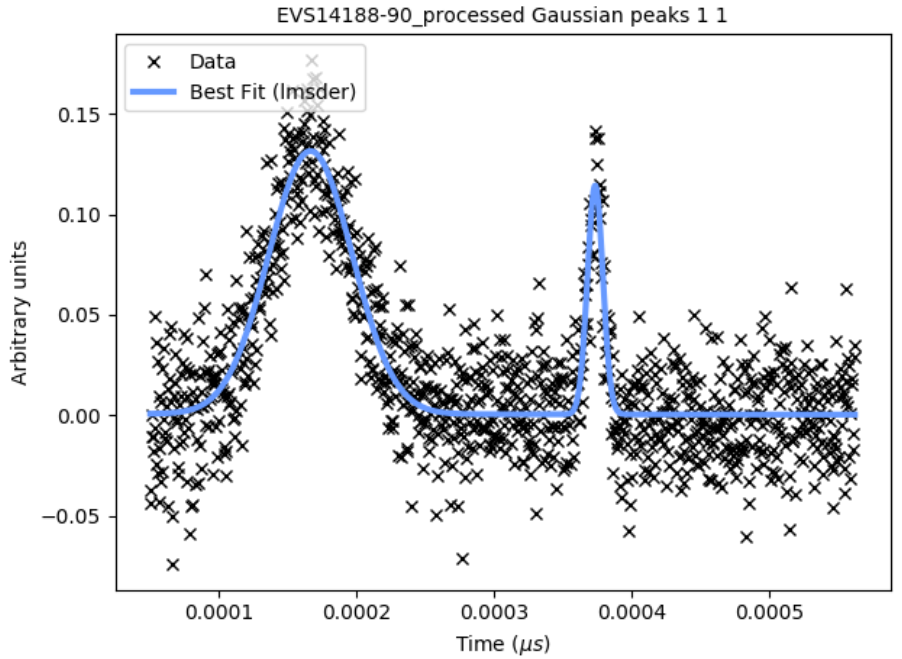


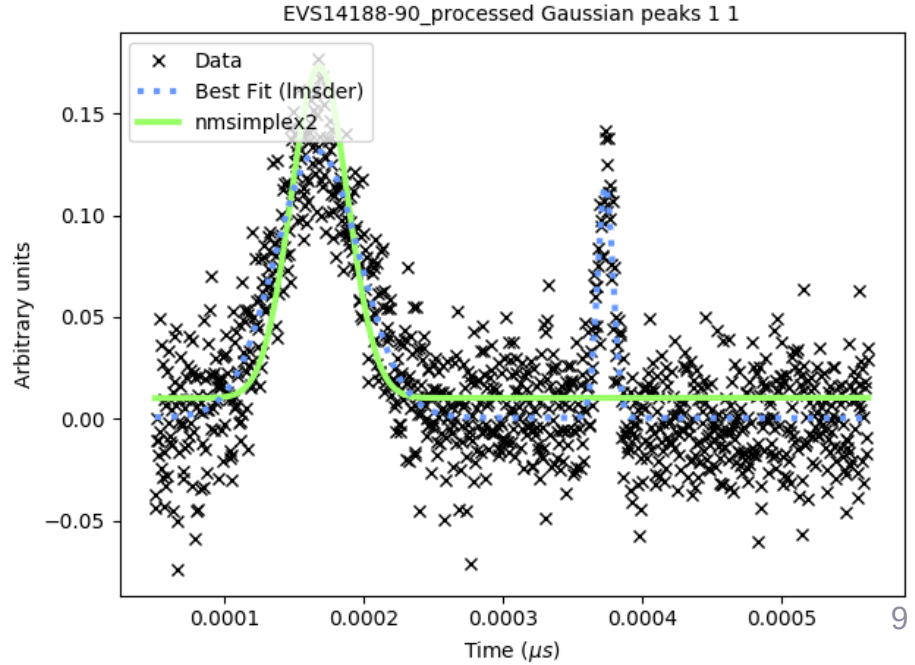
Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

Example workflow



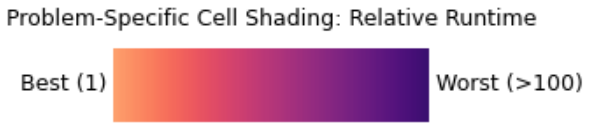




Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

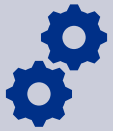
Existing results table metrics



			scipy	
	BFGS: j:best_available	CG: j:best_available	COBYLA	N
ENSO, Start 1	0.02465 (2.874)	0.1274 (14.85)	0.05832 (6.8)[2]	0.1
ENSO, Start 2	0.02697 (1.63)	0.0677 (4.093)	0.1238 (7.486)[2]	0.1
Gauss3, Start 1	0.01475 (19.59)	0.6222 (826.1)[2]	0.0007532 (1)[2]	0.1
Gauss3, Start 2	0.03022 (38.8)	0.1072 (137.6)[2]	0.0007788 (1)[2]	0.(
Hahn1, Start 1	0.134 (25.37)[2]	0.8061 (152.6)[1]	0.0153 (2.895)	0.(
Hahn1, Start 2	0.1301 (44.37)[2]	1.072 (365.5)[1]	0.008123 (2.771)	0.2
Kirby2, Start 1	0.02923 (16.5)[2]	0.004618 (2.607)[2]	0.003126 (1.765)	0.(

Adding the energy usage table

Used CodeCarbon to embed energy usage tracking into FitBenchmarking – initially reported carbon emissions, but this is not a consistent metric



Only tracking energy usage for call to fitting function – explicit object used but was slow initially…

Science and Technology Facilities Council
Scientific Computing

ada lovelace centre

Implementation Pitfalls

- For Windows, obtaining an accurate measurement relies on using a deprecated tool
- On Linux, root access required to change permissions of Intel RAPL files
- Fallback mode estimates energy usage based on runtime, so not very useful for FitBenchmarking users!

Accuracy

	WeightedNLLSCostFunc				
	lmfit		scipy		scipy-go
	differential_evolution	dual_annealing	COBYLA	Powell	differential_evolution
ECKERLE4	0.04663 (1)	2.653 (56.9)	2.653 (56.9)	2.653 (56.91)	2.653 (56.9)
MGH09	0.002846 (1)	0.002846 (1)	0.003973 (1.396)^2	0.002846 (1)	0.002846 (1)
MGH10	7372 (1.385e+06)^2	1.104 (207.4)	568.5 (1.068e+05)	1.332e+04 (2.502e+06)	0.005325 (1)
RAT42	0.3439 (1)	0.3439 (1)	52.07 (151.4)^2	0.3439 (1)	0.344 (1)
RAT43	17.5 (1)	17.5 (1)	3541 (202.4)^2	17.5 (1)	17.5 (1)
THURBER	51.05 (6.725)^2	7.591 (1)	2648 (348.8)^5	3.656e+05 (4.816e+04)	7.599 (1.001)

Runtime

	WeightedNLLSCostFunc				
	lmfit		scipy		scipy-go
	differential_evolution	dual_annealing	COBYLA	Powell	differential_evolution
ECKERLE4	0.0745 (15.64)	1.045 (219.5)	0.004762 (1)	0.01659 (3.483)	0.1189 (24.96)
MGH09	1.432 (25.78)	1.427 (25.7)	0.05555 (1)[2]	0.09937 (1.789)	0.8144 (14.66)
MGH10	1.887 (360.5)[2]	1.324 (253)	0.005234 (1)	0.06169 (11.79)	3.747 (715.9)
RAT42	0.3601 (11.35)	1.022 (32.21)	0.05546 (1.747)[2]	0.03174 (1)	0.2804 (8.835)
RAT43	0.8459 (14.77)	1.422 (24.82)	0.05728 (1)[2]	0.0641 (1.119)	0.5537 (9.667)
THURBER	4.018 (163.9)[2]	3.261 (133.1)	0.05751 (2.347)[5]	0.02451 (1)	4.818 (196.6)

Energy usage

	WeightedNLLSCostFunc				
	lmfit		scipy		scipy-go
	differential_evolution	dual_annealing	COBYLA	Powell	differential_evolution
ECKERLE4	2.165e-06 (8.526)	2.963e-05 (116.7)	2.539e-07 (1)	5.78e-07 (2.277)	3.432e-06 (13.52)
MGH09	3.889e-05 (35.99)	2.726e-05 (25.23)	1.081e-06 $(1)^2$	1.872e-06 (1.732)	1.568e-05 (14.51)
MGH10	3.486e-05 $(206.4)^2$	2.565e-05 (151.8)	1.689e-07 (1)	1.2e-06 (7.103)	6.891e-05 (408)
RAT42	6.789e-06 (10.37)	1.902e-05 (29.04)	1.092e-06 $(1.667)^2$	6.549e-07 (1)	5.296e-06 (8.088)
RAT43	1.567e-05 (13.9)	2.647e-05 (23.48)	1.127e-06 $(1)^2$	1.245e-06 (1.105)	1.042e-05 (9.246)
THURBER	7.454e-05 $(142.1)^2$	6.097e-05 (116.2)	1.135e-06 $(2.163)^5$	5.246e-07 (1)	8.972e-05 (171)

Relationship between energy and runtime

- Van Kempen et al [5] claim that energy and runtime are directly proportional – independent of choice of programming language.
- One of the contributors to power draw is the number of active cores
  - In their experiments, parallelizing programs was almost always an energy efficient choice.

Is running in parallel more efficient?

Runtime

		LoglikeNLLSCostFunc	
		bumps	
		dream	parallel_dream
ENSO, Start 1	9 params, 168 points	27.62 (1.09)	25.33 (1)
ENSO, Start 2	9 params, 168 points	27.83 (1.167)	23.86 (1)
Gauss3, Start 1	8 params, 250 points	23.11 (1.06)	21.81 (1)
Gauss3, Start 2	8 params, 250 points	23.62 (1.08)	21.87 (1)

Energy

		LoglikeNLLSCostFunc	
		bumps	
		dream	parallel_dream
ENSO, Start 1	9 params, 168 points	0.0005021 (1.136)	0.0004419 (1)
ENSO, Start 2	9 params, 168 points	0.0005228 (1.173)	0.0004455 (1)
Gauss3, Start 1	8 params, 250 points	0.0004431 (1.055)	0.00042 (1)
Gauss3, Start 2	8 params, 250 points	0.0004575 (1.091)	0.0004192 (1)

How does an emissions table help users?

Carbon budgeting

Raise awareness

Motivation for developers to produce more carbon efficient solvers

Next steps

- Add examples section to documentation to include case studies
- Improve results table outputs so energy usage can be directly compared to other metrics
- Investigate trade-off between energy usage and number of cores

Scientific Computing

Thank you

jessica.huntley@stfc.ac.uk

scd.stfc.ac.uk

@SciComp_STFC

References

[1] Copenhagen Centre on Energy Efficiency. Greenhouse gas emissions in the ICT sector: Trends and methodologies [Internet]. 2020. Available from: https://c2e2.unepdtu.org/wp-content/uploads/sites/3/2020/03/greenhouse-gas-emissions-in-the-ict-sector.pdf

[2] Lannelongue, L., Grealey, J., Inouye, M., Green Algorithms: Quantifying the Carbon Footprint of Computation. Adv. Sci. 2021, 2100707. https://doi.org/10.1002/advs.202100707

[3] Benoit Courty, ‘mlco2/codecarbon: v2.6.0’. Zenodo, Aug. 09, 2024. doi: 10.5281/zenodo.13286710.

[4] https://www.green-algorithms.org/GA4HPC/

[5] Van Kempen, N. et al. It’s Not Easy Being Green: On the Energy Efficiency of Programming Languages, https://arxiv.org/abs/2410.05460

CIUK 2024 Presentations

Vassil Alexandrov and Adriano Agnello (Hartree Centre - STFC)

Harnessing AI to Address Grand Challenges in Scientific Innovation

Abstract: In this talk we will outline the Hartree Centre's engagement with industry, public sector and research communities in driving scientific innovation based on extreme scale computing and scalable AI solutions. In particular, we will discuss our partnerships and collaborations with UKAEA and US National Labs such as ORNL, LLNL, LBL and ANL to illustrate how we are working to address Grand Challenges with high societal impact. There are many interdisciplinary applications this work can address, and we will showcase our work in AI to bypass simulation bottlenecks in magnetic-confinement fusion and materials discovery, and AI for effective elicitation of knowledge.

Bio: Prof. Vassil Alexandrov (VA), Chief Science Officer, Hartree Centre – STFC. His focus is currently on the enhancement and implementation of the comprehensive Hartree Centre Research Strategy that enables to employ advanced HPC, Data Analytics, AI and quantum-classical methods and algorithms to address key industrial, societal and scientific challenges. He is also bringing key expertise in computational science, extreme scale computing, scalable fault-tolerant and resilient algorithms for advanced computer architectures and Monte Carlo methods and algorithms, applied to diverse application areas and extreme scale (peta and exascale) architectures. He leads STFC's collaboration with UKAEA, based at Hartree Centre, and co-leads with Dr. Rob Akers (UKAEA) the UKAEA -Hartree Centre STFC Fusion Computing Lab. He led Hartree Centre's collaboration with US ECP as part of UK and US National Labs collaboration in exascale computing and now is leading Hartree Centre's collaboration with US national labs (LLNL and ORNL) in extreme scale computing and scalable AI.

Bio: Dr Adriano Agnello (AA) is a Principal AI Researcher at the Hartree Centre, supervising research projects in AI for Applied Sciences and Engineering, including a multi-year workstream in collaboration with the UKAEA and the Intelligent Observatory ISPF project with the South African Astronomical Observatory. Together with the AI Group Leader, he liaises with colleagues at US National Labs on common AI initiatives, including workflows for self-driving labs, simulation orchestration, and magnetic confinement fusion. Besides his work in AI (including contributions to the STFC AI/DR Strategy and ETP4HPC SRA6), Dr Agnello has led searches of rare cosmic milestones and worked on spectroscopic instrument design and recommissioning projects.

CIUK 2024 Presentations

Noel O'Connor (Redhat)

Partnering to take LLM training to the Next Level

Abstract: Background: Today, many industries manage large amounts of data and, as their requirements grow, the technology that supports them seeks to keep pace. Large language models (LLMs) such as ChatGPT step up to the challenge by providing a user-friendly way for organisations to interact with their data at the speed of business to automate processes, provide valuable data insights and create customized content based on their data. LLMs are a form of artificial intelligence (AI) trained on massive sets of text-based data to perform various tasks ranging from detecting patterns to crafting original content.

Currently, LLMs are flourishing in the open with limitations and without efficiency and the gains of traditional open source software. Excitement about the potential of this technology is high, but it's met with equal measures of concern. Across industries, a consensus exists that data is a precious asset organisations should handle with care. As a result, worries about privacy, inaccurate data, biases and leaks of sensitive information are at the centre of the conversation.

Objective: Use the power of collaborative partnering to and end to end solution which will deliver open-source concepts to models and the tools for open evolution. In addition, enable community users and developers to create and merge contributions to LLMs, while increasing efficiency and performance.

Methods: CDW present a Red Hat offering, Instruct LAB where users can learn & experiment via limited desktop-scale training method on small datasets. which includes models and capabilities for model evolution and serving, packaged as immutable instances and tuned to specific architectures, with support, indemnification and lifecycle necessary for enterprise deployment. This is deployed onto Dell hardware, a world leader in technology that is trusted and compatible.

Companies are not comfortable using large language models due to the unknown license of the outputted text. RHEL AI includes indemnification and the model itself is licensed under the Apache-2.0 Open Source license.

Results: A fully strategic and holistic method for enabling learners and researchers to develop training of LLMs from a laptop. With CDW delivering InstructLab* and the Granite models, we make the ability to add knowledge and skills to the language model accessible to everyone, allowing companies to add their specific knowledge and skills to align the model to their specific use case. The InstructLab provides a way for people to get started in a cost effective way and scale out when needed under the guidance of CDW, reliability of Dell and the power of Nvidia .

Conclusion: Utilising the skills and expertise of complimentary technologies and services to deliver an innovative and ever evolving LLM with real world impact from day one.

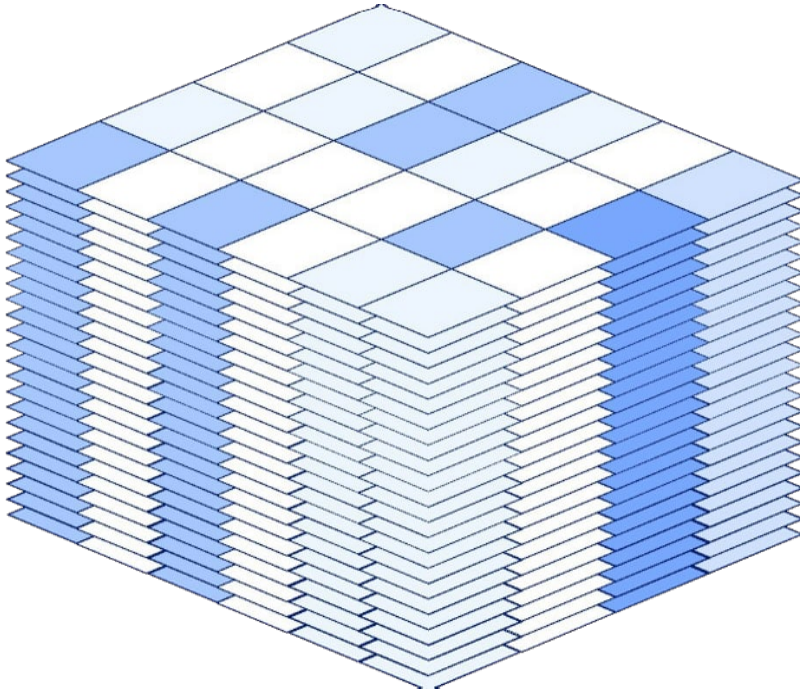
Bio: Noel O'Connor is a Senior Principal Architect in Red Hat's EMEA Solutions Practice specializing in cloud-native application and integration architectures. Since 2008, he has worked with many of Red Hat's global enterprise customers in both EMEA and Asia. He's co-authored a number of books, including "DevOps with OpenShift" and "DevOps Culture and Practice with OpenShift."

InstructLab

Partnering to take LLM training to the Next level

Noel O'Connor
Senior Principal Architect
Red Hat Global Consulting

Large Language Models



Generate Image from prompt

Massive Data beast

Compute

Datasets for Large Language Models: A Comprehensive Survey

Yang Liu[1,3], Jiahuan Cao[1], Chongyu Liu[1], Kai Ding[2,3], Lianwen Jin[1,3]

[1]South China University of Technology
[2]INTSIG Information Co., Ltd
[3]INTSIG-SCUT Joint Lab on Document Analysis and Recognition

https://arxiv.org/pdf/2402.18041

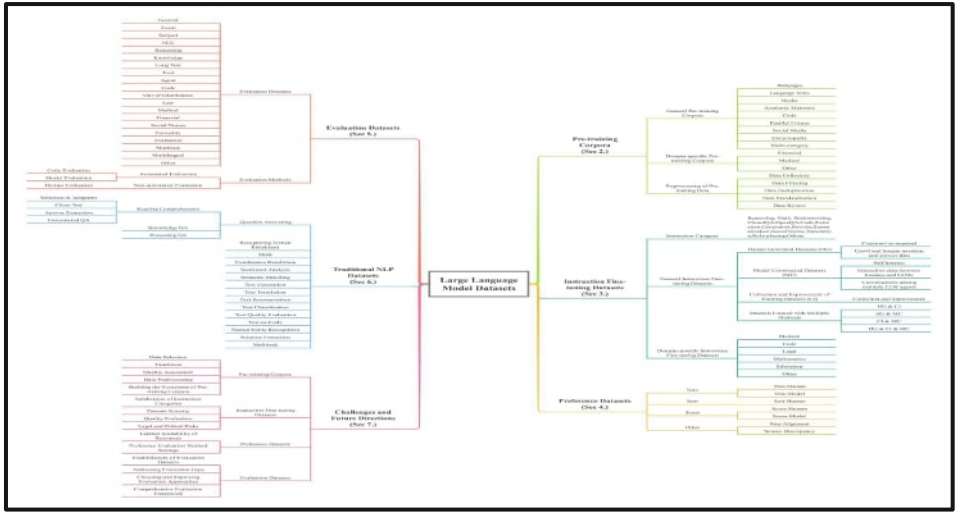


Table 2 Summary of **General Pre-training Corpora** Information **Part II**. Language: "EN" indicates English, "ZH" indicates Chinese, "AR" indicates Arabic, "PL" indicates Programming Language, "Multi" indicates Multilingual, and the number in parentheses indicates the number of languages included. "CM" indicates Construction Methods, where "HG" indicates Human Generated Corpora, "MC" indicates Model Constructed Corpora, and "CI" indicates Collection and Improvement of Existing Corpora

Corpus	Language	CM	Category	Source
ANC	EN	HG	Language Texts	American English texts
Anna's Archive	Multi	HG	Books	Sci-Hub, Library Genesis, Z-Library, etc.
ArabicText 2022	AR	HG & CI	Multi	ArabicWeb, OSCAR, CC100, etc.
arXiv	EN	HG	Academic Materials	arXiv preprint
Baidu baike	ZH	HG	Encyclopedia	Encyclopedic content data
BIGQUERY	PL	CI	Code	BigQuery
BNC	EN	HG	Language Texts	British English texts
BookCorpusOpen	EN	CI	Books	Toronto Book Corpus
CC-Stories	EN	CI	Webpages	Common Crawl
CC100	Multi (100)	CI	Webpages	Common Crawl
CLUECorpus2020	ZH	CI	Webpages	Common Crawl
Common Crawl	Multi	HG	Webpages	Web crawler data
CulturaX	Multi (167)	CI	Webpages	mC4, OSCAR
C4	EN	CI	Webpages	Common Crawl
Dolma	EN	HG & CI	Multi	Project Gutenberg, C4, Reddit, etc.
Github	PL	HG	Code	Various code projects
mC4	Multi (108)	CI	Webpages	Common Crawl
MNBVC	ZH	HG & CI	Multi	Chinese books, webpages, theses, etc.
MTP	EN & ZH	HG & CI	Parallel Corpus	Chinese-English parallel text pairs on the web
MultiUN	Multi (7)	HG	Parallel Corpus	United Nations documents
News-crawl	Multi (59)	HG	Language Texts	Newspapers
OpenWebText	EN	HG	Social Media	Reddit
OSCAR 22.01	Multi (151)	CI	Webpages	Common Crawl
ParaCrawl	Multi (42)	HG	Parallel Corpus	Web crawler data

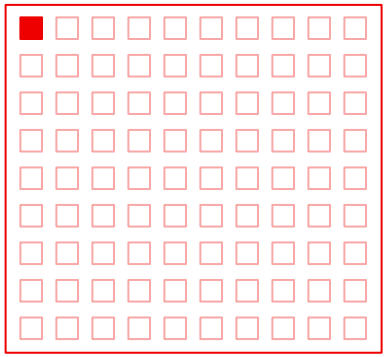
Generate Image from prompt

large beast eating lots of data

Compute

Organisations need models aligned to their **private data**

LLMs are trained with a large range of public data, not ***private*** *enterprise-relevant data*



Less than 1% of all enterprise data
is represented in foundation models

Enterprise organizations need to

1. Start from a **trusted** base model
2. Align the model to **your data & use case**
3. Deploy and scale across the hybrid cloud

Red Hat

Adding knowledge to existing LLMs

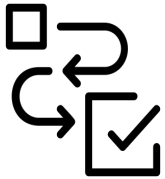
RAG

Retrieval Augmented Generation

Enhance Gen AI model-generated text by retrieving relevant information from external sources, improving accuracy and depth of model's responses.

Fine tuning

Fine Tuning

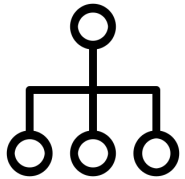


Adjust a pre-trained model on specific tasks or data, improving its performance and accuracy for specialized applications without full retraining.

NEW

InstructLab

Large-scale Alignment for chatBots



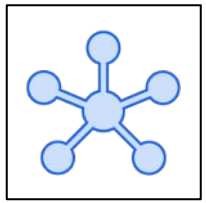
Leverage a taxonomy-guided synthetic data generation process and a multi-phase tuning framework to improve model performance.

Red Hat

InstructLab and the “LAB” method

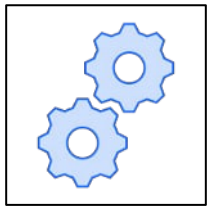
Red Hat

Large-scale **A**lignment for chat**B**ots Methodology



Taxonomy-driven Data Curation

Large-scale Synthetic Data Generation



Phased, Large-scale Alignment Tuning

Available on Arxiv

LAB (Large-scale Alignment for ChatBots) Method

Types Of Information

Skills are performative. When you create a skill for the model, you're teaching it how to do something: "write me a song," "rearrange words in a sentence" or "summarize an email."

Knowledge focuses more on answering questions that involve facts, data, or references. Knowledge is supported by documents, such as a textbook, technical manual, encyclopedia, journal, or magazine.

Knowledge Example

Selection View Go Run Terminal Help

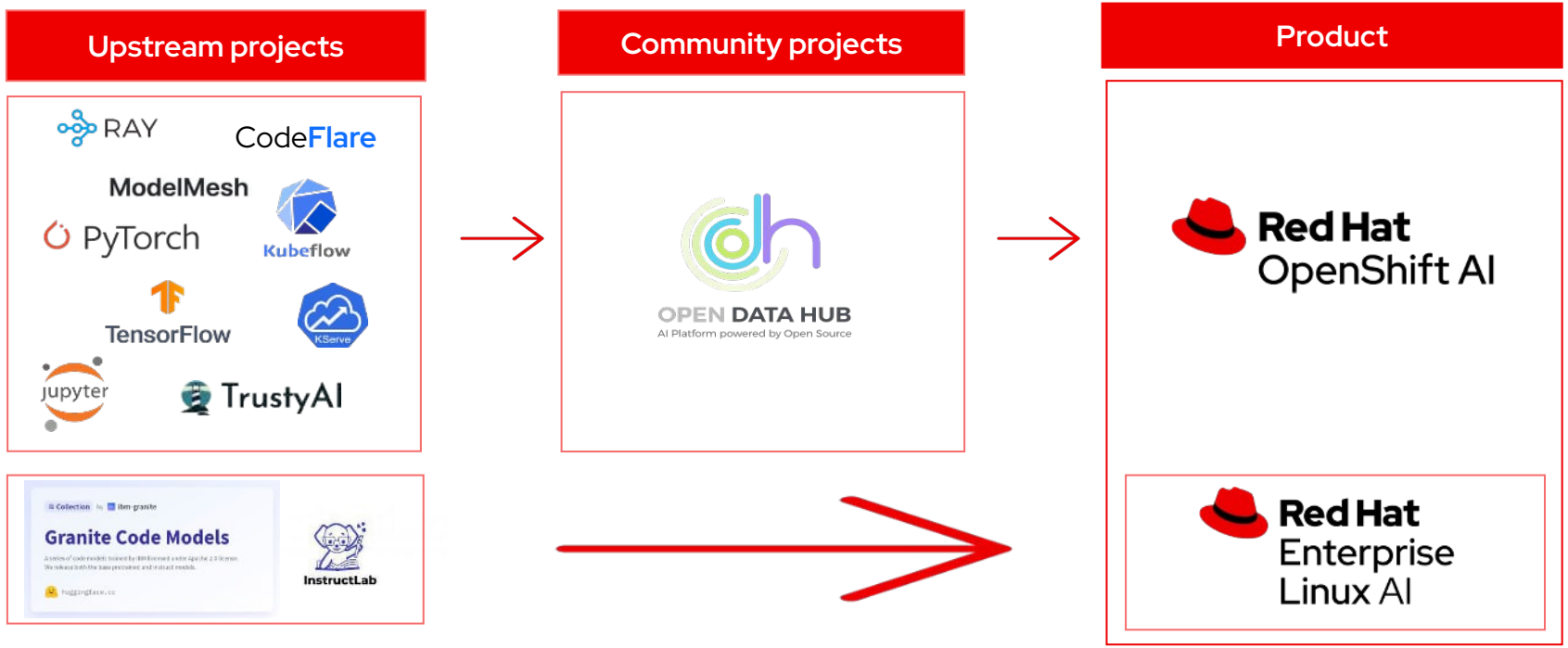
qna.yaml

qna.yaml > created_by

```yaml
created_by: noelo
domain: Batteries
seed_examples:
- context: '# Lithium iron phosphate battery The *lithium iron phosphate battery*
    (*LiFePO4 battery*) or LFP battery (lithium ferrophosphate) is a type of lithium-ion
    battery using lithium iron phosphate (LiFePO4) as the cathode material, and a
    graphitic carbon electrode with a metallic backing as the anode.'
  questions_and_answers:
  - answer: The cathode material in a lithium iron phosphate battery is lithium iron
      phosphate (LiFePO4).
    question: What is the cathode material in a lithium iron phosphate battery?
  - answer: Another name for a lithium ferrophosphate battery is LFP battery.
    question: What is another name for a lithium ferrophosphate battery?
  - answer: The anode material in a lithium iron phosphate battery is a graphitic
      carbon electrode with a metallic backing.
    question: What is the anode material in a lithium iron phosphate battery?
- context: Because of their low cost, high safety, low toxicity, long cycle life and
    other factors, LFP batteries are finding a number of roles in vehicle use, utility-scale
    stationary applications, and backup power. LFP batteries are cobalt-free. As of
    September 2022, LFP type battery market share for EVs reached 31%, and of that,
    68% were from EV makers Tesla and BYD alone. Chinese manufacturers currently hold
    a near monopoly of LFP battery type production. With patents having started to
    expire in 2022 and the increased demand for cheaper EV batteries, LFP type production
    is expected to rise further and surpass lithium nickel manganese cobalt oxides
    (NMC) type batteries in 2028. The specific energy of LFP batteries is lower than
    that of other common lithium-ion battery types such as nickel manganese cobalt
    (NMC) and nickel cobalt aluminum (NCA).
```

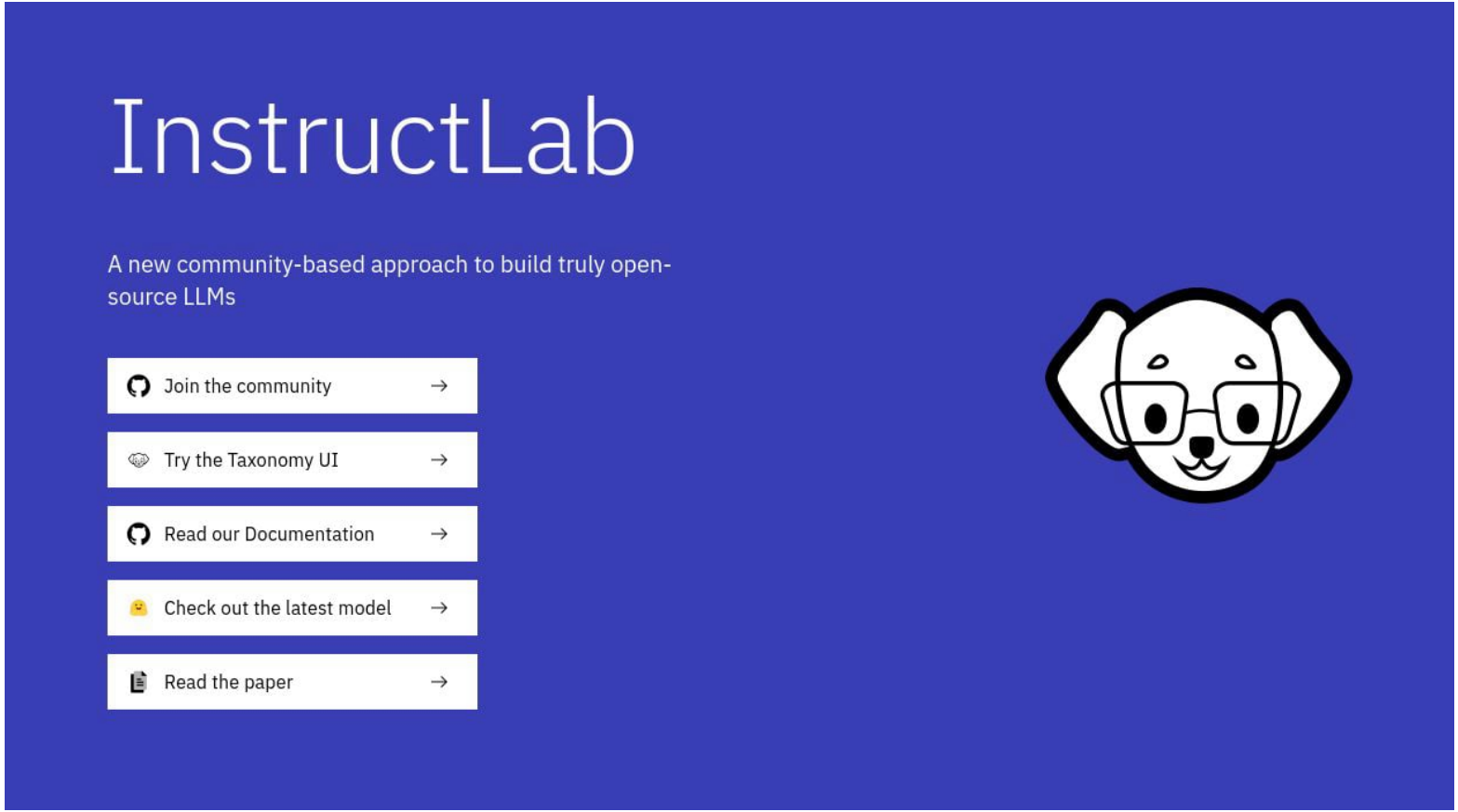
https://github.com/noelo/taxonomy/blob/main/knowledge/energy/electricity/batteries/qna.yaml

Red Hat's AI/ML engineering is 100% open source



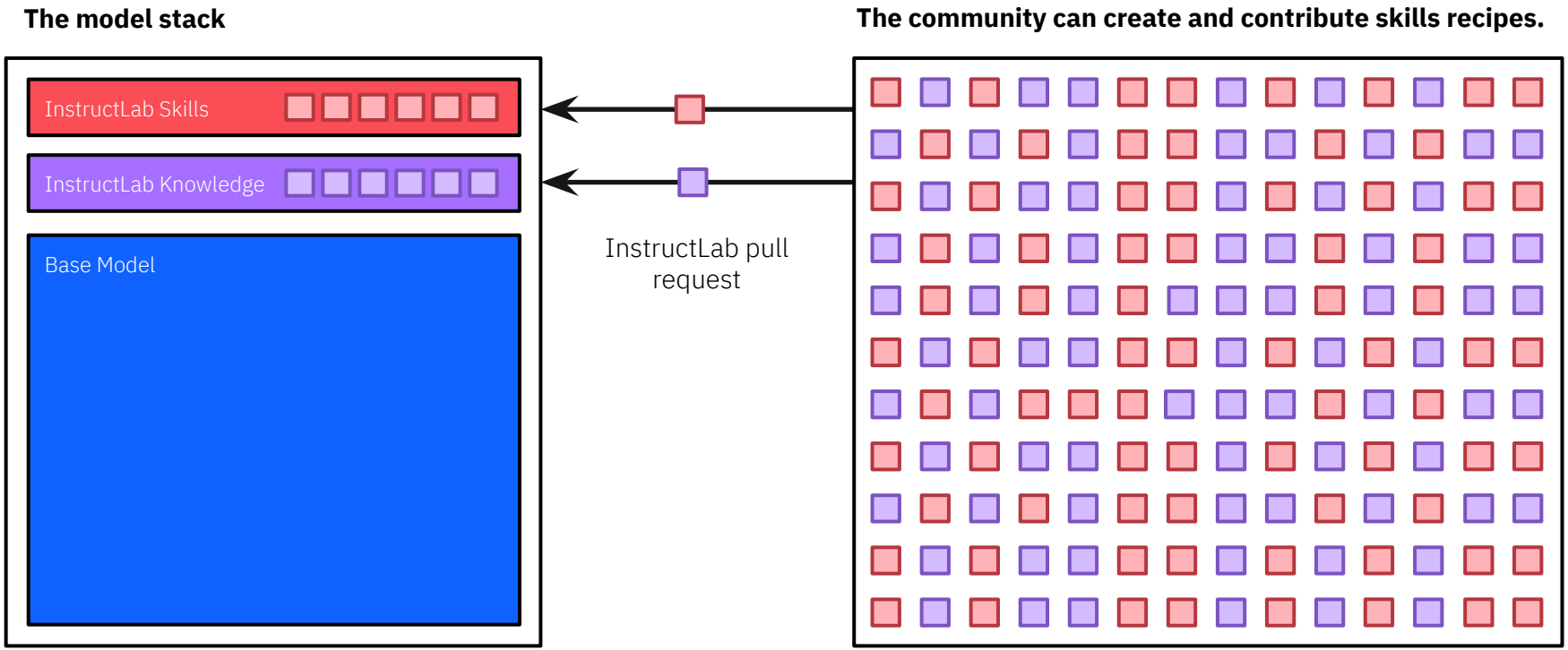
Red Hat

InstructLab Project: OSS community project for GenAI model development

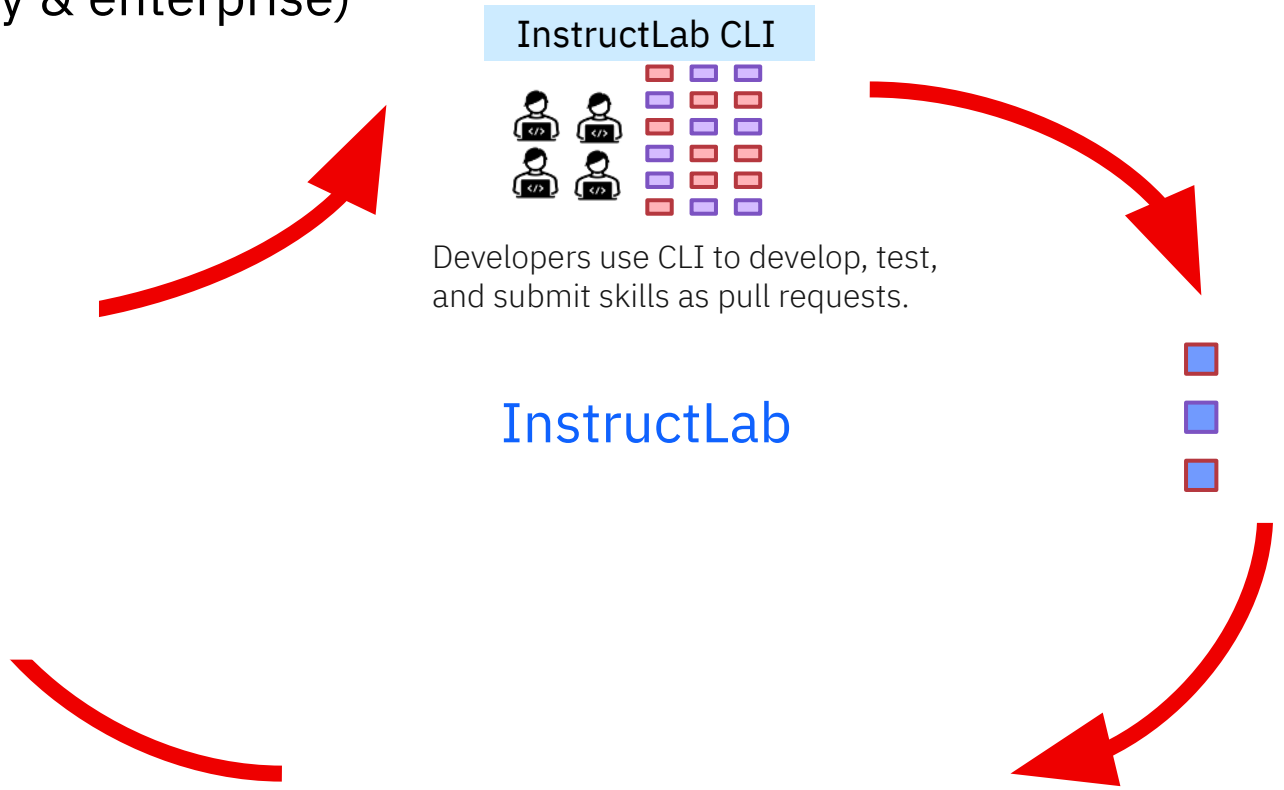


https://instructlab.ai/

InstructLab: Leverages the LAB method to enable community-driven development and evolution of models

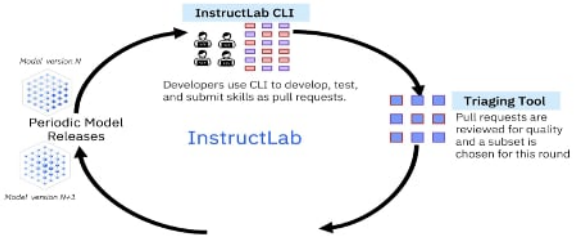


InstructLab: a flywheel for rapid open source innovation (community & enterprise)

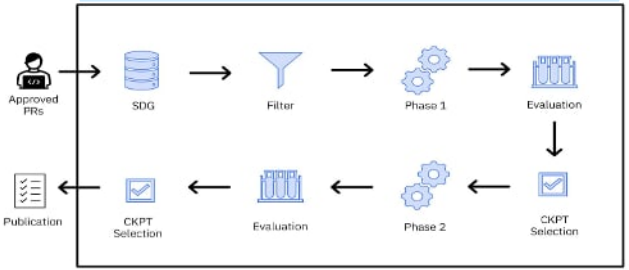


Community Instance

operated by Red Hat with support from IBM Research team



InstructLab Backend for open community



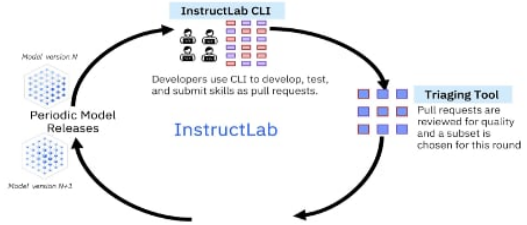
Local

Knowledge/Skills

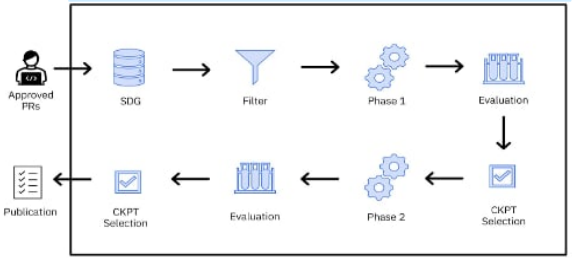
GIT

Customer Instance

operated by Customer or RedHat/Partner_aaS with commercial support from Red Hat



InstructLab Backend for customers



Red Hat OpenShift AI

Red Hat Enterprise Linux AI

Red Hat

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make
Red Hat a trusted adviser to the Fortune 500.

linkedin.com/company/red-hat

facebook.com/redhatinc

twitter.com/redhat

Red Hat

CIUK 2024 Presentations

Simon McIntosh-Smith (The Bristol Centre for Supercomputing, University of Bristol), Pryia Sharma (DSIT) and Christopher Edsall (University of Cambridge)

The National AI Research Resource (AIRR) – The journey to launch

Abstract: Join the Universities of Bristol and Cambridge, the Department of Science, Innovation and Technology (DSIT) and UK Research and Innovation (UKRI) to hear about the launch of the national AI Research Resource (AIRR). Find out about the journey so far, from construction and deployment of Dawn at Cambridge and Isambard-AI at Bristol to early access use cases and their experiences and lessons learnt, as well as the future outlook.

Bio:

Professor Simon McIntosh-Smith is the founder and Director of the Bristol Centre for Supercomputing, which runs the UK's Isambard-AI service. He began his career in industry as a microprocessor architect, first at Inmos and STMicro in the 1990s, before co-designing the world's first fully programmable GPU at Pixelfusion in 1999. In 2002 he co-founded ClearSpeed Technology where, as Director of Architecture and Applications, he co-developed the first modern many-core HPC accelerators. He previously founded the HPC Research Group in Bristol, where his research interests include advanced computer architectures and performance portability.

Priya Sharma is the Compute Delivery Lead at the Department for Science, Innovation and Technology (DSIT), overseeing the delivery of high-performance computing projects, particularly the AI Research Resource project in collaboration with the University of Bristol and the University of Cambridge. Since joining the Civil Service in 2016, Priya has successfully led and delivered complex digital infrastructure projects, including the UK Telecoms Lab (UKTL) and the UK Emergency Alerts Project.

Chris Edsall is the Head of Research Software Engineering at Research Computing Services, a co-director of the Institute of Computing for Climate Science, and a bye-fellow of Queens' College Cambridge. He leads several teams of research software engineers and HPC consultants with the goal of delivering better software enabling better research.

Chris studied physics at the University of Canterbury and then worked in several research institutions in New Zealand (NIWA) and the UK (National Oceanography Centre and the University of Bristol) administering their HPC systems and upskilling researchers in software engineering to make best use of the supercomputing facilities.

COMPUTING INSIGHT UK 2024

Science and Technology Facilities Council

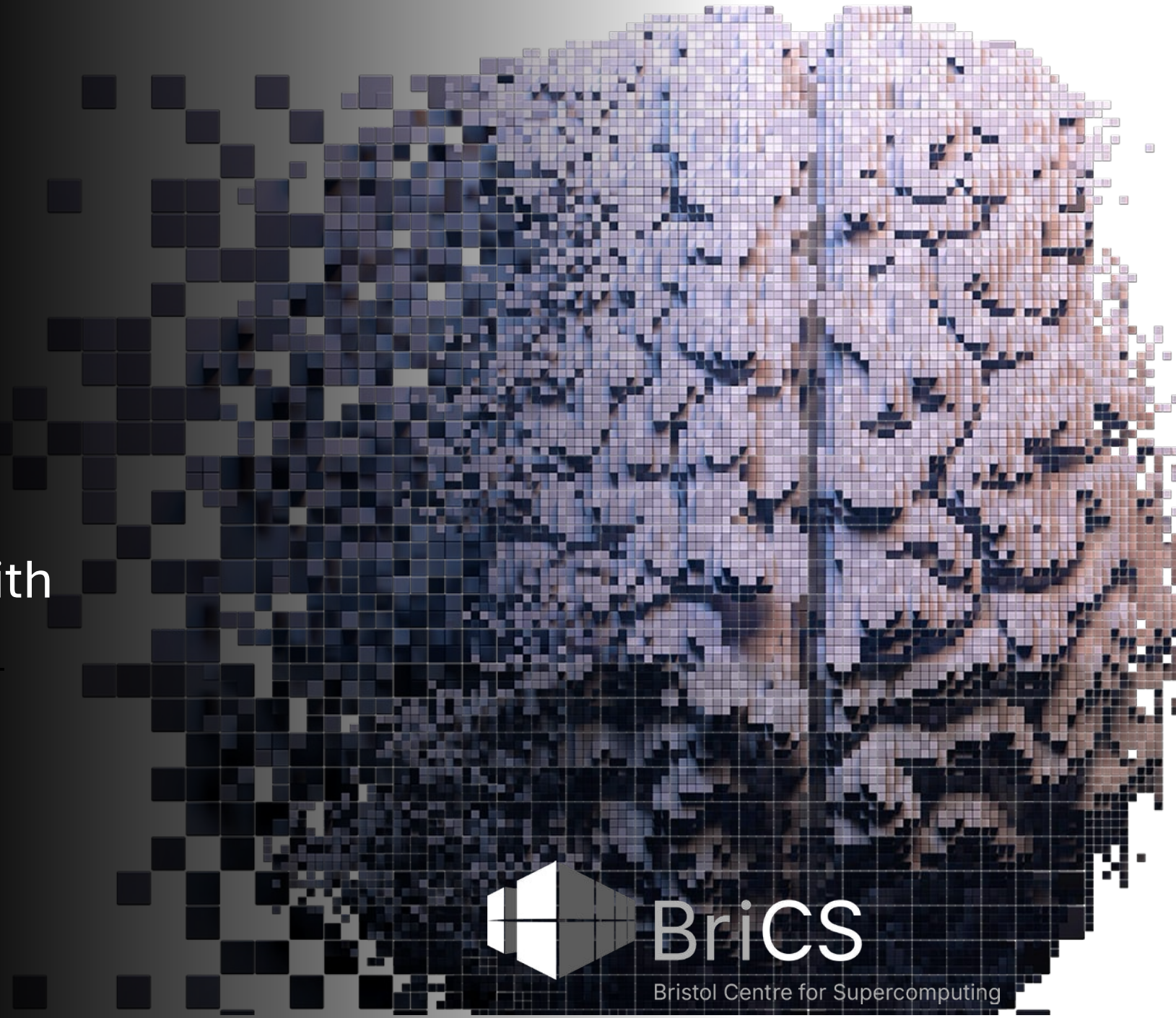
The National AI Research Resource (AIRR)
The journey to launch

Isambard-AI update

Prof. Simon McIntosh-Smith

Bristol Centre for Supercomputing (BriCS)

University of BRISTOL



What is Isambard-AI?

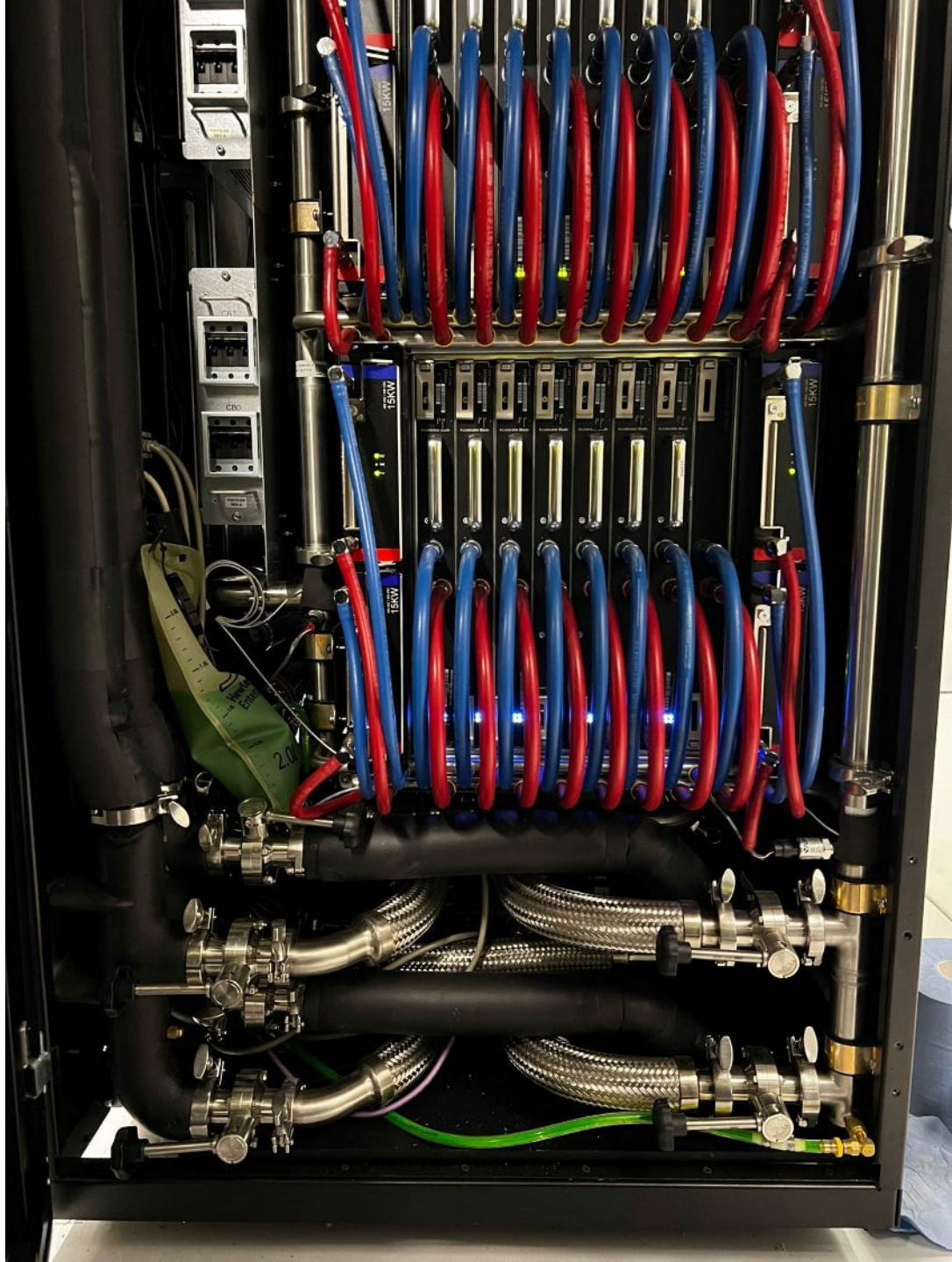
- **>£300M** investment by UK Government in AI capability

- **5,448 NVIDIA Grace-Hopper GH200 GPUs** in a new, 5MW HPE modular data centre (MDC) facility in Bristol, UK
  - ~21 ExaFLOP/s of 8-bit for AI, ~250 PFLOP/s 64-bit
  - 20PB ClusterStor E1000, ~5PB VAST, both all SSD
  - **Phase 1 system of 168 GPUs in service since June 2024**
  - **Phase 2 system** ground broken on 3rd June 2024, being installed winter 2024, bring up spring 2025, in service summer 2025

Department for Science, Innovation, & Technology

Isambard-AI phase 1

- Running AI science since June 2024
- Over 300 registered users and over 30 projects
- Delivering ~110,000 GPU hours per month
- Debuted at #2 on Green500 and #128 on Top500
  - 68 GFLOP/s per Watt
  - 7.4 PetaFLOP/s 64-bit on Linpack
  - 630 PetaFLOP/s 8-bit for AI training and inference



University of BRISTOL

Isambard-AI phase 2 system completed testing in HPE's factory in the Czech Republic

Factory visit on June 11th 2024

University of BRISTOL

BriCS
Bristol Centre for Supercomputing

Isambard-AI phase 2 MDC build – started 3rd June ‘24

Isambard-AI phase 2 MDC being delivered

Food security

Our work will demonstrate the applicability of state-of-the art AI forcefields for drug binding simulations. This is of **direct relevance to the pharmaceuticals and agrochemicals industrial sectors**.

This project focuses on the **automated identification and behaviour analysis** of hundreds of **dairy cattle** on farms using computer vision to analyse their individual coat patterns. Each cow has a unique pattern, which can be automatically identified, enabling long-term monitoring and analysis of individual and herd behaviours. Changes in social behaviour can serve as **early indicators** of **subclinical diseases, such as early-stage mastitis and lameness.**

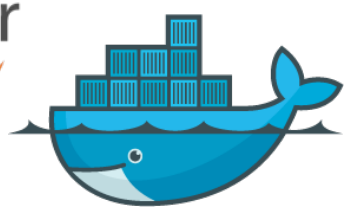
The prevalence of lighter skin tones in open-source **skin cancer** datasets has the potential to introduce **bias into artificial intelligence (AI) models**, which can negatively impact the accuracy of these models for individuals with **darker skin tones**. This project aims to further analyse these datasets for skin tone bias and address previous computational resource limitations.

We will use AF2 derived models to better understand multiprotein complexes. Alphafold Pull Down (AP) will drive novel hypotheses in the **inflammatory process associated with heart disease**. Alphafold Unmasked (AU) will allow us to assess the **impacts of common disease-causing mutations** on predicted complexes. We will look to fine-tune these models to improve performance on a specific sub-class of **disease relevant proteins**

Health

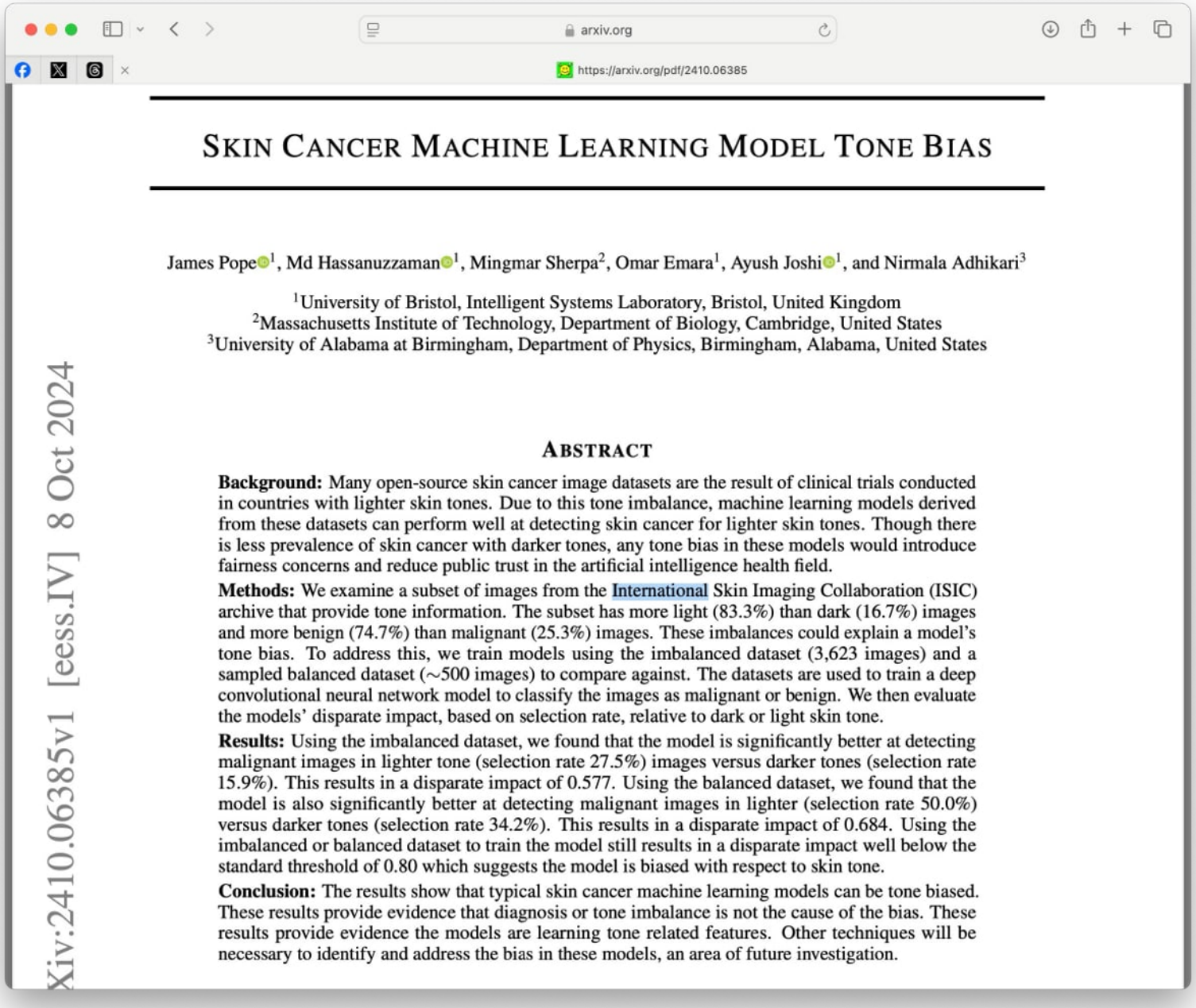
Conservation

Behavioural analysis of great apes is critical to **conservation** efforts since it is one of the earliest indicators of population health. However, manual analysis of video footage is time consuming and limited by the availability of human resource and expertise. This project aims to **develop AI to automate behavioural classification** of great apes in camera trap video footage.



BriCS

Bristol Centre for Supercomputing



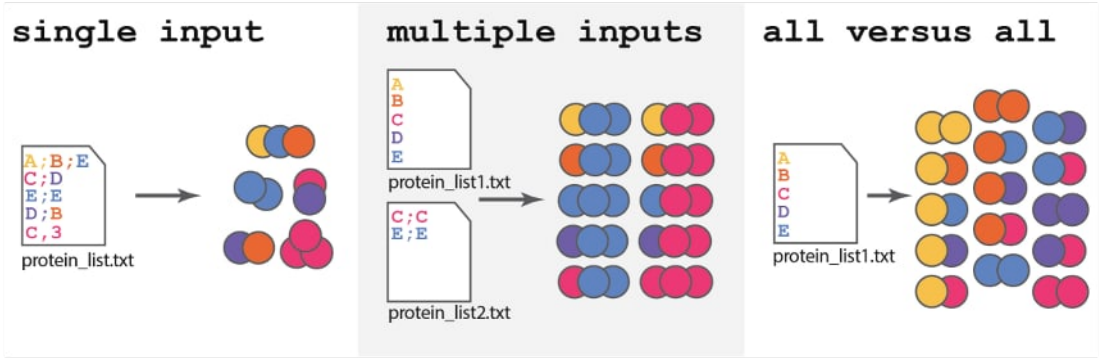
University of BRISTOL

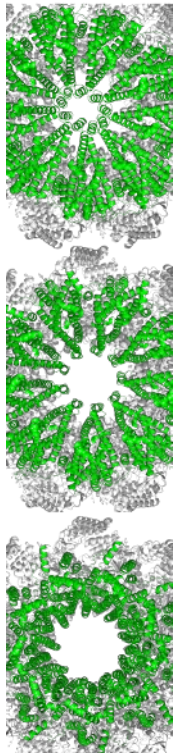
https://arxiv.org/pdf/2410.06385

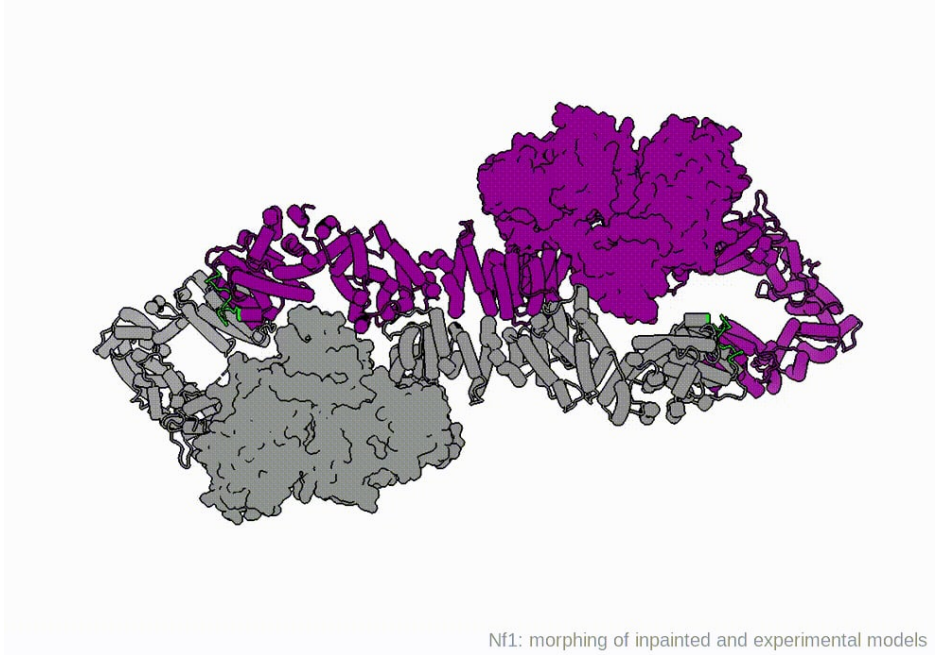
BriCS
Bristol Centre for Supercomputing

AlphaFold for Cardiac disease

- Using AlphaFold and OpenFold
- AlphaFold Unmasked: Predicting how protein **complexes** are formed.
- AlphaPulldown: high-throughput screening of protein-protein **interactions**.
- Using these tools to understand how gene mutations change protein complexes and cause inflammation.
- In diseases such as heart disease, cancer, and diabetes.



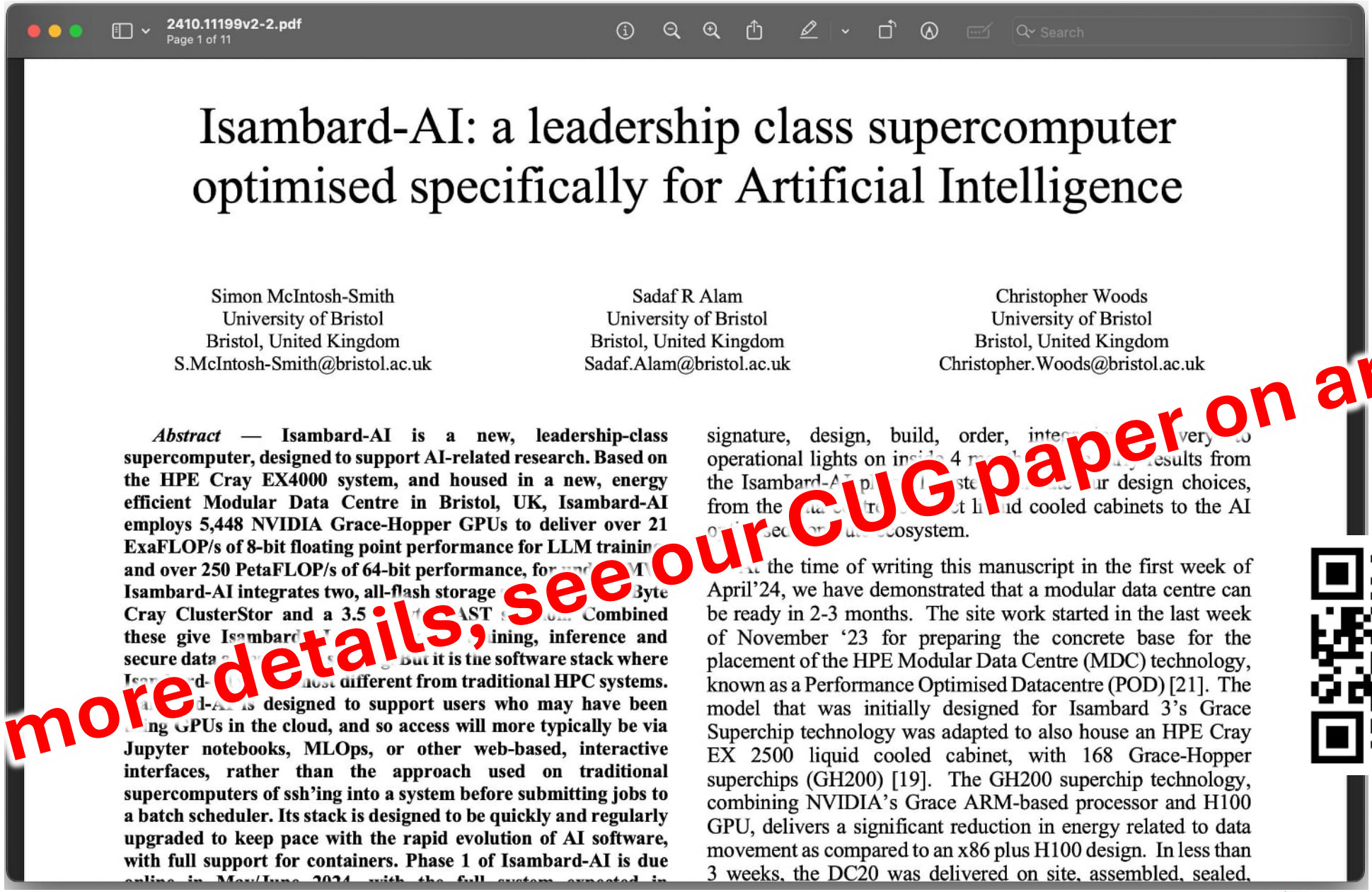




Nf1: morphing of inpainted and experimental models

University of BRISTOL

BriCS
Bristol Centre for Supercomputing



Isambard-AI: a leadership class supercomputer optimised specifically for Artificial Intelligence

Simon McIntosh-Smith
University of Bristol
Bristol, United Kingdom
S.McIntosh-Smith@bristol.ac.uk

Sadaf R Alam
University of Bristol
Bristol, United Kingdom
Sadaf.Alam@bristol.ac.uk

Christopher Woods
University of Bristol
Bristol, United Kingdom
Christopher.Woods@bristol.ac.uk

Abstract **— Isambard-AI is a new, leadership-class supercomputer, designed to support AI-related research. Based on the HPE Cray EX4000 system, and housed in a new, energy efficient Modular Data Centre in Bristol, UK, Isambard-AI employs 5,448 NVIDIA Grace-Hopper GPUs to deliver over 21 ExaFLOP/s of 8-bit floating point performance for LLM trainin and over 250 PetaFLOP/s of 64-bit performance, for MT Isambard-AI integrates two, all-flash storage Byte Cray ClusterStor and a 3.5 AST Combined these give Isambard ining, inference and secure data But it is the software stack where Isambard- ost different from traditional HPC systems. d-AI is designed to support users who may have been ng GPUs in the cloud, and so access will more typically be via Jupyter notebooks, MLOps, or other web-based, interactive interfaces, rather than the approach used on traditional supercomputers of ssh'ing into a system before submitting jobs to a batch scheduler. Its stack is designed to be quickly and regularly upgraded to keep pace with the rapid evolution of AI software, with full support for containers. Phase 1 of Isambard-AI is due**

signature, design, build, order, inte very to operational lights on ins 4 r esults from the Isambard-A p ste r design choices, from the t t li d cooled cabinets to the AI o ecosystem.

t the time of writing this manuscript in the first week of April'24, we have demonstrated that a modular data centre can be ready in 2-3 months. The site work started in the last week of November '23 for preparing the concrete base for the placement of the HPE Modular Data Centre (MDC) technology, known as a Performance Optimised Datacentre (POD) [21]. The model that was initially designed for Isambard 3's Grace Superchip technology was adapted to also house an HPE Cray EX 2500 liquid cooled cabinet, with 168 Grace-Hopper superchips (GH200) [19]. The GH200 superchip technology, combining NVIDIA's Grace ARM-based processor and H100 GPU, delivers a significant reduction in energy related to data movement as compared to an x86 plus H100 design. In less than 3 weeks, the DC20 was delivered on site, assembled, sealed,

For more details, see our CUG paper on arXiv!

https://arxiv.org/pdf/2410.11199

University of BRISTOL

BriCS
Bristol Centre for Supercomputing

Acknowledgements

Isambard-AI is funded by the UK Government's Department of Science, Innovation and Technology (DSIT) via UKRI / STFC. The project has only been possible thanks to a very talented and extremely hardworking team of people across the University of Bristol, HPE, Contour, NVIDIA, the National Composites Centre, and our building contractors, Oakland.

brics-enquiries@bristol.ac.uk

https://www.bristol.ac.uk/supercomputing/

University of BRISTOL

BriCS
Bristol Centre for Supercomputing

Cambridge

Chris Edsall

Dawn Use Cases

- AI@Cam is the cross university organization coordinating AI research
- Headed by Deepmind Professor of Machine Learning Neil Lawrence
- Open call for projects to run on Dawn
  - 47 applications
  - Onboarded in three groups
  - We gained insight in to communities needs
- https://ai.cam.ac.uk/

Stats

- 78 Projects
- 198 Users
- More than 50 000 completed jobs

Dawn Use Cases – UKAEA, Fusion

NekRS simulation of a hypervaportron

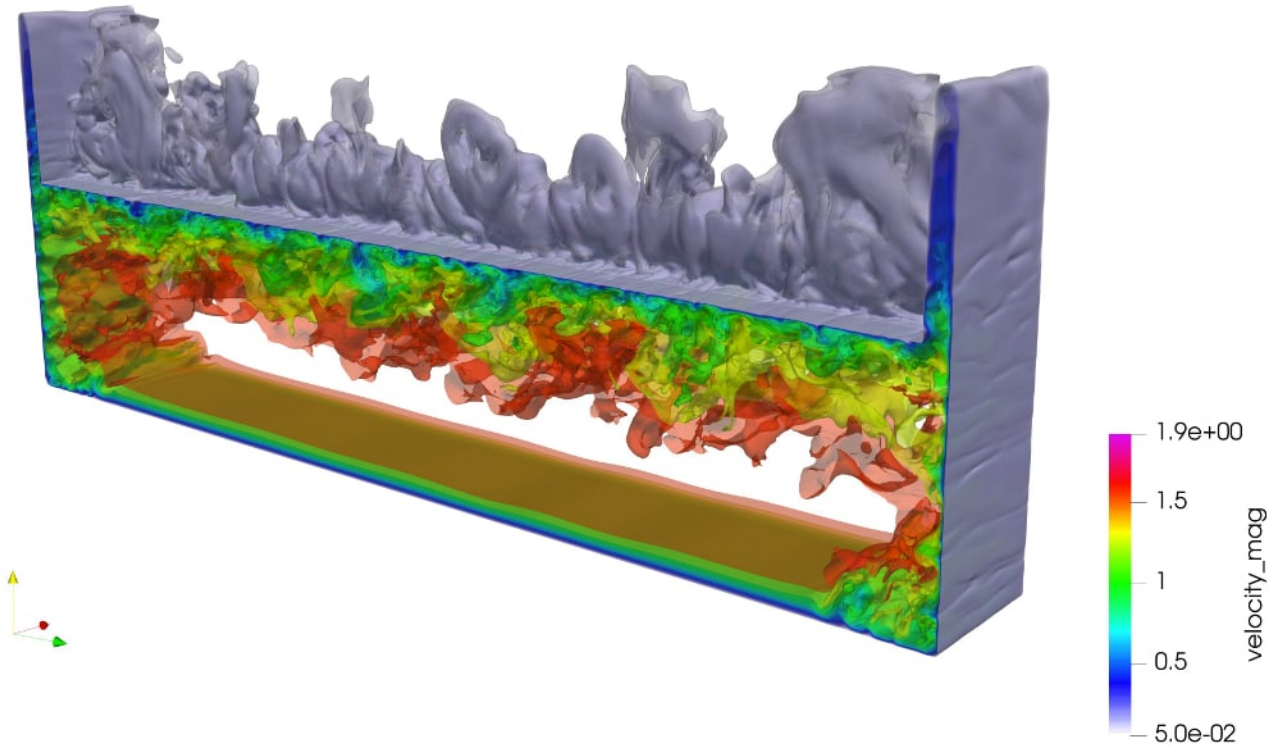


Image: Rupert Eardley-Brunt

Dawn Use Cases – Radiology

Researchers in the Department of Oncology at the University of Cambridge use convolutional neural networks to detect cancer in CT scans.

Dawn Use Cases - LLMs

- **The Alan Turing Institute** is focusing on workloads relevant to the AI research community, e.g. transformers, graph neural networks and computer vision.

Early Access Performance

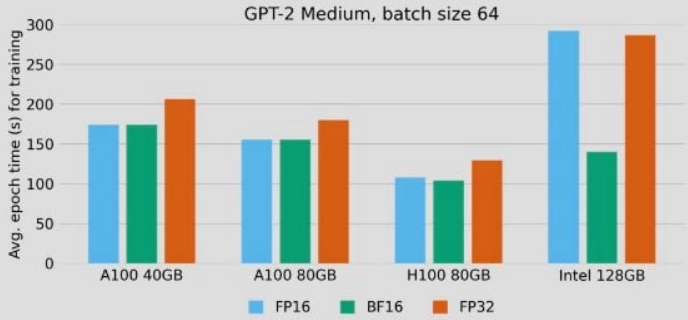
Very pleased with the BF16 results

- Distributed training works well
- Good scaling across nodes
- However, it is only using KYEV*

Results:

- Preliminary, with minimal tuning
- Favorable when using BF16
- Effective for AI workloads

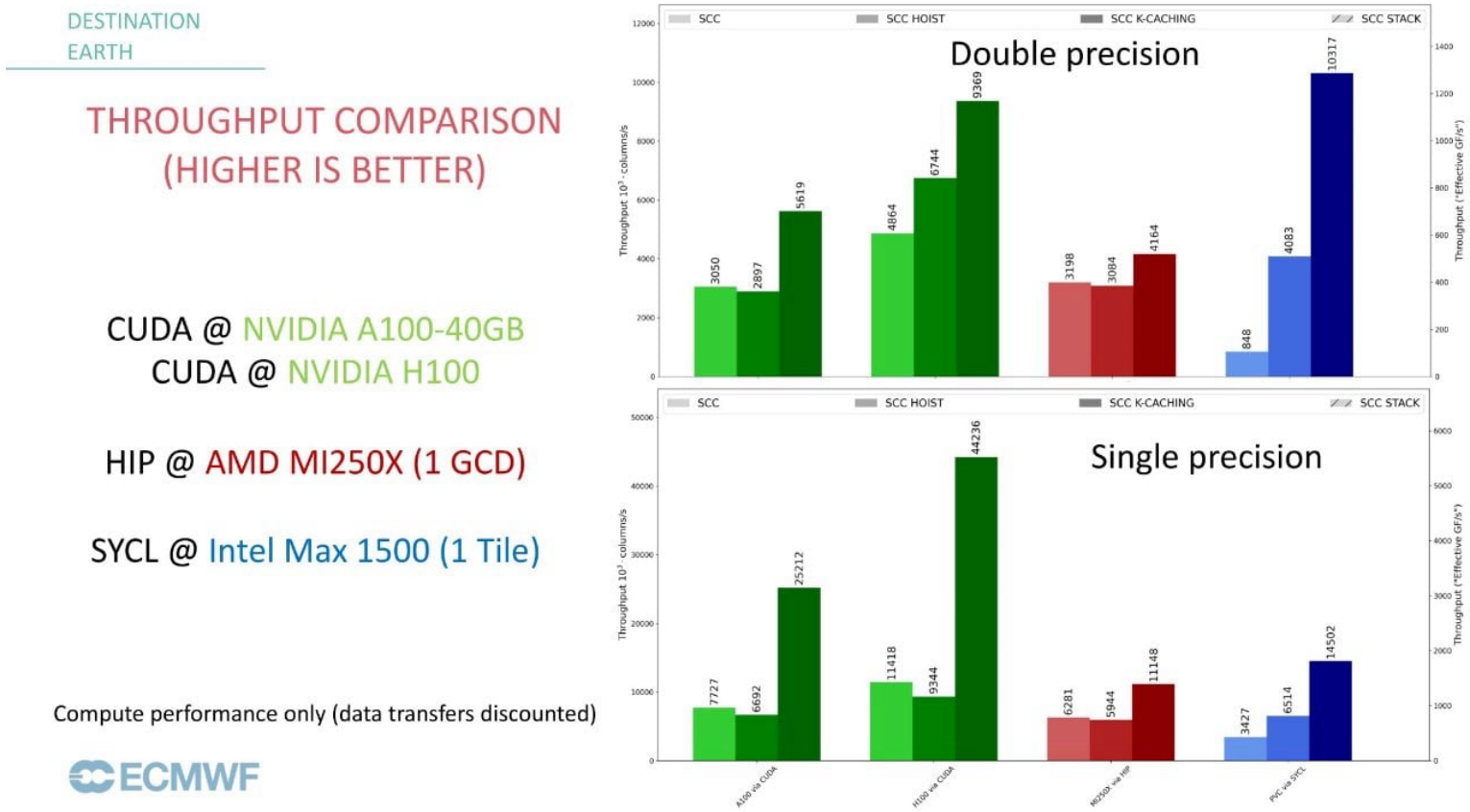
* Know Your Environment Variables



Source: 10.5281/zenodo.11105323

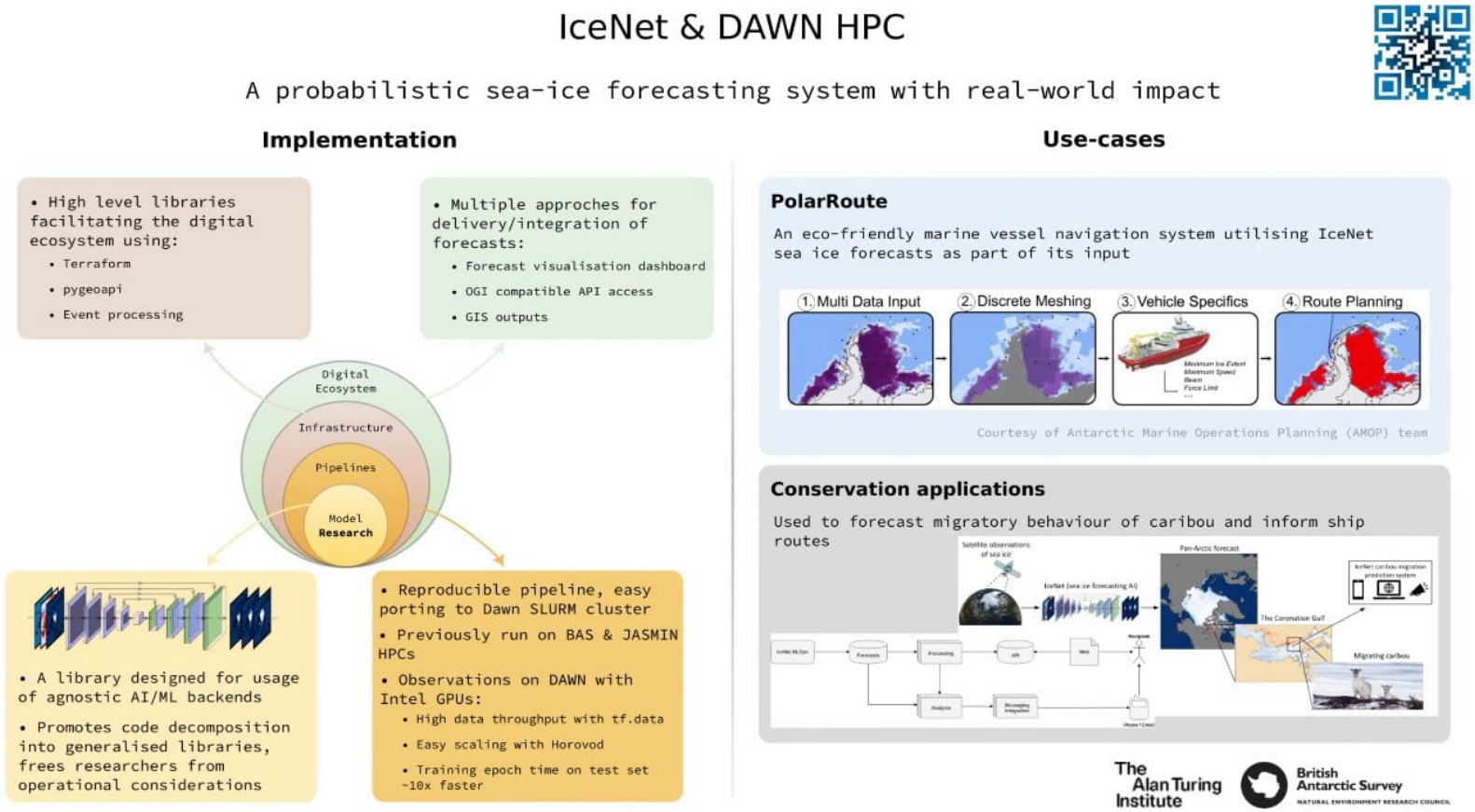
Dawn Use Cases – Weather (cloud microphysics)

- Researchers at the **European Centre for Medium-Range Weather Forecasts** use **Dawn** for the optimisation and performance assessment of a proxy-app for a class of algorithmic components in the IFS.



- Dawn Use Cases – Sea Ice Forecasts

- **British Antarctic Survey** and **The Alan Turing Institute** use **Dawn** to generate AI-powered sea ice forecasts.



- Dawn Use Cases – Conservation Evidence

  - Researchers on the **Nature and AI Project** at the **University of Cambridge** are using **Dawn** to synthesise the global evidence for the effects of conservation efforts on biodiversity.

Ponte Vecchio Across the Atlantic: Single-Node Benchmarking of Two Intel GPU Systems

Thomas Applencourt
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
tapplencourt@anl.gov

Aditya Sadawarte
School of Computer Science
University of Bristol
Bristol, UK
aditya.sadawarte@bristol.ac.uk

Servesh Muralidharan
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
servesh@anl.gov

Colleen Bertoni
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
bertoni@anl.gov

JaeHyuk Kwack
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
jkwack@anl.gov

Ye Luo
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
yeluo@anl.gov

Esteban Rangel
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
erangel@anl.gov

John Tramm
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
jtramm@anl.gov

Yasaman Ghadar
Argonne Leadership Computing Facility
Argonne National Laboratory
Lemont, IL, USA
ghadar@anl.gov

Arjen Tamerus
University of Cambridge
Cambridge, UK
at748@cam.ac.uk

Chris Edsall
University of Cambridge
Cambridge, UK
cje57@cam.ac.uk

Tom Deakin
School of Computer Science
University of Bristol
Bristol, UK
tom.deakin@bristol.ac.uk

https://doi.org/10.1109/SCW63240.2024.00184

Software Solutions

- **MyAccessID**: Federated Identity Provider
- **Keycloak**: Serves as the main broker for authentication
- **LDAP**: Stores user information for central account management.
- **Waldur**: Handles resource offering, project management, and integrates with SLURM. Provides an interface for Principal Investigators (PIs) to manage research projects and invite users
- **Open OnDemand**: Provides web-based HPC/AI access
- **SLURM Cluster**: Provides SSH Certificates based access to resources for job scheduling
- **K8s**: Providing K8s as a service

UK Research and Innovation

UKRI's Digital Research Infrastructure Programme

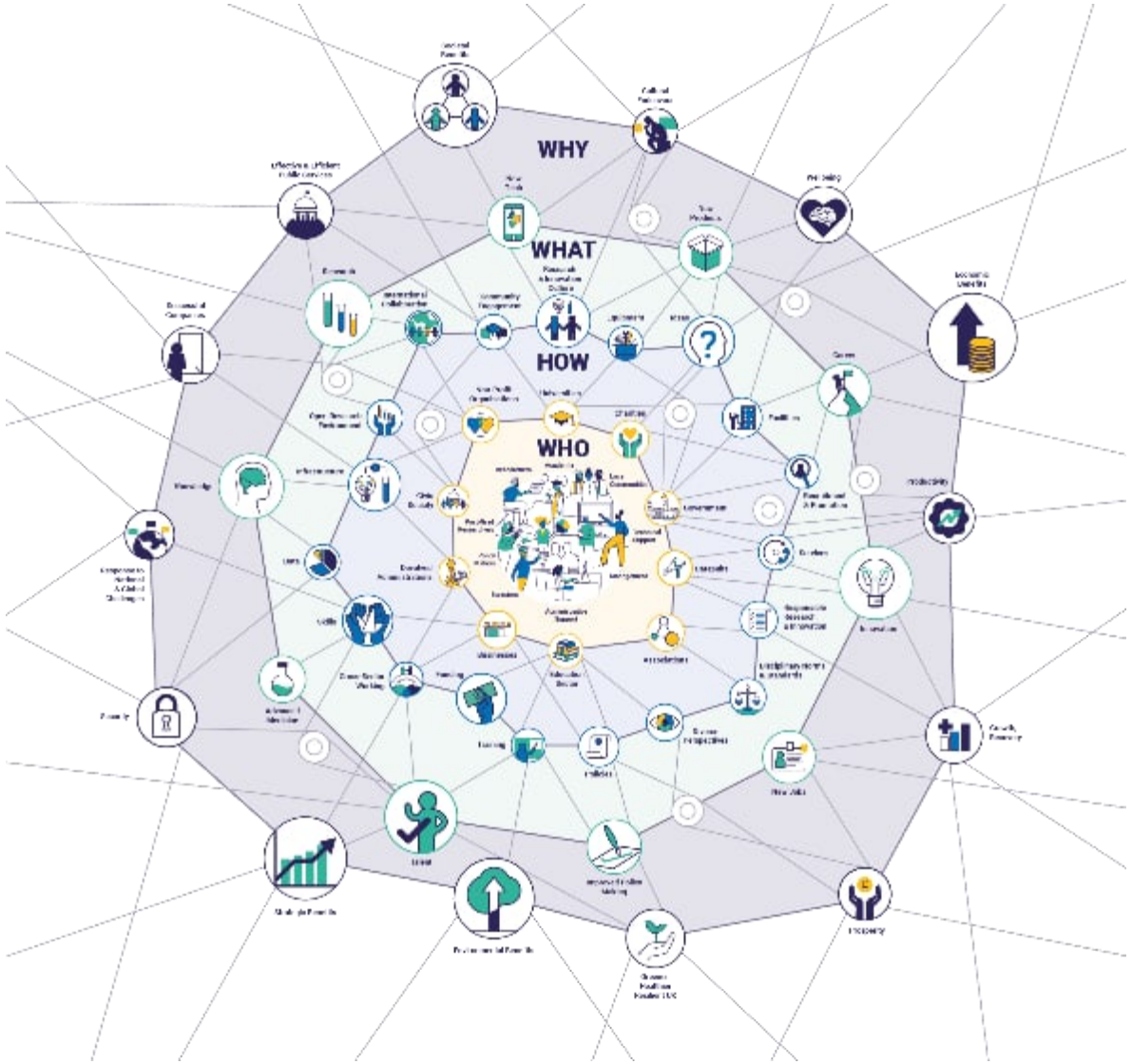
Richard Gunn, DRI Programme Director
richard.gunn@ukri.org

CIUK, Manchester Central
5 December 2024

UK Research and Innovation

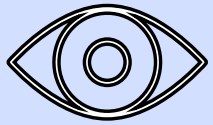
UKRI supports all parts of the R&I system in a variety of ways

- The UKRI portfolio provides support for all parts of the research and innovation system including:
  - Grants to fund research projects
  - Funding the building of new infrastructure necessary for R&I
  - Supporting the training of new generations of researchers
  - Grants to fund innovation and the creation of new products
- It also funds research across all areas/disciplines of research, including large societal challenges such as Net-Zero, Artificial Intelligence, or Healthy Ageing.



UK Research and Innovation

UKRI's vision for Digital Research Infrastructure



Our Vision

A coherent state-of-the-art national Digital Research Infrastructure (DRI) that will seamlessly connect researchers and innovators to the computers, data, tools, techniques and skills that underpin the most ambitious and creative research

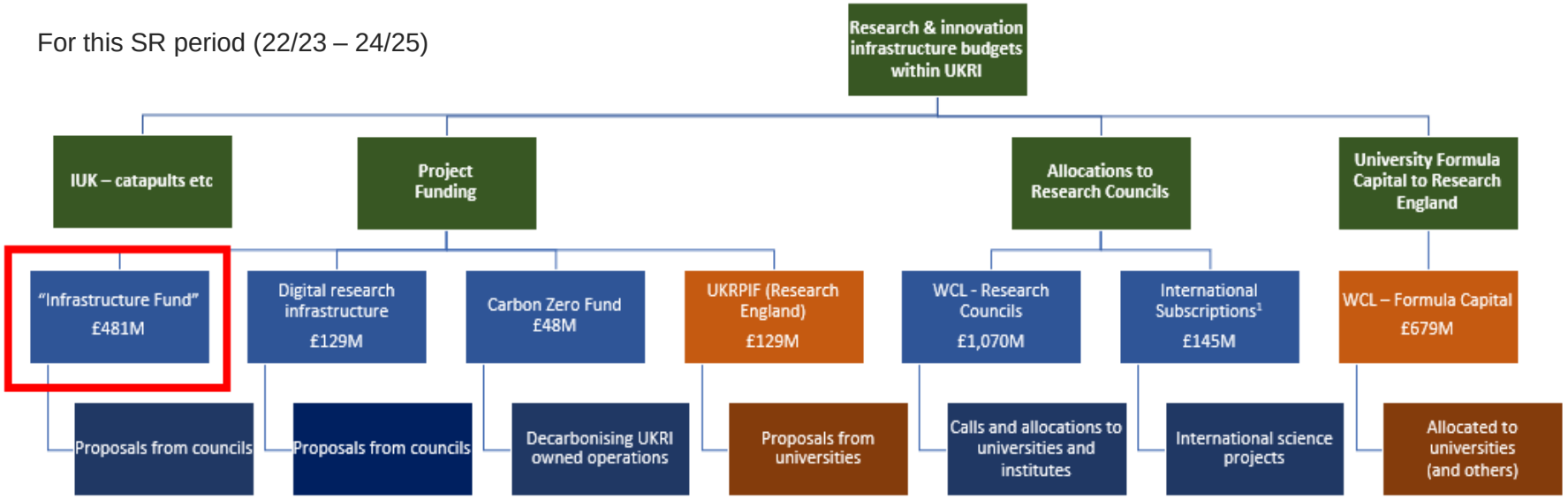
Our Approach

DRI is a system that includes large-scale computing (LSC), data storage, facilities, software, networks, skilled DRI professionals, and other components.
Working with Councils, we will achieve our vision by evolving existing infrastructures to support new communities of practice and, subject to funding, by investing in new capabilities.

UK Research and Innovation

UKRI Infrastructure Funding Landscape

For this SR period (22/23 – 24/25)



Over this SR we will grow our infrastructure budget by over £200m to reach over £1.1bn in 2024/25

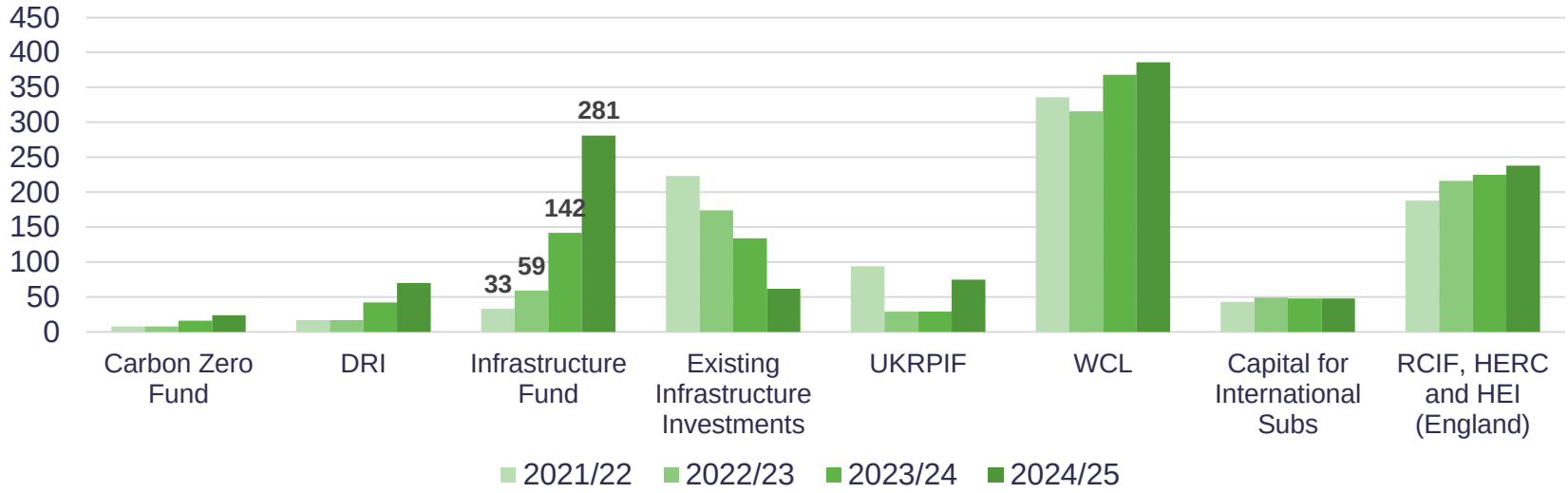
[1]Includes resource funding.

UKRPIF = UK Research Partnership Investment Fund
WCL = World Class Labs

UKRI Infrastructure Fund

A pan-UKRI budget to support step-changes in research and innovation infrastructure capability across the disciplinary spectrum.

UKRI Infrastructure Funding Allocations 2021/22 – 2024/25



UKRI Digital Infrastructure Phase 2 Portfolio

Federated Data Services

To deliver aspects of the DRI strategy relating to data infrastructure and implementation of FAIR.

Total for this SR	Indicative forward profile
£38.1M	£31.6M

National Computational Research Services

Ensuring the UK has the compute services to facilitate the ambitions of our R&I community.

Total for this SR	Indicative forward profile
£23M	£16M

Software for future Large-scale compute

Creation of a mission-led software programme for large-scale accelerated computing.

Total for this SR	Indicative forward profile
£8.5M	£13.5M

Supporting DRI Professionals

Ensuring the UK has the skills base to deliver effective world-leading research and innovation.

Total for this SR	Indicative forward profile
£12.2M	£21.5M

Software, Networks, Security and Net Zero

Secure access and connectivity, and software as an infrastructure and UKRI net zero ambitions.

Total for this SR	Indicative forward profile
£11M	£18.2M

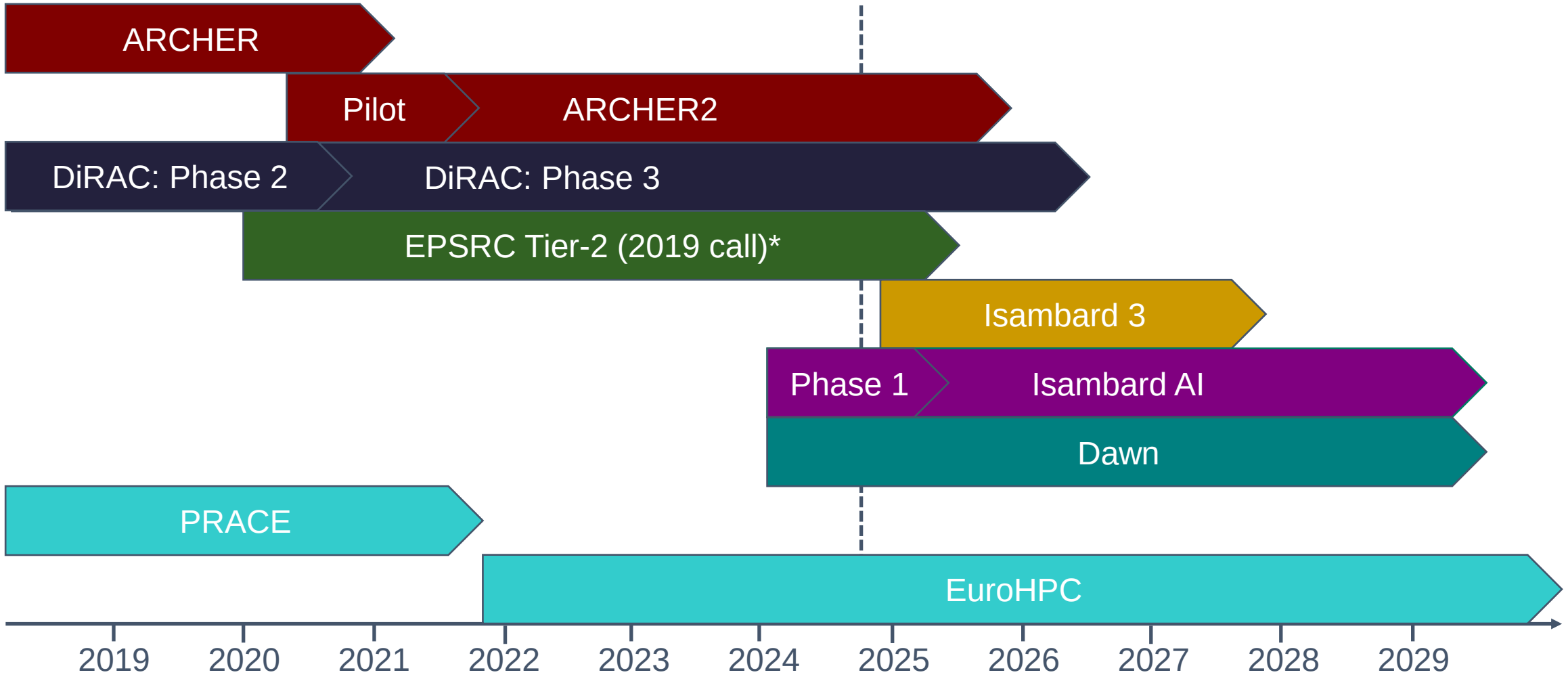
UK Research and Innovation

2024 in review

- Launched funding opportunities for Net-Zero DRI leadership and community building, Research Technical Professional (RTP) support, federation of compute services, and software support based on usage.
- Launched the first DRI programme led funding competition for novel approaches to software and skills and first cross-UKRI ‘Access to HPC’ call.
- Published a comprehensive Jisc-led study and recommendations on federating our DRIs. See: https://www.ukri.org/publications/dri-programme-response-to-mapping-federation-journeys-report/
- Installation of Isambard 3: UKRI’s first compute service funded through the DRI programme and accessible to all of UKRI’s communities from the outset.
- Commenced scoping work (partnering with OpenUK) to develop a strategic approach to open technologies, focussing on sustainability, curation and leadership.
- Recruited a second cohort of members of the Advisory Group for DRI (AGD).
- Commenced the development of a ‘living library’ of open source compute benchmarks spanning UKRI’s communities of practice.

UK Research and Innovation

Existing UK R&I compute provision



UK Research and Innovation

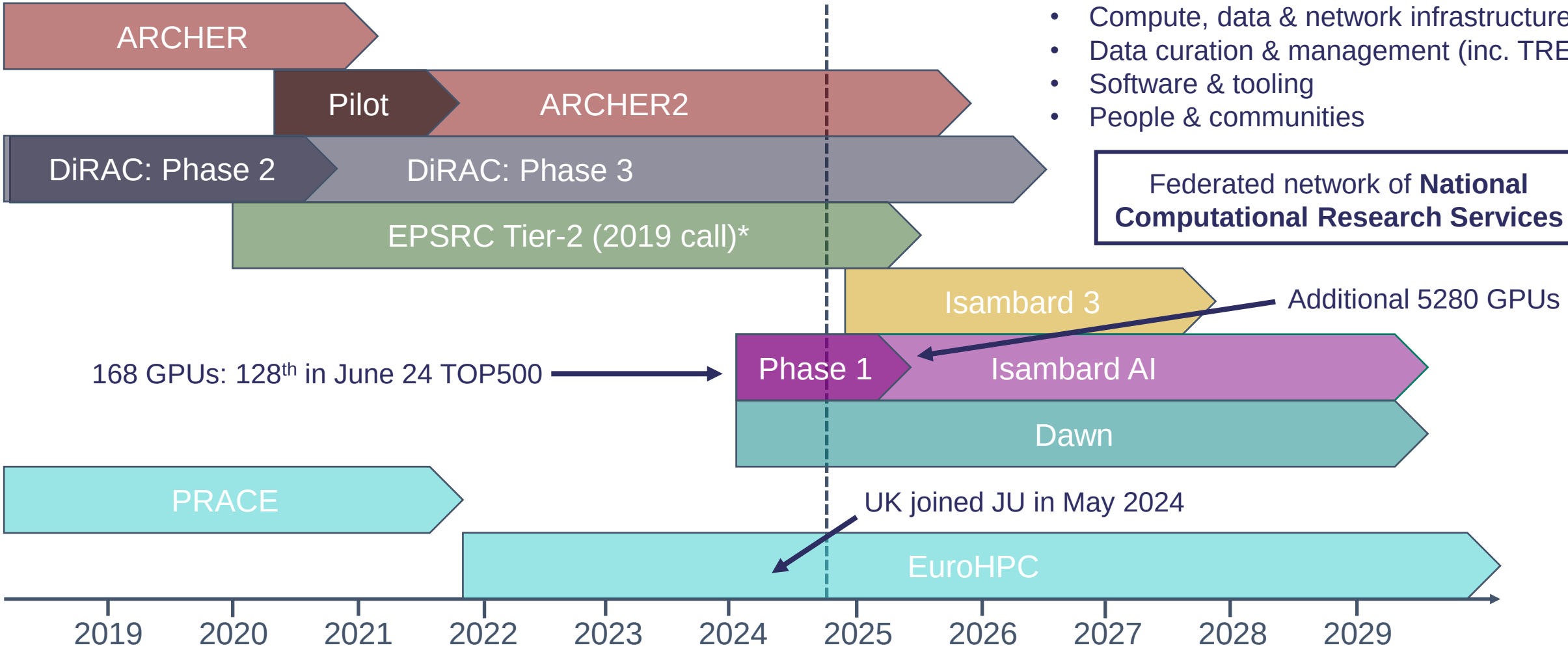
*Isambard 2 ended March 2024, JADE2 ends January 2025

Existing UK R&I compute provision

UKRI DRI investment in:

- Compute, data & network infrastructure
- Data curation & management (inc. TREs)
- Software & tooling
- People & communities

Federated network of **National Computational Research Services**



*Isambard 2 ended March 2024, JADE2 ends January 2025

UK Research and Innovation

Next steps

- We are developing the case for further phases of investment in the UKRI DRI programme across the two phases of the Spending Review, in order to deliver a ‘**data as a service’** model for the UK community, **sustaining and enhancing our large-scale compute system**, and developing the underpinning tools, techniques, services and skills required for a federated **secure, sustainable and collaborative DRI**.
- We are engaging on **how we invest**, as well as **what we are investing in**.
    - Continue to work closely with DSIT to deliver major compute investment, including implementation of AIRR access policy.
    - Working with OpenUK, SSI & ReSA partners on software sustainability & open technologies.
    - We will launch a ‘statement of requirements’ process for external parties to contribute ideas for the development of our DRI in early 2025 and host a DRI Congress in 2H 2025.

UK Research and Innovation

Priorities of the DRI Programme

A globally competitive Digital Research Infrastructure (DRI) is essential for the UK to support cutting-edge research and innovation. The DRI Programme aims to provide a mechanism for promoting collaboration and eliminating research siloes throughout the research and innovation communities, ensuring that DRI is made to efficiently fit the needs of its users and addresses current challenges and opportunities.

Through a consultation process involving researchers, council representatives, and external experts, the Programme has identified four key priorities to guide its efforts. These priorities, while distinct, are interconnected and work together to enhance the DRI and deliver better research solutions for the UK research community.

Interconnected DRI

Human DRI

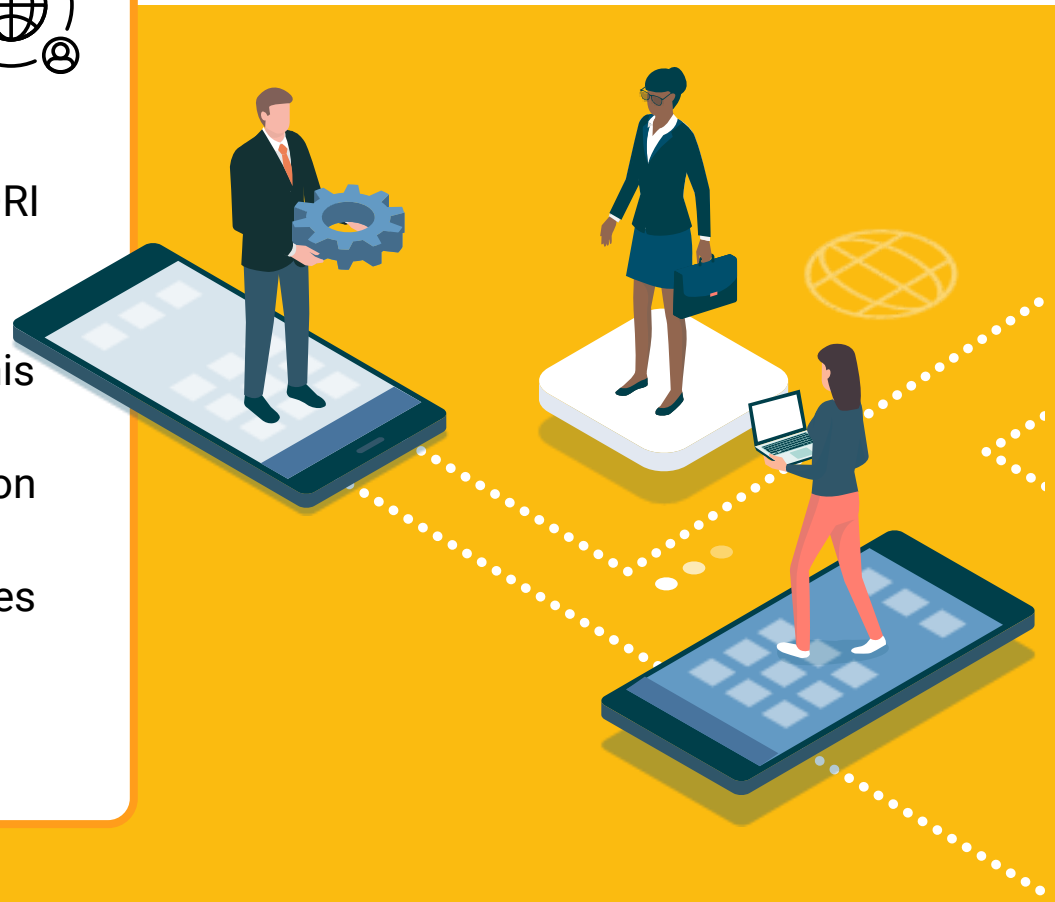
FAIR DRI

Sustainable DRI

UK Research and Innovation

Interconnected DRI

We aim to break down silos and promote collaboration within the DRI landscape by creating a more interconnected, open, inclusive, and user-friendly infrastructure. This involves fostering community engagement, promoting the creation of federated infrastructure and ensuring that DRI funding addresses the technical challenges of operating across disciplines.



UK Research and Innovation

Human DRI

We understand that DRI starts with the people that build, maintain, and upgrade it. By recognising their crucial role within DRI, promoting their careers, and supporting their professional development, we aim to foster a supportive culture and ensure a skilled and diverse DRI workforce.

UK Research and Innovation

FAIR DRI

We will ensure that research infrastructure is easily accessible, interoperable, and reusable by driving adoption of the FAIR principles, enhancing research efficiency, and fostering a more collaborative research environment.

UK Research and Innovation

Sustainable DRI

We focus on ensuring the long-term health of DRI by addressing efficiency, security, environmental and financial challenges to create a sustainable DRI ecosystem.

UK Research and Innovation

UK Research and Innovation

Panel discussion:

Building FAIRer, human-centric, interoperable and sustainable DRI

UK Research and Innovation

CIUK 2024 Presentations

Martyn Guest (University of Cardiff)

Community Code Performance on Multi-core Processors. An Analysis of Computational Chemistry and Ocean Modelling Applications

Abstract: This session will overview the measured performance of a number of popular community codes on a variety of HPC systems, with our established analysis based on both computational chemistry and ocean modelling applications. The former features codes from Molecular Dynamics (DL_POLY, AMBER, and GROMACS), molecular electronic structure (GAMESS-UK) and Materials Science (VASP), while NEMO is the representative code from the ocean modelling community.

The variety of systems considered focus on both the Intel Emerald Rapids and recently released Granite Rapids processors, together with the AMD EPYC Genoa family of CPUs. Using the Intel Skylake Gold 6148 and AMD EPYC Rome 7502 as the baselines, a performance assessment is made across a variety of Granite Rapids (6980, 6972 and 6960) and Emerald Rapids (8562Y, 8568Y and 8592) SKUs, with system interconnects from both NVIDIA Networks and Cornelis Networks. Attention is also focused on systems featuring the AMD Genoa EPYC processors, including the Genoa 32-core 9354, the 48-core 9454 and 96-core 9654 and 9684X SKUs.

The benefits of the Intel® oneAPI Toolkit and Linaro Performance Reports are demonstrated throughout this analysis. To best capture a 'like for like' comparison amidst the extensive array of core densities, our analysis remains based on both a “node-by-node” and the more traditional “core-by-core” consideration.

Bio: Professor Martyn Guest has led a variety of high performance and distributed computing initiatives in the UK. He spent three years as Senior Chief Scientist and HPC Chemistry Group Leader at PNNL, before returning to the UK as Associate Director of Daresbury’s Computational Science and Engineering Department. Martyn joined Cardiff University in April 2007 as their Director of Advanced Research Computing, retaining this position until February 2023. He is also Technical Director of the Supercomputing Wales programme and is co-I on the Isambard-2 and Isambard-3 systems at the GW4 Tier-2 HPC regional centre.

COMPUTING INSIGHT UK 2024
Catalysing Research
5 - 6 DECEMBER 2024
Manchester Central, UK
www.ukri.org/CIUK

UKRI Science and Technology Facilities Council

Community Code Performance on Multi-core Processor

An Analysis of Computational Chemistry and Ocean Modelling Applications

Martyn Guest & Jose Munoz Criollo

Advanced Research Computing @ Cardiff (ARCCA) &

Supercomputing Wales

Introduction and Overview

- Presentation part of our ongoing assessment of the performance of **community codes** on multi-core processors.
- Regular presentations at Daresbury's MEW and successor CIUK conferences – **nine "new" systems** under analysis since CIUK'23.
- Focus on systems featuring **processors from Intel** (Emerald Rapids & Granite Rapids SKUs) and **AMD** (EPYC Genoa SKUs) with Infiniband (EDR, HDR, NDR) & Cornelis Networks interconnects.
  - Baseline clusters: Skylake (SKL) **Gold 6148/2.4 GHz** and **AMD EPYC Rome 7502 2.5Gz** cluster – "Hawk" – at Cardiff University.
  - **Three Intel Granite Rapids** clusters – the 72-core Platinum **6960** (2.7 GHz), 96-core **6972** (2.4 GHz) and 128-core **6980** (2.0 GHz).
  - **Three** Intel Xeon Emerald Rapids clusters, the 32-core Platinum **8562Y** (2.8 GHz), 48-core **8568Y** (2.3GHz) and 64-core **8592** (1.9 GHz) plus other Sapphire-Rapids systems.

Introduction and Overview

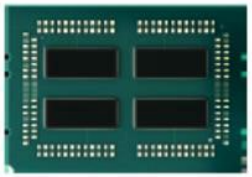
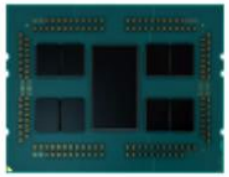
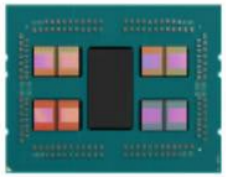
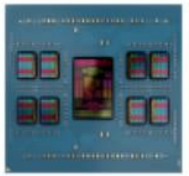
- **Four** AMD **Genoa clusters** featuring the 32-core **9354** (3.25 GHz), 48-core **9454** (2.85 GHz) and the 96-core **9654** (2.4 GHz) and **9684X** (2.55 GHz) SKUs.
- **NVIDIA Arm Grace access** through Isambard 3 project underway.

- Consider performance of both synthetic and **end-user applications:**
  - Molecular simulation (**DL_POLY, AMBER & GROMACS** Molecular Dynamics codes);
  - Materials modelling (**VASP**) & electronic structure (**GAMESS-UK**);
  - Ocean modelling codes including **NEMO.**
- Scalability analysis by **processing elements (cores)** and by **nodes** (ARM Performance Reports). Baselined against **V100** NVIDIA GPUs.
- **Pricing** – remains of course a key issue but lies outside the scope of this presentation.

Methodology and Approach

1. Provide guidance based on evaluating performance that a **standard user** would experience on the systems
2. Target performance regime – **mid-range clusters**. No real effort invested in optimising applications having used standard implementations when available
3. All benchmarks run on systems in general production i.e. not dedicated to this exercise – used standard Slurm job schedulers
4. **Performance comparisons** across a spectrum of MPI versions with Intel Parallel Studio XE e.g. 2018/4, 2019/5, 2019/12 & 2020/4 PLUS OneAPI proved **challenging**.
    - Initial problems encountered on **AMD Milan** systems when working code using Intel 2019/5 on AMD Rome systems failed on Milan. Resolved using. **Intel oneAPI. But performance issues remain e.g.,** major decline in VASP performance on AMD EPYC when using later versions of Intel MPI.
5. **Performance issues** encountered when using the **Intel Endeavour system.**
6. Consistency through use of **EasyBuild** & **SPACK Package Manager for HPC** demonstrated throughout this analysis.

AMD “GENOA” EPYC SERVER CPUS

	AMD EPYC 7001 ‘NAPLES’	AMD EPYC 7002 ‘ROME’	AMD EPYC 7003 ‘MILAN’	AMD EPYC 9004, 8004 ‘GENOA’, ‘SIENA’
				
Core Architecture	‘Zen’	‘Zen 2’	‘Zen 3’	‘Zen 4’ and ‘Zen 4c’
Cores	8 to 32	8 to 64	8 to 64	8 to 128
IPC Improvement Over Prior Generation	N/A	~24%[ROM-236]	~19%[MLN-003]	~14%[EPYC-038]
Max L3 Cache	Up to 64 MB	Up to 256 MB	Up to 256 MB	Up to 384 MB (EPYC 9004) Up to 128 MB (EPYC 8004)
Max L3 Cache with 3D V-Cache™ technology			768 MB	Up to 1152 MB
PCIe® Lanes	Up to 128 Gen 3	Up to 128 Gen 3	Up to 128 Gen 4	Up to 128 Gen 5 8 bonus lanes Gen 3
CPU Process Technology	14nm	7nm	7nm	5nm
I/O Die Process Technology	N/A	14nm	14nm	6nm
Power (Configurable TDP [cTDP])	120-200W	120-280W	155-280W	70-400W
Max Memory Capacity	2 TB DDR3-2400/2666	4 TB DDR4-3200	4 TB DDR4-3200	6 TB DDR5-4800

AMD EPYC Genoa: IPC Improvements

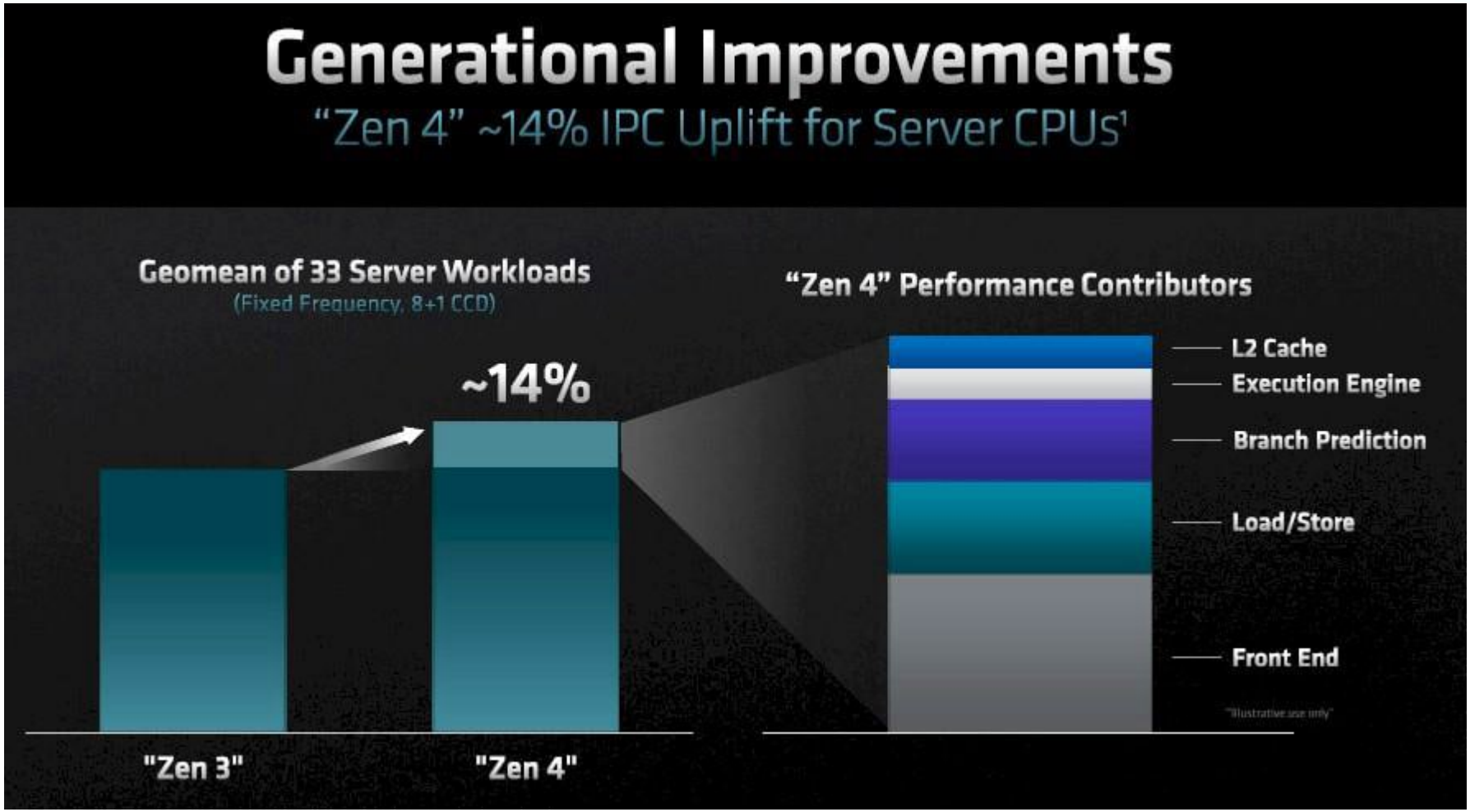


Figure. The move to Genoa provided a big leap in performance, starting with the move to the "Zen 4" cores, which are provided a 14 percent increase in the instructions per clock (IPC) compared to the prior "Zen3" cores used in the Milan Epyc 7003s.

AMD EPYC Turin: IPC Improvements

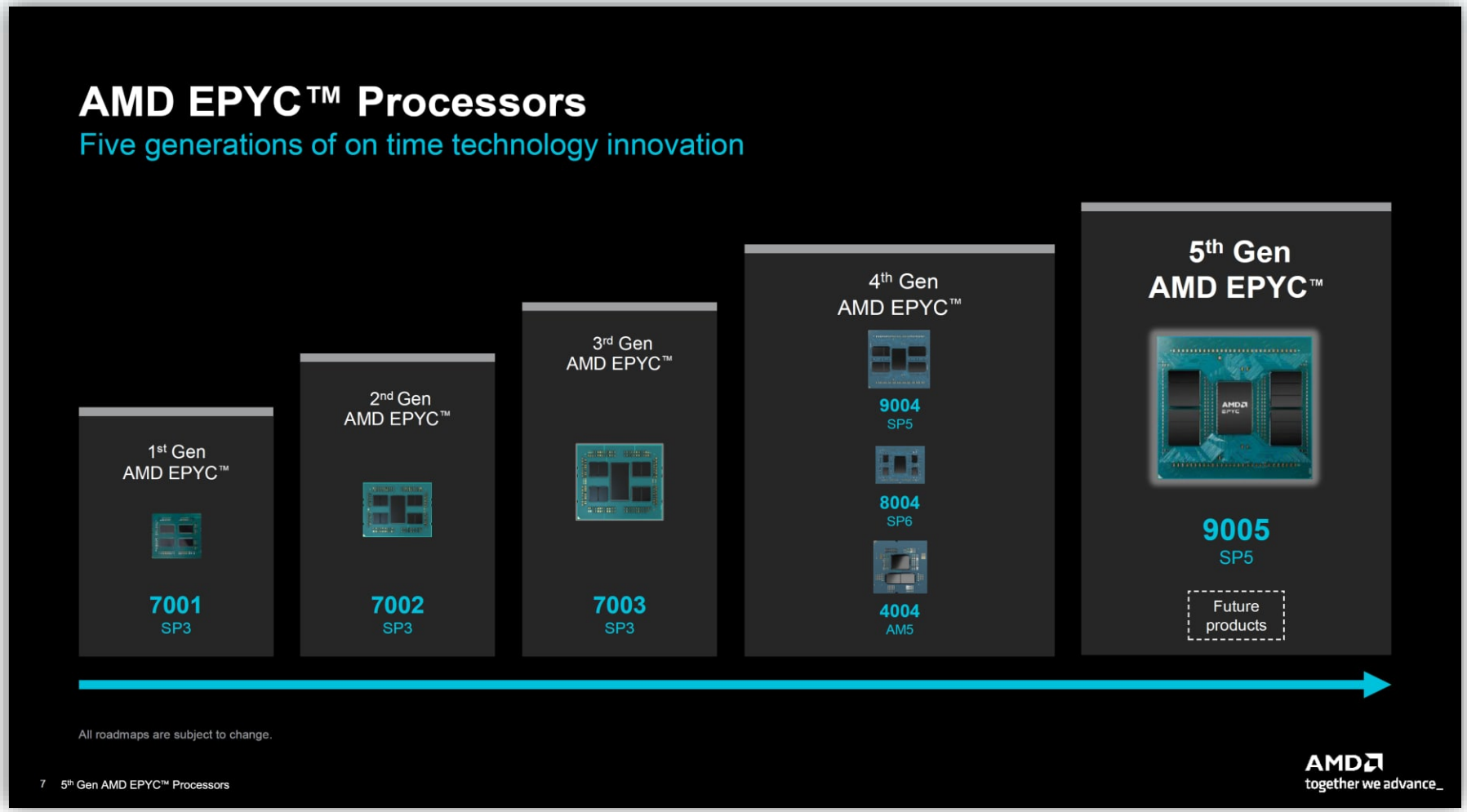


Figure. The move to Turin is a further big leap in performance, starting with the move to the "Zen 5" cores, which are providing a 17 percent increasing in the instructions per clock (IPC) compared to the prior "Zen4" cores used in the Genoa EPYC 9004s.

AMD EPYC Turin: IPC Improvements

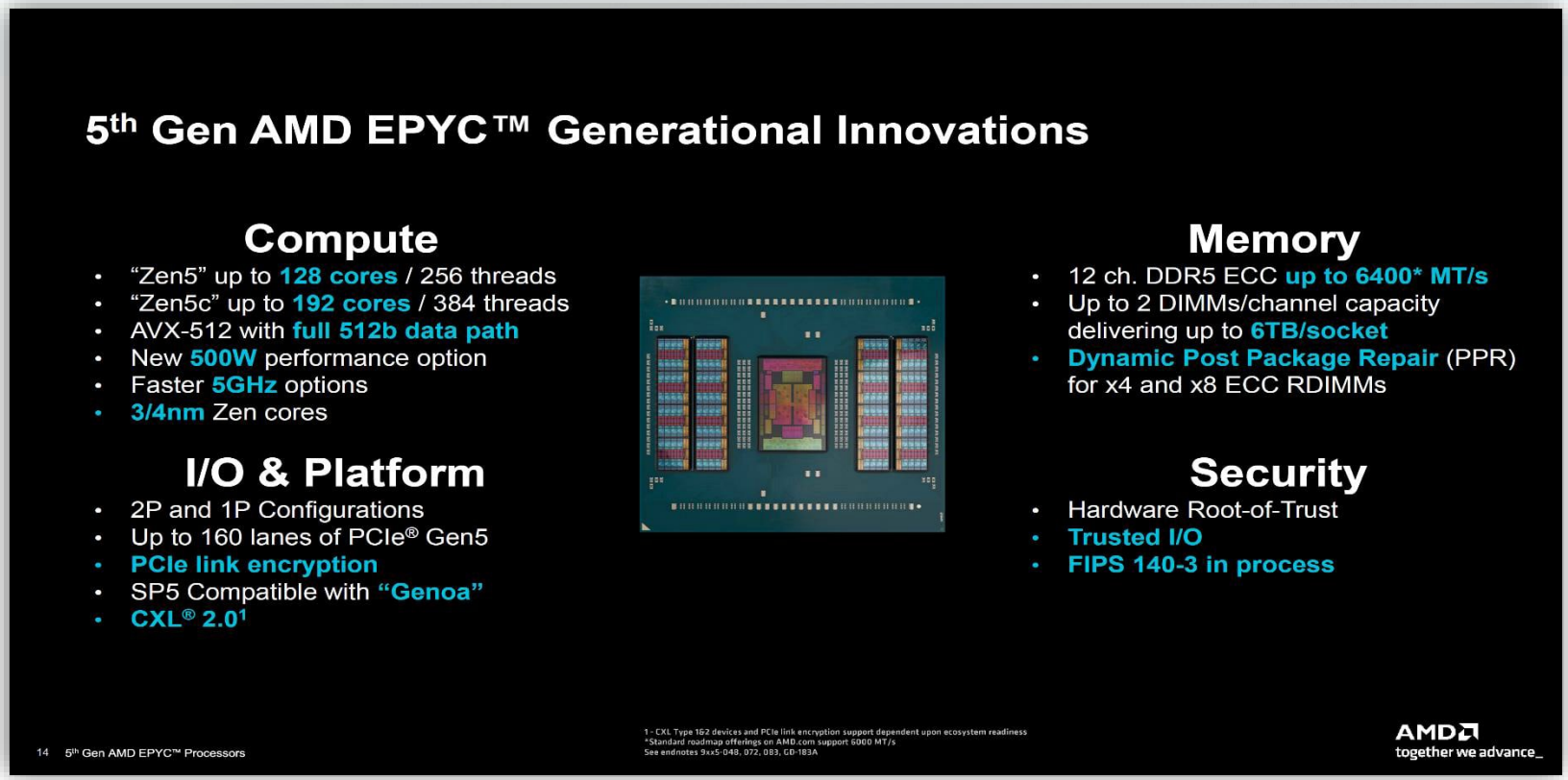
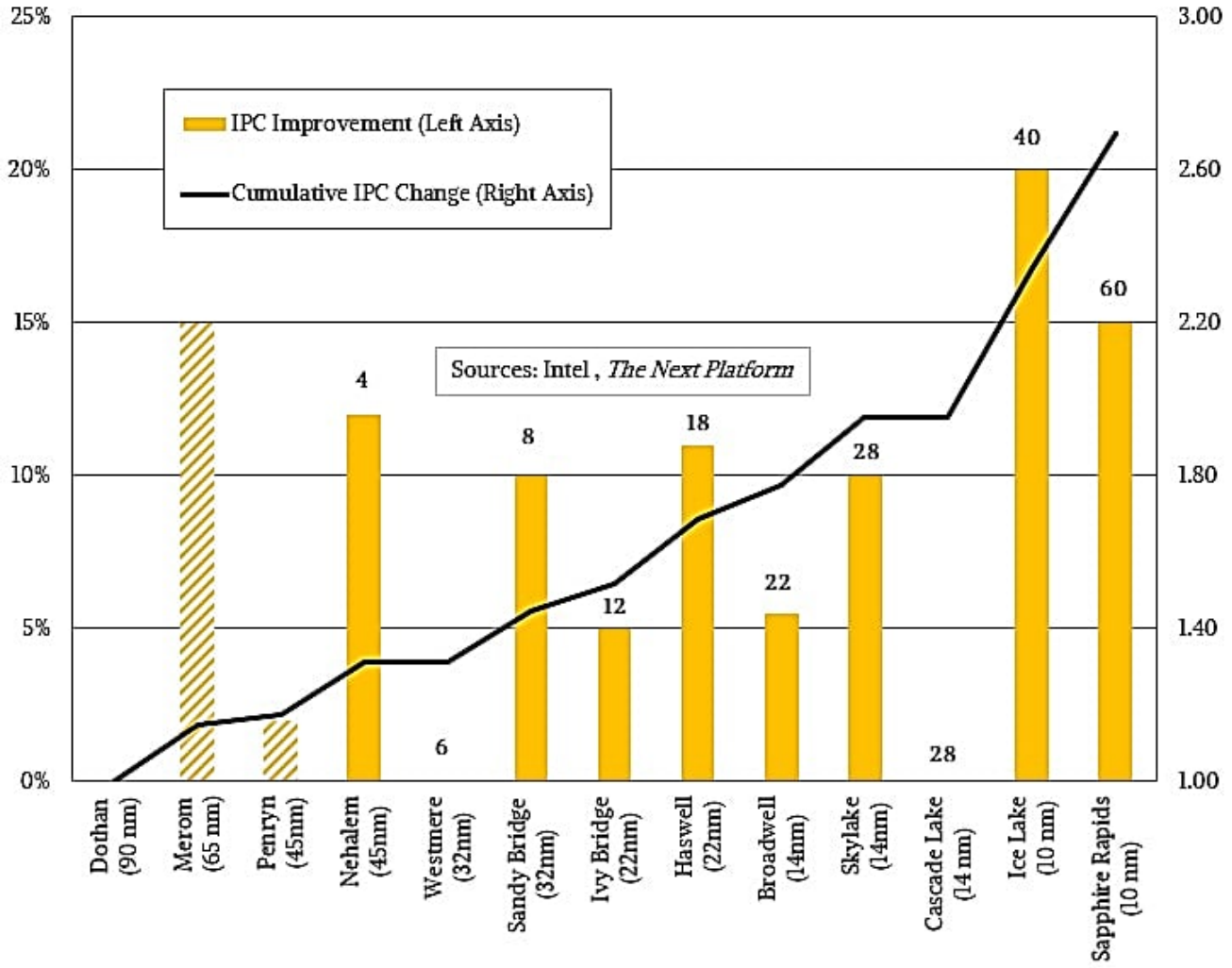


Figure. 5th Gen EPYC "Turin" Family: Up To 192 Cores, 500W TDPs and 5 GHz Clock Speeds. Two solutions (i) the 4nm version with up to 16 "Zen 5" CCDs, offering up to 128 cores and 256 threads "Scale-Up" variant, and (ii) the "Scale-Out" variant which utilizes the 3nm "Zen 5C" cores with up to 12 CCDs, offering up to 192 cores and 384 threads.

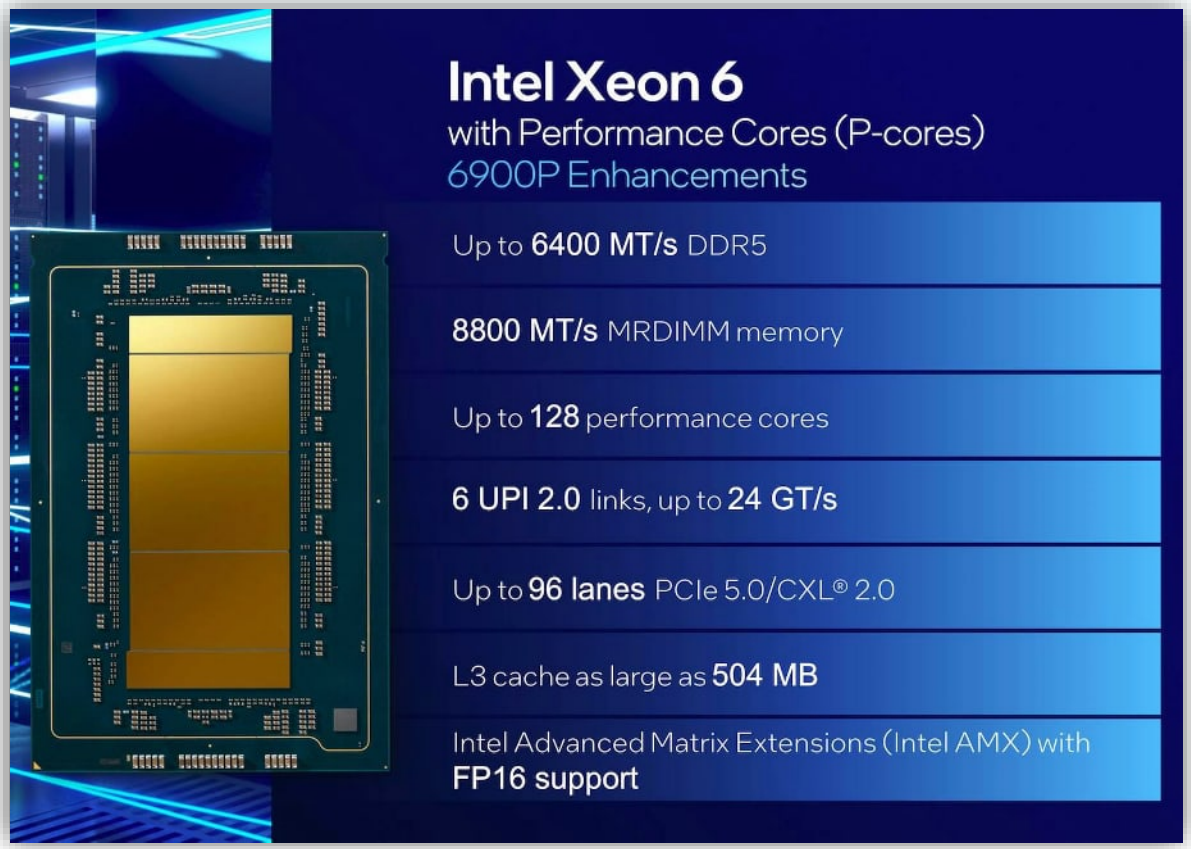
- **EPYC Zen 5C:** Up To 192 Cores, 384 MB L3 Cache (+50% More Cores / L3 Cache Versus Zen 4C)
- **EPYC Zen 5:** Up To 128 Cores, 512 MB L3 Cache (+33% More Cores / L3 Cache Versus Zen 4)

IPC Improvements - Intel Core Generations



Instructions per clock (IPC) improvement per generation versus cumulative IPC over time. Maximum core count per generation shown above the bars for each Xeon chip.

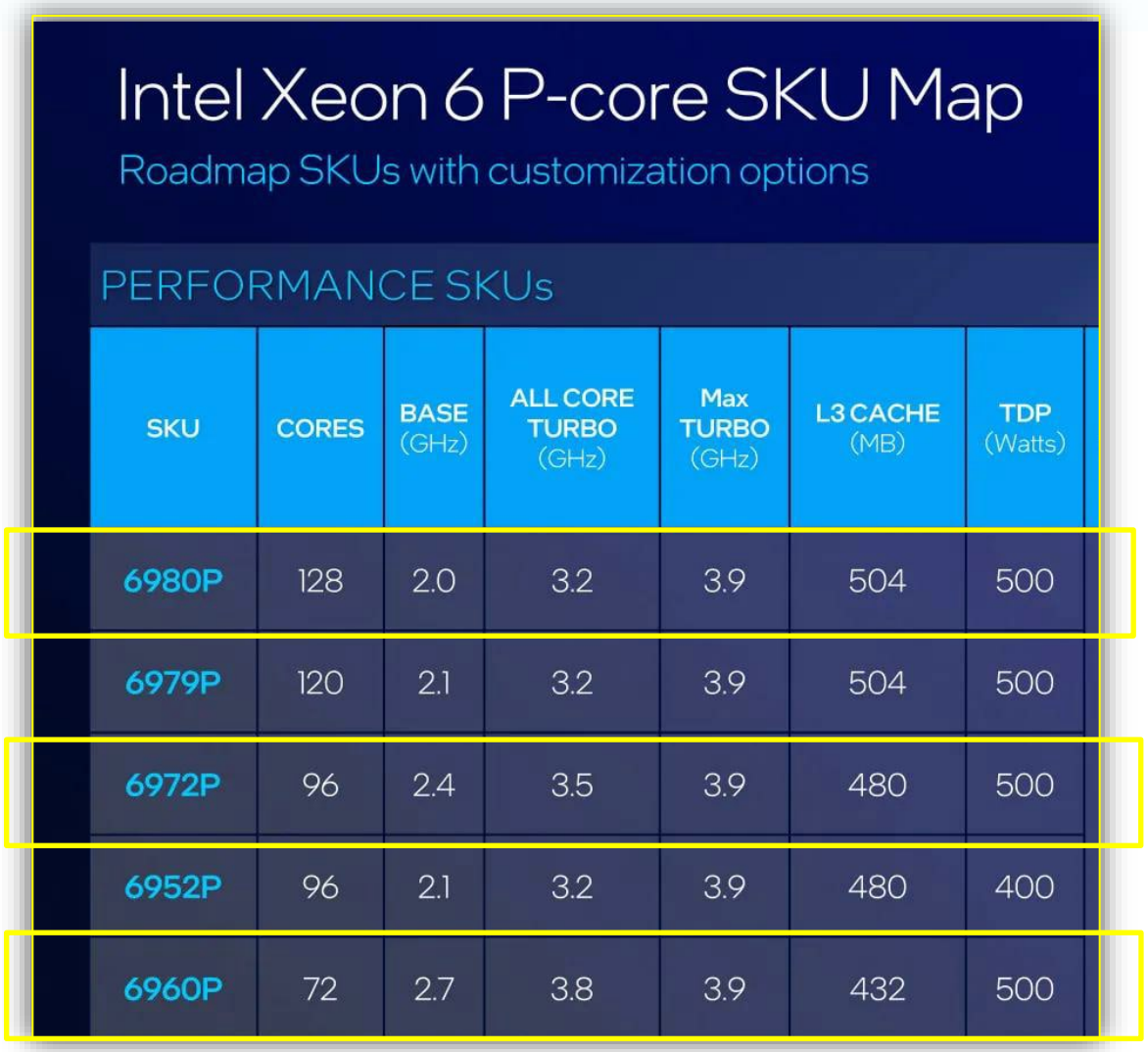
Intel Xeon 6 Granite Rapids Enhancements



- The **Xeon 6900P Granite Rapids (GNR)** processor family features up to 128 performance cores (P-cores), 256 threads, and clock speeds peaking at 3.9 GHz.
- Substantially **larger L3 cache**, up to 504 MB, with 12 memory channels, & support for both **6,400 MT/s DDR5 and 8,800 MT/s Multiplexer Combined Rank DIMMS (MCR-DIMMS),** delivers major improvements in memory B/W.
- Support for **six UPI 2.0 links** and up to **96 lanes of PCI Express 5.0**, featuring CXL 2.0 support

- **Large generational gains** going from prior **Emerald Rapids** (**EMR) processors to GNR**.
- Modest improvement in **instructions-per-clock (IPC)** compared to the Raptor Cove cores in the EMR Xeons, amounting to **< 10 %,** but there are **twice the number of cores!**
- Comes at the expense of **higher power consumption**. Compared to EMR, Intel's 6900P-series GNR processors are consuming an **extra 50-150 watts**.

Intel Xeon 6 Pcore SKU Map



Intel Xeon 6 P-core SKU Map

Roadmap SKUs with customization options

PERFORMANCE SKUs

SKU	CORES	BASE (GHz)	ALL CORE TURBO (GHz)	Max TURBO (GHz)	L3 CACHE (MB)	TDP (Watts)
6980P	128	2.0	3.2	3.9	504	500
6979P	120	2.1	3.2	3.9	504	500
6972P	96	2.4	3.5	3.9	480	500
6952P	96	2.1	3.2	3.9	480	400
6960P	72	2.7	3.8	3.9	432	500

The flagship **Intel Xeon 6980P** features

- **128 cores / 256 threads**, and a **2.0GHz base clock** with 3.2GHz all-core turbo frequency and 3.9GHz maximum turbo frequency.
- **504MB of L3 cache** and a **500 Watt TDP**.
- As with the rest of the Xeon 6 P-core SKUs, there is support for **12 channel DDR5-6400 memory or Multiplexed Rank DIMMS (MR-DIMM) 8800MT/s**

Performance of Computational Chemistry and Ocean Modelling Codes

Systems, Software and Installation

Baseline Cluster System

Supercomputing Wales "Hawk" Cluster Configuration	
"Phase-1" - Intel Skylake Partition	201 nodes, totalling 8,040 cores, 46.080 TB total memory. • CPU: 2 x Intel Xeon Skylake Gold 6148 CPU @ 2.40GHz with 20 cores each; RAM: 192 GB, 384GB on high memory and GPU nodes; GPU: 26 x nVidia P100 GPUs with 16GB of RAM on 13 nodes. • Mellanox IB/EDR infiniband interconnect.
"Phase-2" AMD Rome Partition	64 nodes, totalling 4,096 cores, 32 TB total memory. • CPU: 2 x AMD EPYC Rome 7502 CPU @ 2.50GHz with 32 cores each; RAM: 512 GB, and GPU nodes; GPU: 30 x nVidia V100 GPUs with 16GB of RAM on 15 nodes
Researcher Funded Partitions	• 4,616 cores – Intel Skylake dedicated researcher expansion • 5,288 cores – Intel CSL and AMD Milan SKUs • 2,064 cores – Intel Broadwell and Haswell Raven migrated sub-system nodes (now decommissioned)

The available compute hardware is managed by the **Slurm job scheduler** and organised into 'partitions' of similar type/purpose.

Intel Xeon Emerald Rapids Clusters

Cluster / Configuration
Dell Zenith cluster at the Dell Technologies HPC & AI Innovation Lab – Intel Xeon sub-systems with Mellanox HDR and NDR interconnect fabrics running Slurm
• 50 nodes × Intel **Xeon Platinum 8592 Processor / 1.90 GHz**; # of CPU Cores: **64**; # of Threads: 128; Max Turbo Frequency: 3.90 GHz Base Clock: **1.90 GHz;** Cache **320 MB;** Default TDP / TDP: 350W; **DDR5 4800 MT/s**; **Mellanox NDR 400Gb/s** • The 8592 systems are connected to NDR InfiniBand, configured in a fat tree, with each rack of nodes generally using a single edge switch.
• 20 nodes × Intel **Xeon Platinum 8562Y Processor / 2.80 GHz**; # of CPU Cores: **32**; # of Threads: 64; Max Turbo Frequency: 4.10 GHz Base Clock: **2.80 GHz;** Cache **60 MB;** Default TDP / TDP: 300W; **DDR5 6400 MT/s**; **Mellanox NDR 400Gb/s**
Intel's Endeavour cluster with Mellanox HDR and Cornelis OPE interconnect fabrics running Slurm
• 150 nodes × Intel **Xeon Platinum 8592 Processor / 1.90 GHz**; # of CPU Cores: **64**; # of Threads: 128; Max Turbo Frequency: 3.90 GHz Base Clock: **1.90 GHz;** Cache **320 MB;** Default TDP / TDP: 350W; **DDR5 4800 & 6400 MT/s; Mellanox HDR 200Gb/s; Cornelis OPE**
• 20 nodes × Intel **Xeon Platinum 8568Y Processor / 2.30 GHz**; # of CPU Cores: **48**; # of Threads: 96; Max Turbo Frequency: 4.00 GHz Base Clock: **2.30 GHz;** Cache **300 MB;** Default TDP / TDP: 350W; **DDR5 6400 MT/s; Mellanox HDR 200Gb/s; Cornelis OPE**
Intel's EMR 8592 cluster (4 nodes) with IB backend network fabric running Slurm

Intel Xeon Granite Rapids Clusters

Cluster / Configuration
Intel's Endeavour cluster with Mellanox HDR and Cornelis OPE interconnect fabrics running Slurm
• 150 nodes × Intel **Xeon 5 Platinum 8592 Processor / 1.90 GHz**; # of CPU Cores: **64**; # of Threads: 128; Max Turbo Frequency: 3.90 GHz Base Clock: **1.90 GHz;** Cache **320 MB;** Default TDP / TDP: 350W; **DDR5 4800 & 6400 MT/s; Max # of Memory Channels 6; Max # of UPI Links 4; Intel® UPI Speed 20 GT/s; Mellanox HDR 200Gb/s; Cornelis OPE**
• 20 nodes × Intel **Xeon 6 6980P Processor / 2.00 GHz**; # of CPU Cores: **128**; # of Threads: 256; Max Turbo Frequency: 3.90 GHz Base Clock: **2.00 GHz;** Cache **504 MB;** Default TDP / TDP: 500W; **DDR5 6400 MT/s & MRDIMM (8800MHz); Max # of Memory Channels 8; Max # of UPI Links 6; Intel® UPI Speed 24 GT/s; Mellanox HDR 200Gb/s; Cornelis OPE**
• 20 nodes × Intel **Xeon 6 6972P Processor / 2.40 GHz**; # of CPU Cores: **96**; # of Threads: 192; Max Turbo Frequency: 3.90 GHz Base Clock: **2.40 GHz;** Cache **480 MB;** Default TDP / TDP: 500W; **DDR5 6400 MT/s & MRDIMM (8800MHz); Max # of Memory Channels 12; Max # of UPI Links 6; Intel® UPI Speed 24 GT/s; Mellanox HDR 200Gb/s; Cornelis OPE**
• 20 nodes × Intel **Xeon 6 6960P Processor / 2.70 GHz**; # of CPU Cores: **72**; # of Threads: 144; Max Turbo Frequency: 4.00 GHz Base Clock: **2.30 GHz;** Cache **480 MB;** Default TDP / TDP: 350W; **DDR5 6400 MT/s & MRDIMM (8800MHz); Max # of Memory Channels 12; Max # of UPI Links 6; Intel® UPI Speed 24 GT/s; Mellanox HDR 200Gb/s; Cornelis OPE**

AMD EPYC Genoa Clusters

Cluster / Configuration
Dell "Minerva" cluster – Dell Technologies HPC & AI Innovation Lab – AMD Genoa sub-system with Mellanox NDR interconnect fabric [Slurm]. systems are connected to NDR InfiniBand configured on a single switch
• **22 nodes × AMD EPYC Genoa 9354 / 3.25 GHz**; # of CPU Cores: **32**; # of Threads: 64; Max Turbo Frequency: 3.8 GHz Base Clock: **3.25 GHz;** L3 Cache 256 MB**;** Default TDP / TDP: 280W; **Mellanox NDR 400Gb/s**
• **45 nodes × AMD EPYC Genoa 9654 / 2.4 GHz**; # of CPU Cores: **96**; # of Threads: 192; Max Turbo Frequency: 3.7 GHz Base Clock: **2.4 GHz;** L3 Cache 256 MB**;** Default TDP / TDP: 280W; **Mellanox NDR 400Gb/s**.
• **8 nodes × AMD EPYC Genoa 9684X / 2.55 GHz**; # of CPU Cores: **96**; # of Threads: 192; Max Turbo Frequency: 3.7 GHz Base Clock: **2.55 GHz;** L3 Cache 1152 MB**;** Default TDP / TDP: 400W; **Mellanox NDR 400Gb/s**
AMD 7008-core Genoa cluster, "Ada", at Nottingham University with Mellanox NDR interconnect fabric running Slurm.
• **AMD EPYC Genoa 9454 / 2.75 GHz Processor**; # of CPU Cores: **48**; # of Threads: 96; Max Turbo Frequency: 3.80 GHz Base Clock: **2.75 GHz;** L3 Cache 256 MB**;** Default TDP / TDP: 290W; **Mellanox NDR200 200Gb/s.**
AMD "muncl02" cluster – AMD HPC Center of Excellence in Munich – variety of AMD Genoa SKUs – 9354, 9454, 9554, 9654 and 9684X

NVIDIA HPC-X

NVIDIA HPC-X: Increased use of NVIDIA HPC-X that includes MPI, SHMEM and PGAS communications libraries, and various acceleration packages.

- **Key Features**
  - Offloads collectives communications from MPI onto NVIDIA InfiniBand networking hardware
  - Multiple transport support, including Reliable Connection (RC), Dynamic Connected (DC), and Unreliable Datagram (UD)
  - Intra-node shared memory communication
  - Native support for MPI-3
  - Multi-rail support with message striping
  - NVIDIA GPUDirect with CUDA support
  - NCCL-RDMA-SHARP plug-in support
- Experience suggests that this toolkit enables MPI & SHMEM/PGAS programming languages to achieve **higher performance, scalability, and efficiency.**
- Notable performance impact in both **CASTEP and VASP**. ***(Rev 2.16)***

The Performance Benchmarks

- The ***Test suite*** comprises both **synthetics & end-user applications**. Synthetics limited to **IMB** (*http://software.intel.com/en-us/articles/intel-mpi-benchmarks*), **IOR** (https://github.com/hpc/ior) and **STREAM**
- Variety of “open source” & commercial end-user application codes:

DL_POLY, AMBER & GROMACS (Molecular Dynamics)

GAMESS-UK (molecular electronic structure)

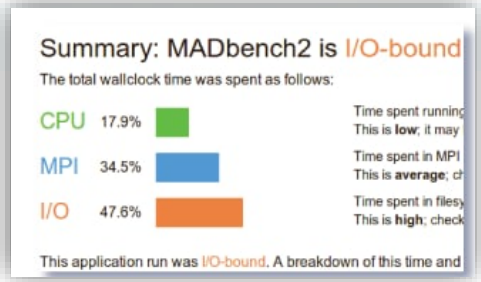
VASP (ab initio Materials properties)

NEMO (ocean modelling code)

- These stress various aspects of the architectures under consideration and should provide a level of insight into why particular levels of performance are observed e.g., ***memory bandwidth and latency, node floating point performance and interconnect performance (both latency and B/W) and sustained I/O performance***.

Analysis Software - Allinea|ARM|Linaro Performance Reports

Provides a mechanism to characterize and understand the performance of HPC application runs through a single-page HTML report.



- Based on Allinea MAP's adaptive sampling technology that keeps data volumes collected and **application overhead low**.
- Modest application slowdown (ca. 5%) even with 1000's of MPI processes.
- **Runs on existing codes: a single command added to execution scripts.**
- If submitted through a batch queuing system, then the submission script is modified to load the Allinea module and add the 'perf-report' command in front of the required mpirun command.

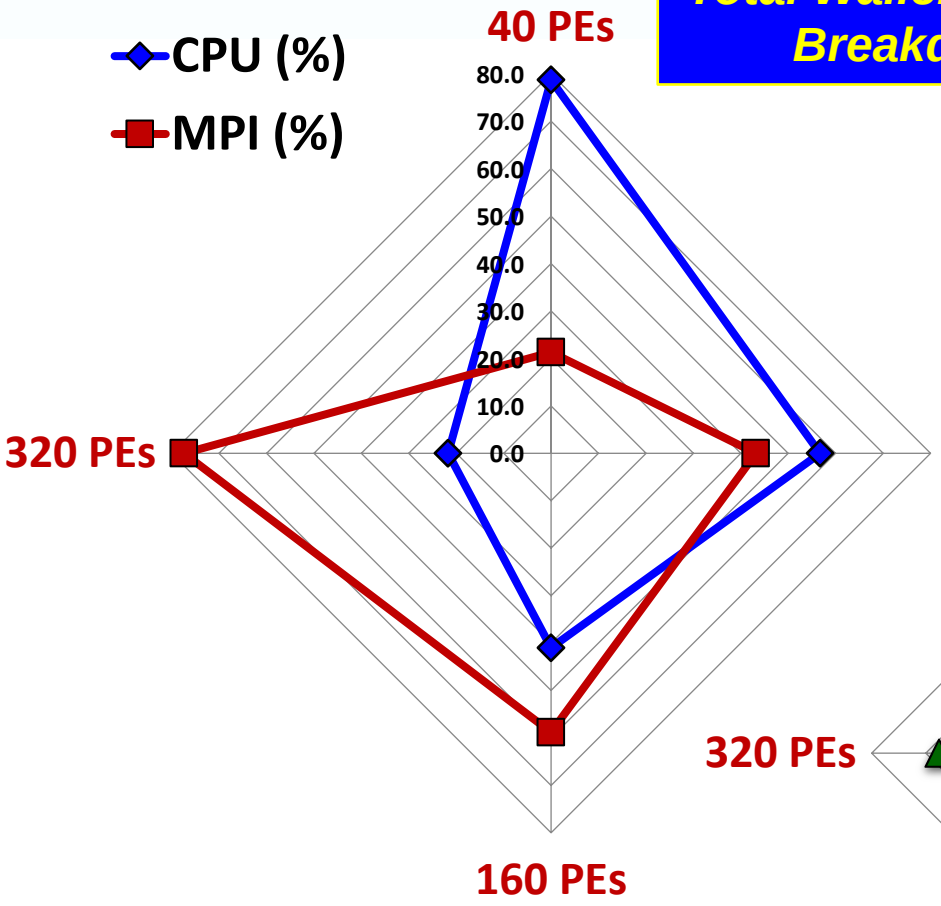
  perf-report mpirun $code

- ***A Report Summary:*** This characterizes how the application's wallclock time was spent, broken down into CPU, MPI and I/O
- All examples from the **Hawk Cluster (SKL Gold 6148 / 2.4GHz)**

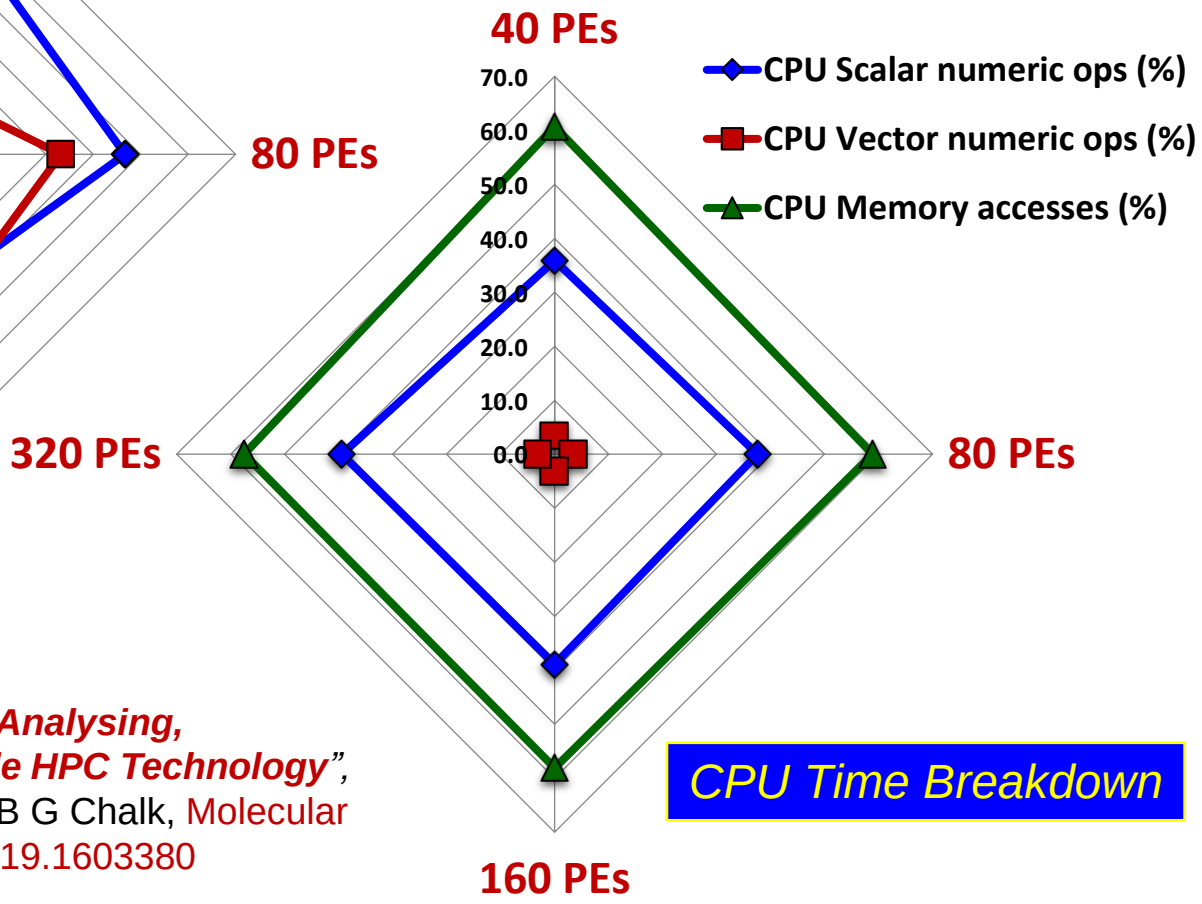
DLPOLY4 – Performance Report: Gramicidin Simulation

Total Wallclock Time Breakdown

Performance Data (40-320 PEs)



Smooth Particle Mesh Ewald Scheme



"DL_POLY - A Performance Overview. Analysing, Understanding and Exploiting available HPC Technology", Martyn F Guest, Alin M Elena and Aidan B G Chalk, Molecular Simulation, (2019) 10.1080/08927022.2019.1603380

CPU Time Breakdown

EPYC - Compiler and Run-time Options

STREAM (AMD Minerva Cluster):

```
icc stream.c -DSTATIC -Ofast -march=core-avx2 -DSTREAM_ARRAY_SIZE=2500000000 -
DNTIMES=10 -mcmodel=large -shared-intel -restrict -qopt-streaming-stores always
-o streamc.Rome
icc stream.c -DSTATIC -Ofast -march=core-avx2 -qopenmp -
DSTREAM_ARRAY_SIZE=2500000000 -DNTIMES=10 -mcmodel=large -shared-intel -restrict
-qopt-streaming-stores always -o streamcp.Rome

# When using IntelMPI on AMD Rome/Milan
export I_MPI_FABRICS=shm:ofi
export I_MPI_SHM=clx_avx2
export FI_PROVIDER=mlx
```

Compilation:

INTEL SKL: –O3 –xCORE-AVX512

AMD EPYC: –O3 –march=core-avx2 -align array64byte -fma -ftz -fomit-frame-pointer

STREAM (AMD Muncl02 Cluster):

```
Build with AOCC and non-temporal store enabled
CPPFLAGS="-DSTREAM_ARRAY_SIZE=2500000000 -DNTIMES=10 -DSTREAM_TYPE=double -
DSTATIC"
CFLAGS="-march=znver4 -ffp-contract=fast -O3 -m64 -Ofast -ffast-math -fPIC -
mcmodel=large –fopenmp"
clang ${CFLAGS} ${CPPFLAGS} -fno-unroll-loops -fnt-store=aggressive -c stream.c
-o stream_55GB.o
clang ${CFLAGS} stream_55GB.o -o /bin/stream_55GB
```

2P EPYC 9654 node with 1 threads / CCD i.e., 2x12 scattered threads

```
OMP_NUM_THREADS=24OMP_PLACES='{0},{8},{16},{24},{32},{40},{48},{56},{64},{72},{8
0},{88},{96},{104},{112},{120},{128},{136},{144},{152},{160},{168},{176},{184}'
```

Using the Spack package manager

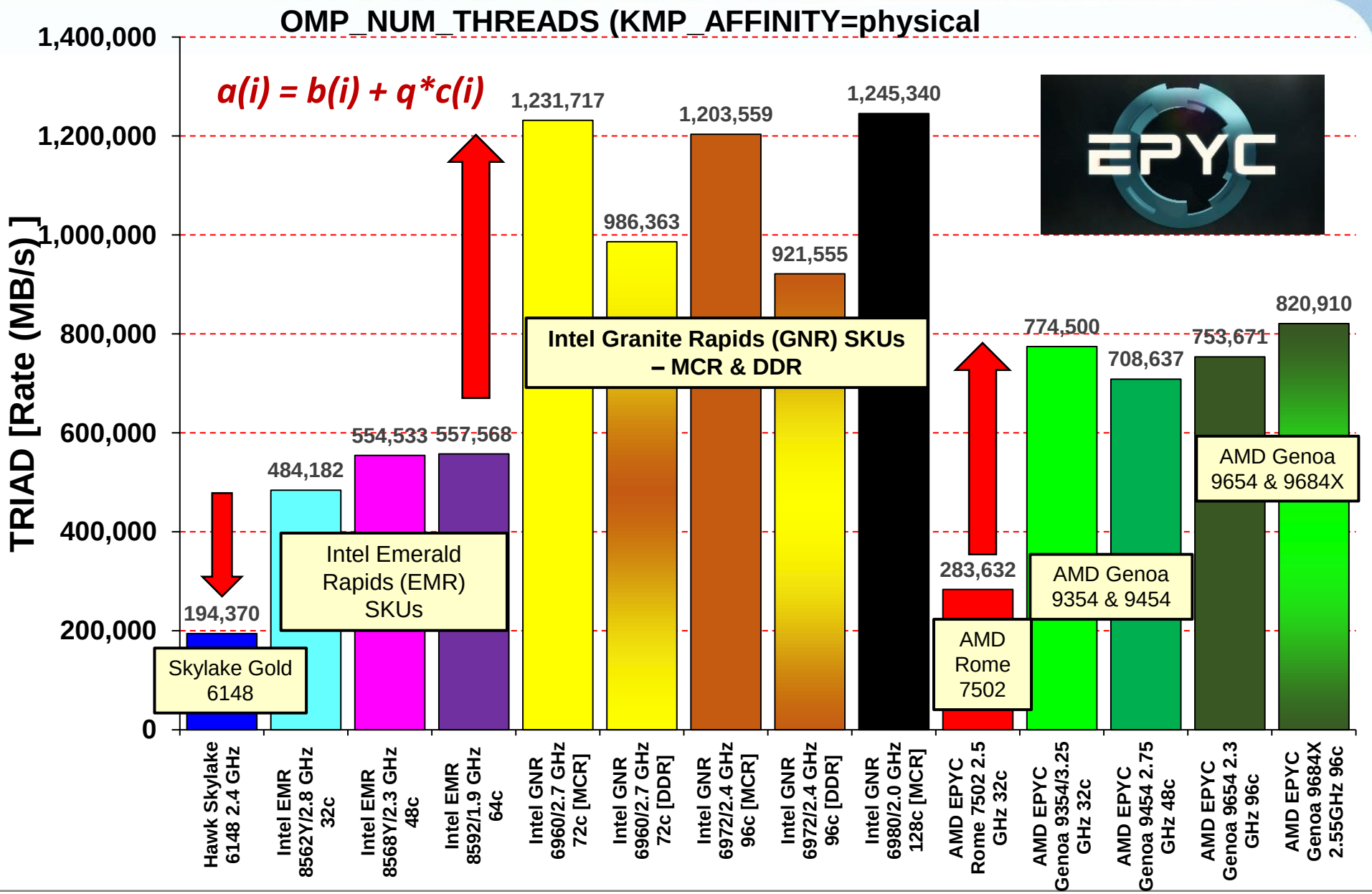
- Like EasyBuild (1), Spack (2) Spack is a multi-platform package manager that builds and installs multiple versions and configurations of software. **Spack** resolves dependencies and installs them like any other package manager you can find on a linux platform.
- The definition provided by the official documentation is as follows:

  "Spack is a multi-platform package manager that builds and installs multiple versions and configurations of software. It works on Linux, macOS, and many supercomputers. Spack is non-destructive: installing a new version of a package does not break existing installations, so many configurations of the same package can coexist"
- *Spack offers a simple "spec" syntax that allows users to specify versions and configuration options. Package files are written in pure Python, and specs allow package authors to write a single script for many different builds of the same package. With Spack, you can build your software as you wish".*

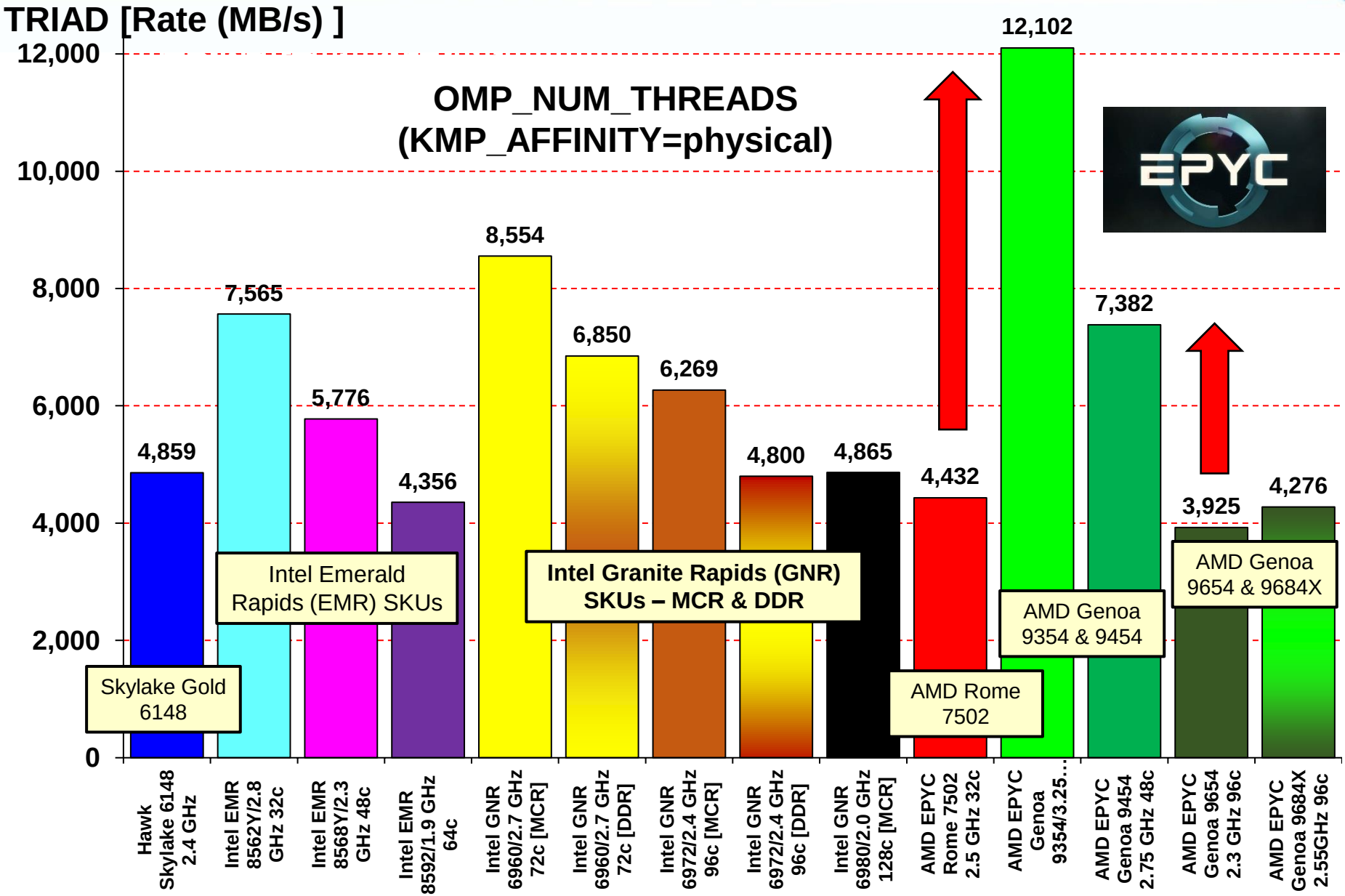
[1] *https://docs.easybuild.io/installation/*

[2] *https://spack.readthedocs.io/en/latest/index.html#*

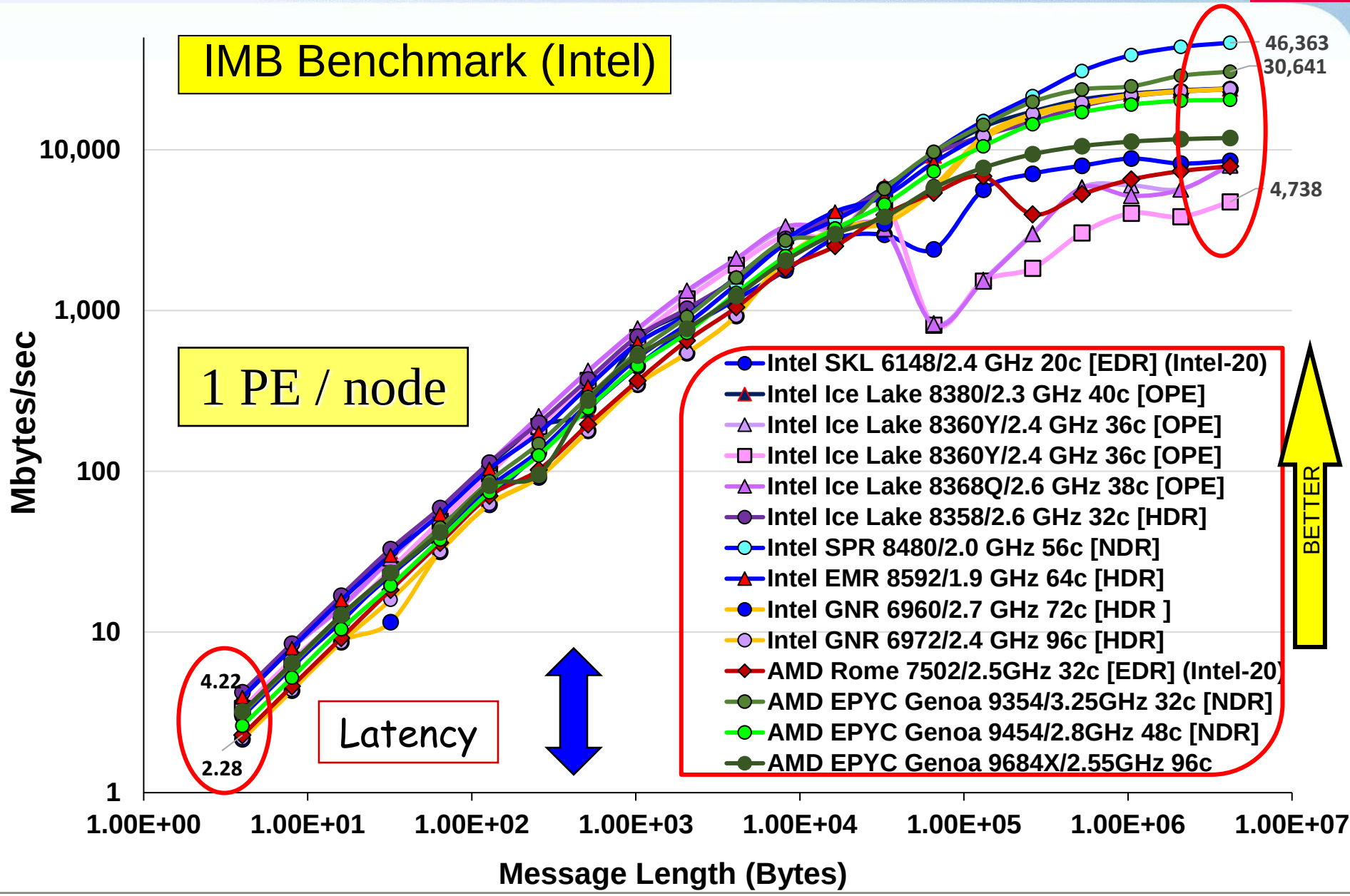
Memory B/W – STREAM performance



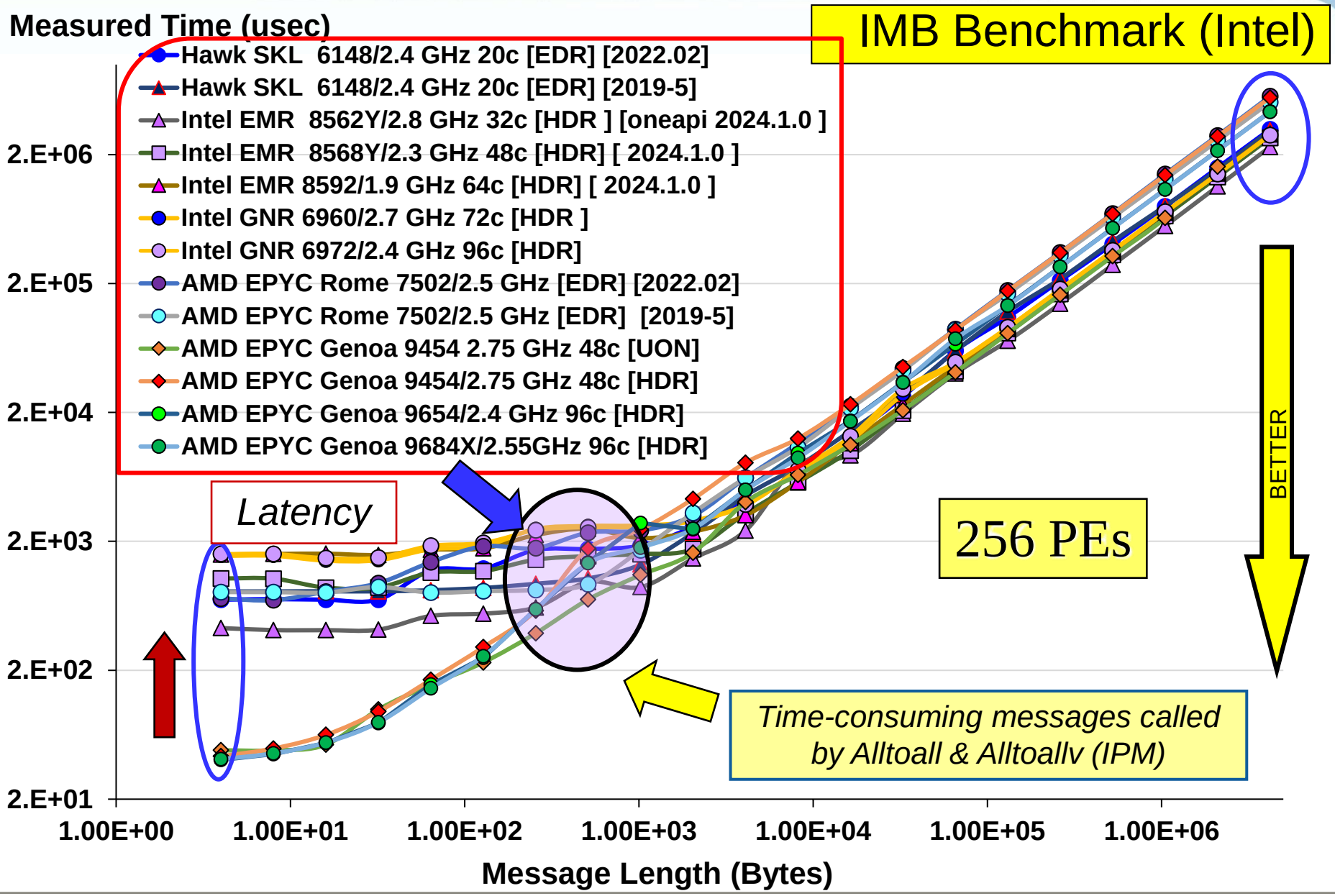
Memory B/W – STREAM / core performance



MPI Performance – PingPong



MPI Collectives – Alltoallv (256 PEs)

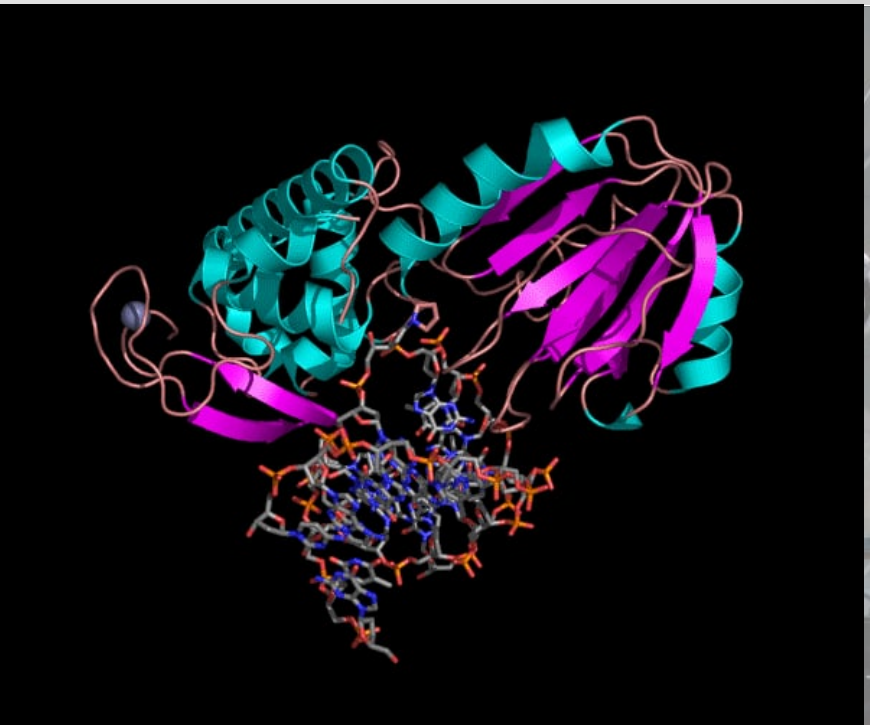


Performance Metrics – “Core to Core” & “Node to Node”

- Analysis of performance Metrics across a variety of data sets
  - ❑ “**Core to core**” and “**node to node**” workload comparisons
    - ***Core to core*** comparison i.e. performance for jobs with a fixed number of cores
    - ***Node to Node*** comparison typical of the performance when running a workload (real life production). Expected to reveal the major benefits of **increasing core count per socket**
  - ❑ Focus on a variety of “**node to node**” and “**core-to-core**” comparisons e.g., :

1	***Hawk - Dell |EMC Skylake Gold 6148 2.4GHz (T) EDR with 40 cores / node***	***AMD EPYC Genoa 9654 nodes with 192 cores per node. [1-8 nodes]***
2	***Hawk - Dell |EMC Skylake Gold 6148 2.4GHz (T) EDR with 40 cores / node***	***Intel Xeon Granite Rapids 6972 nodes with 192 cores per node. [1-8 nodes]***

Performance of Computational Chemistry and Ocean Modelling Codes

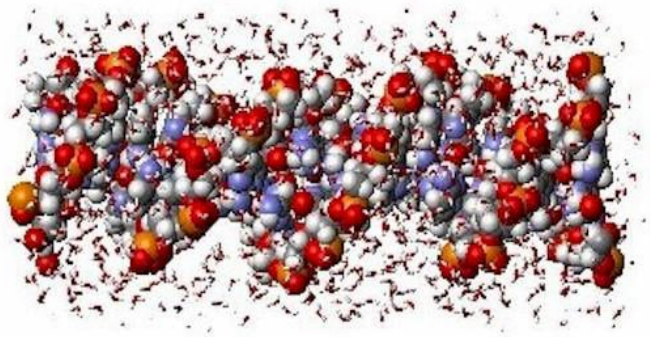


Molecular Simulation;

1. DL_POLY

Molecular Simulation I. DL_POLY

Molecular Dynamics Codes:
***AMBER**, **DL_POLY**, CHARMM,*
*NAMD, LAMMPS, **GROMACS** etc*



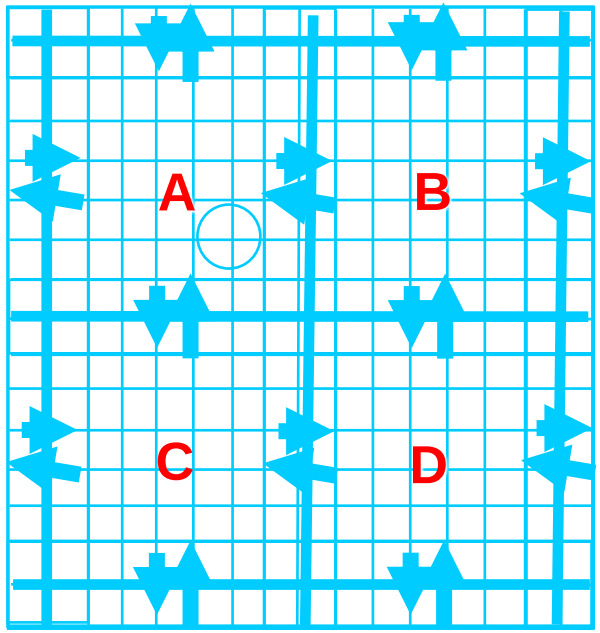
DL_POLY

- Developed as CCP5 parallel MD code by W. Smith, T.R. Forester and I. Todorov
  - UK CCP5 + International user community
  - DLPOLY_classic (replicated data) and DLPOLY_3 & _4 (distributed data – domain decomposition)
- Areas of application:
  - liquids, solutions, spectroscopy, ionic solids, molecular crystals, polymers, glasses, membranes, proteins, metals, solid and liquid interfaces, catalysis, clathrates, liquid crystals, biopolymers, polymer electrolytes.

DL_POLY 4 – Distributed data

Domain Decomposition - Distributed data:

- Distribute atoms, forces across the nodes
  - More memory efficient, can address much larger cases (10^5-10^7)
- Shake and short-ranges forces require only neighbour communication
  - communications scale linearly with number of nodes
- Coulombic energy remains global
  - Adopt **Smooth Particle Mesh Ewald** scheme
    - includes Fourier transform smoothed charge density (reciprocal space grid typically 64x64x64 - 128x128x128)



W. Smith and I. Todorov

Benchmarks

1. NaCl Simulation; 216,000 ions, 1,000 time steps, Cutoff=12Å

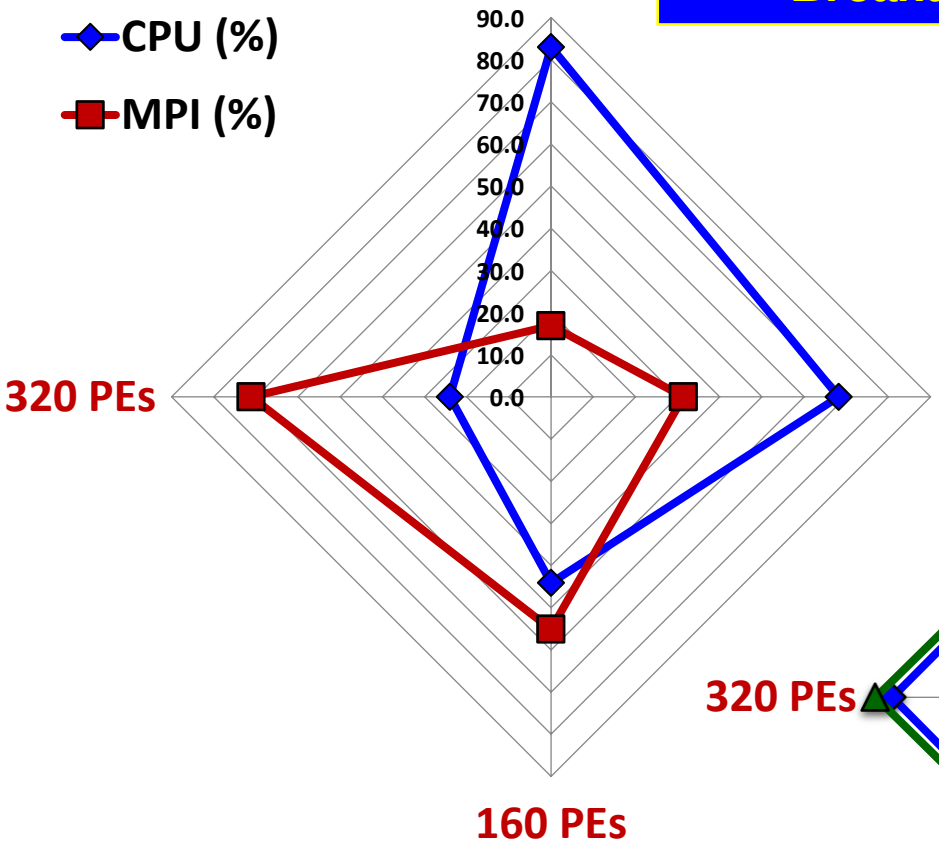
2. Gramicidin in water; rigid bonds + SHAKE: 792,960 ions, 1,000 time steps

https://www.scd.stfc.ac.uk/Pages/DL_POLY.aspx

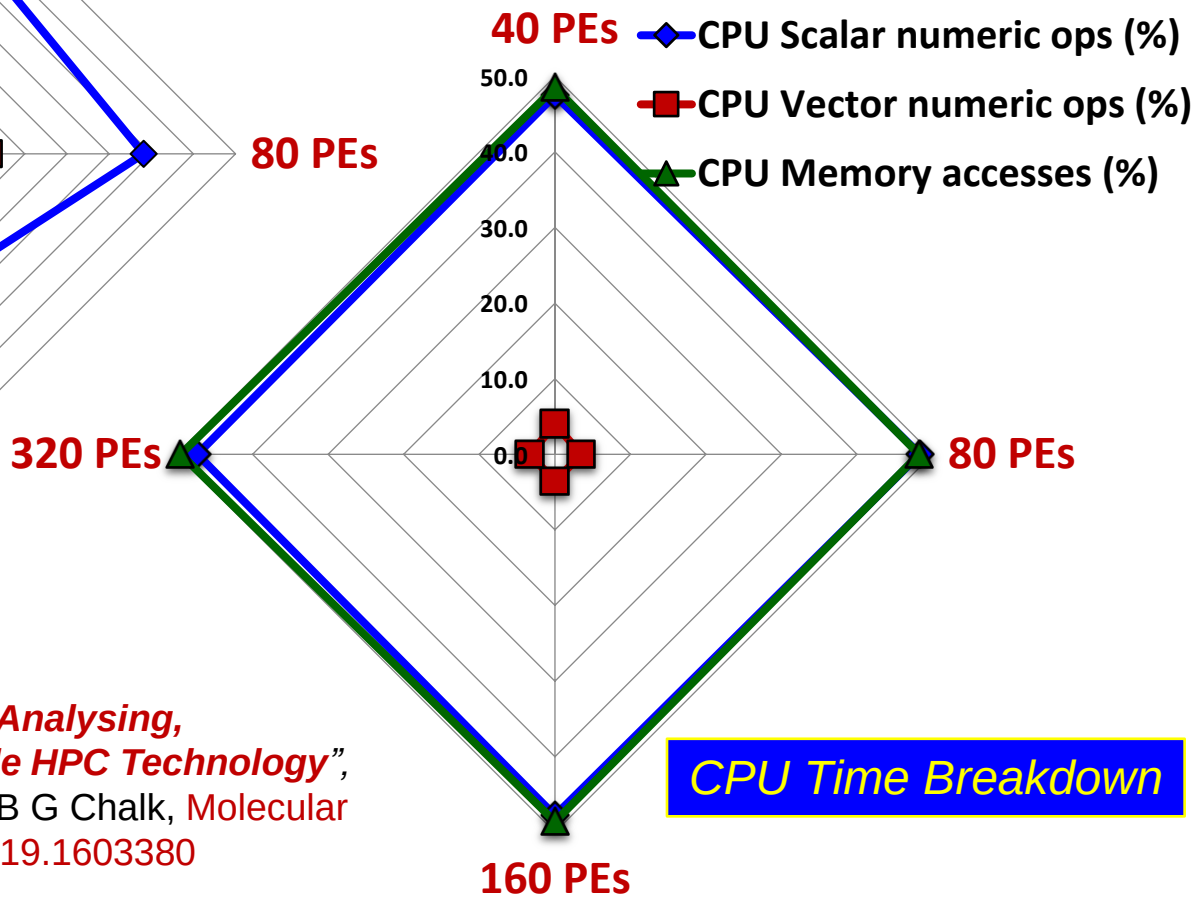
DLPOLY4 – Performance Report: NaCl Simulation

Total Wallclock Time Breakdown

Performance Data (40-320 PEs)



Smooth Particle Mesh Ewald Scheme



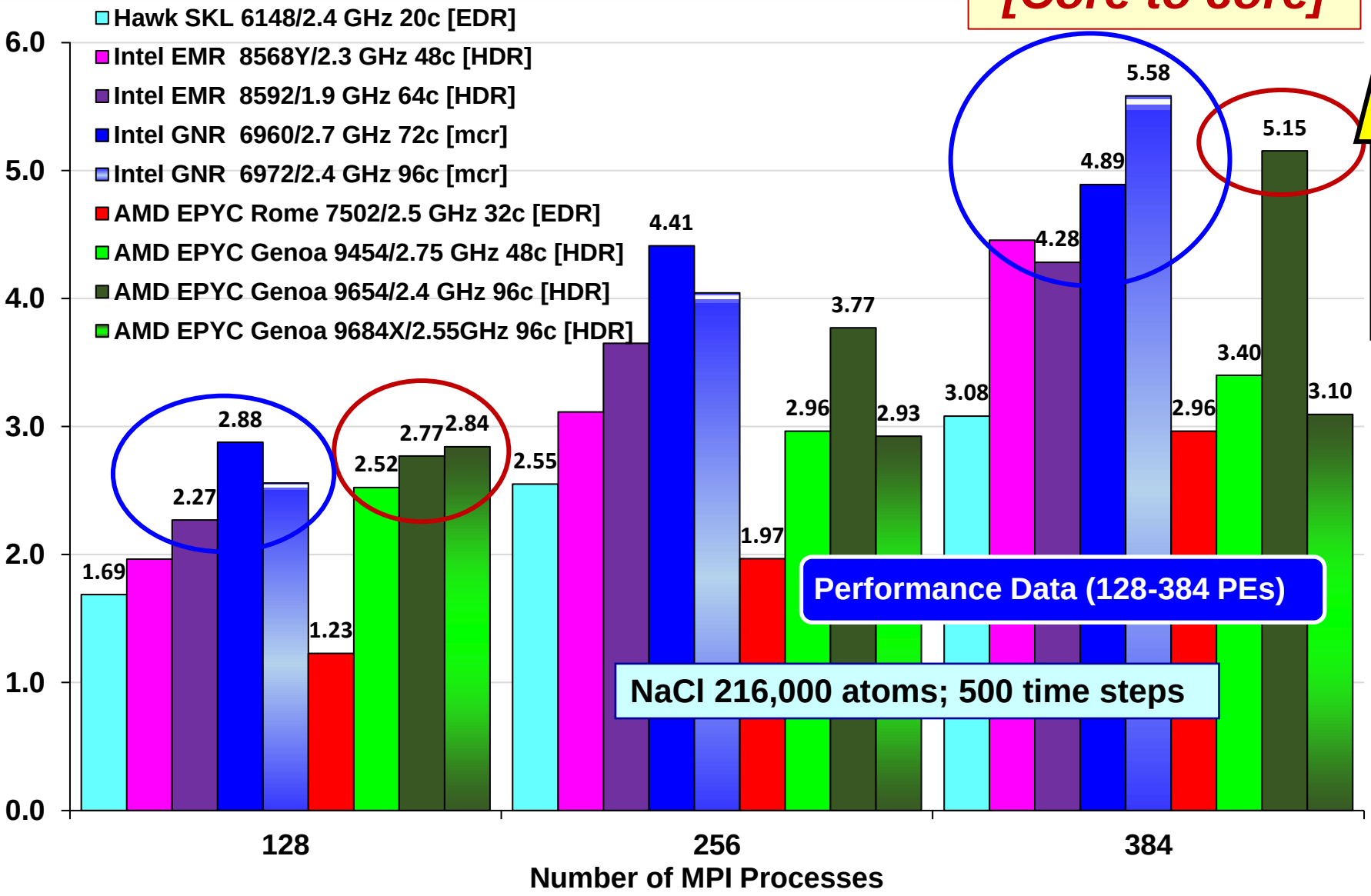
"DL_POLY - A Performance Overview. Analysing, Understanding and Exploiting available HPC Technology", Martyn F Guest, Alin M Elena and Aidan B G Chalk, Molecular Simulation, (2019) 10.1080/08927022.2019.1603380

CPU Time Breakdown

DL_POLY 4 – NaCl Simulation

Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*

[Core to core]

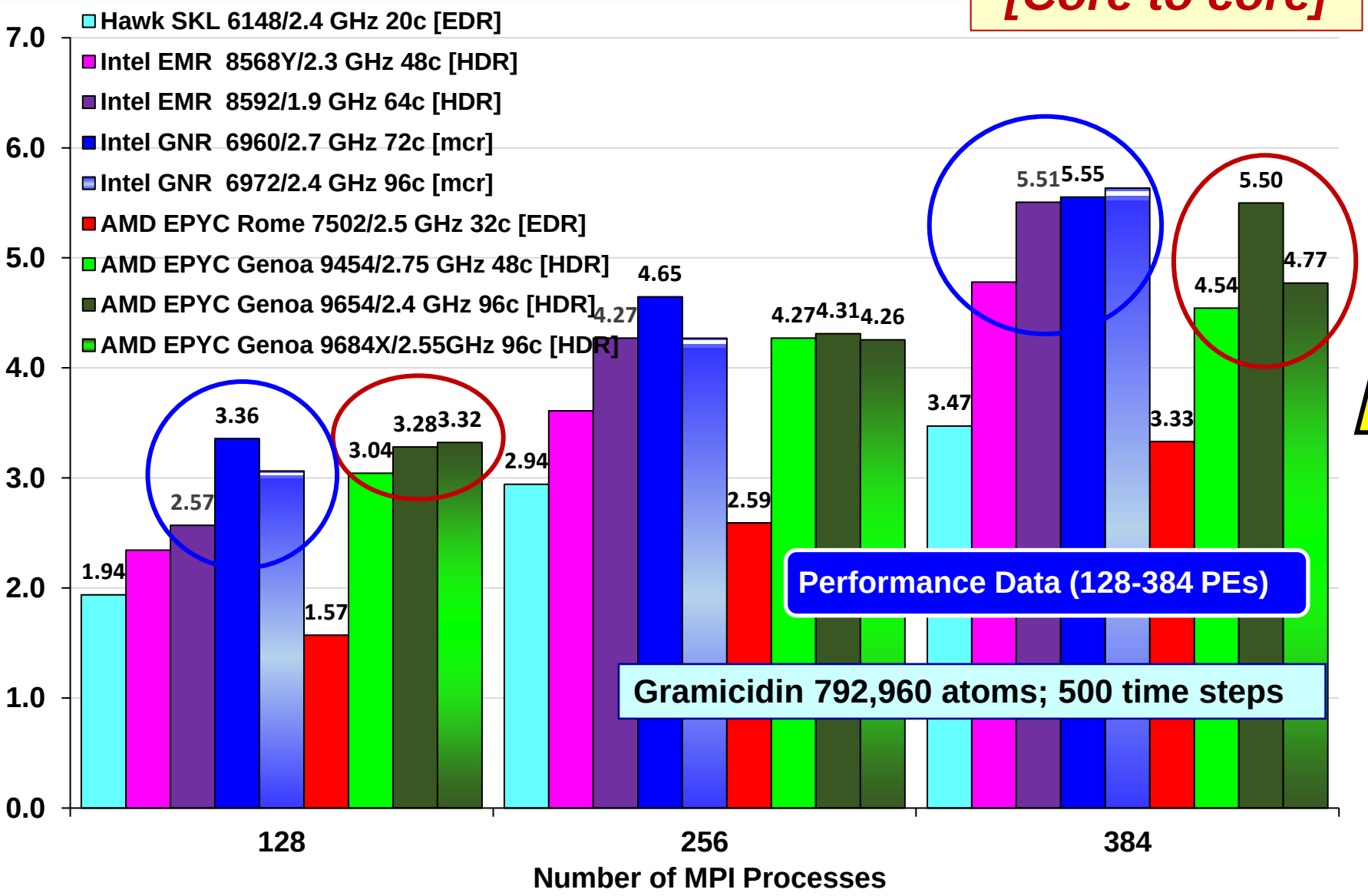


DL_POLY 4 – Gramicidin Simulation

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*

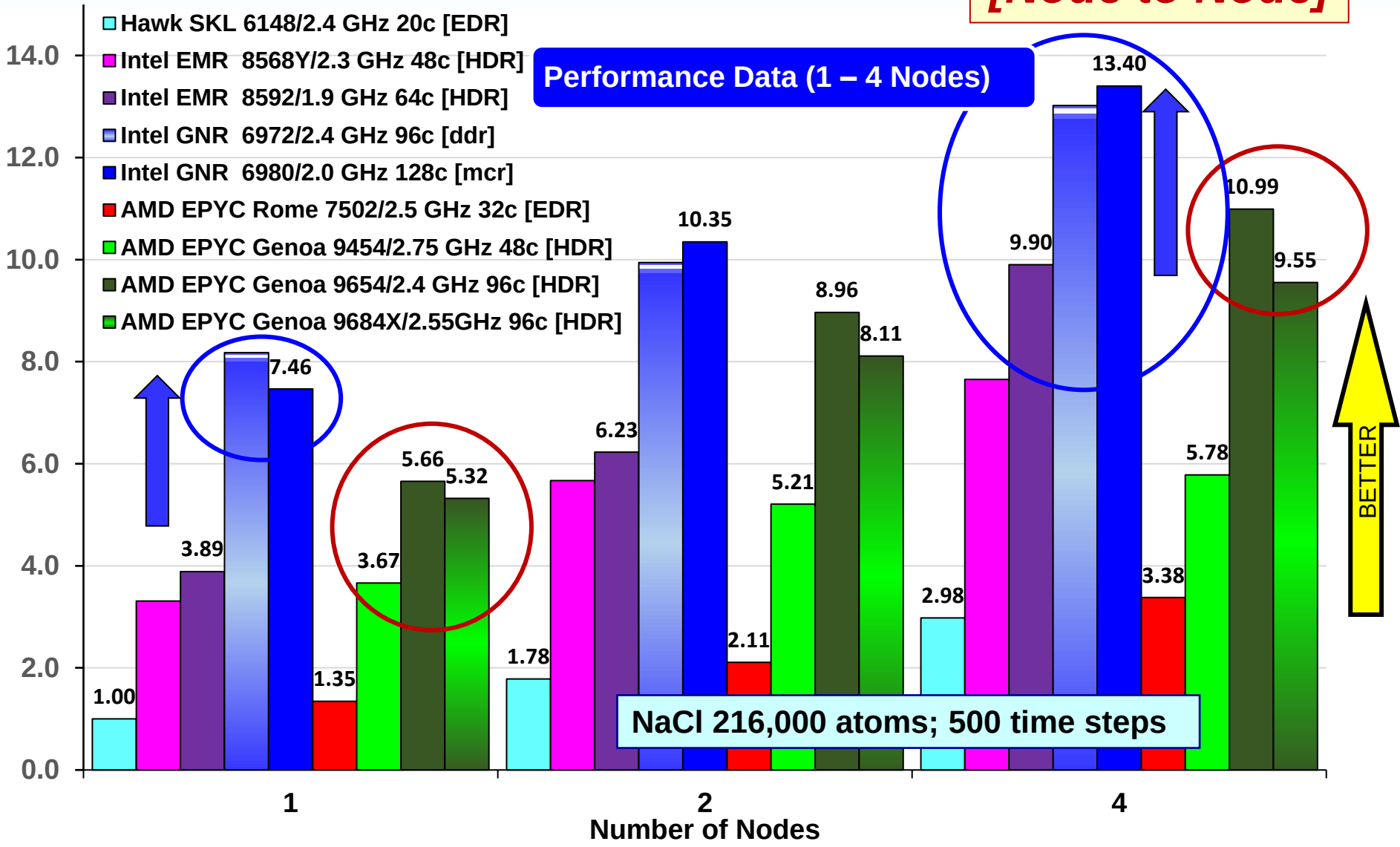
[Core to core]



DL_POLY 4 – NaCl Simulation

Performance

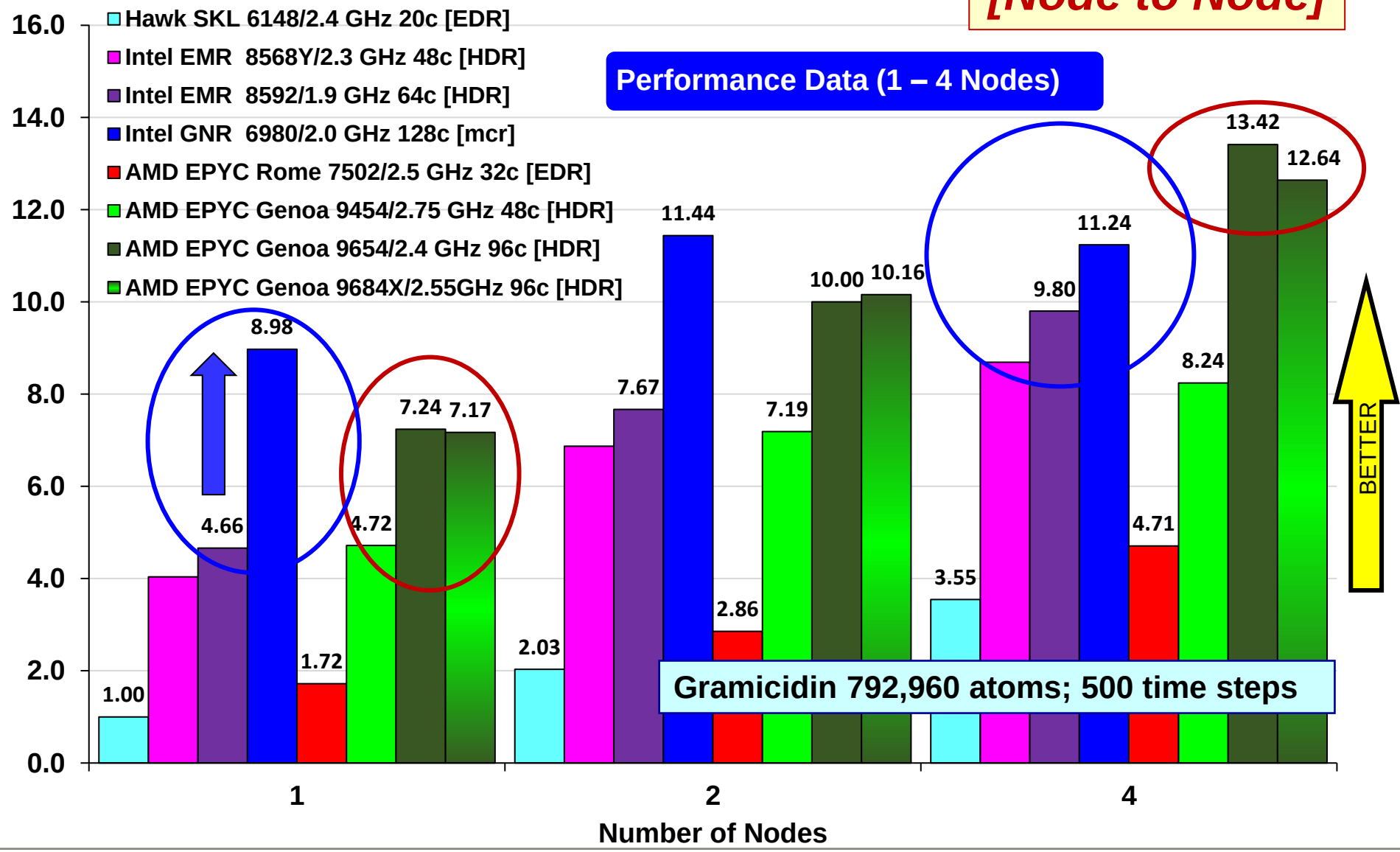
[Node to Node]



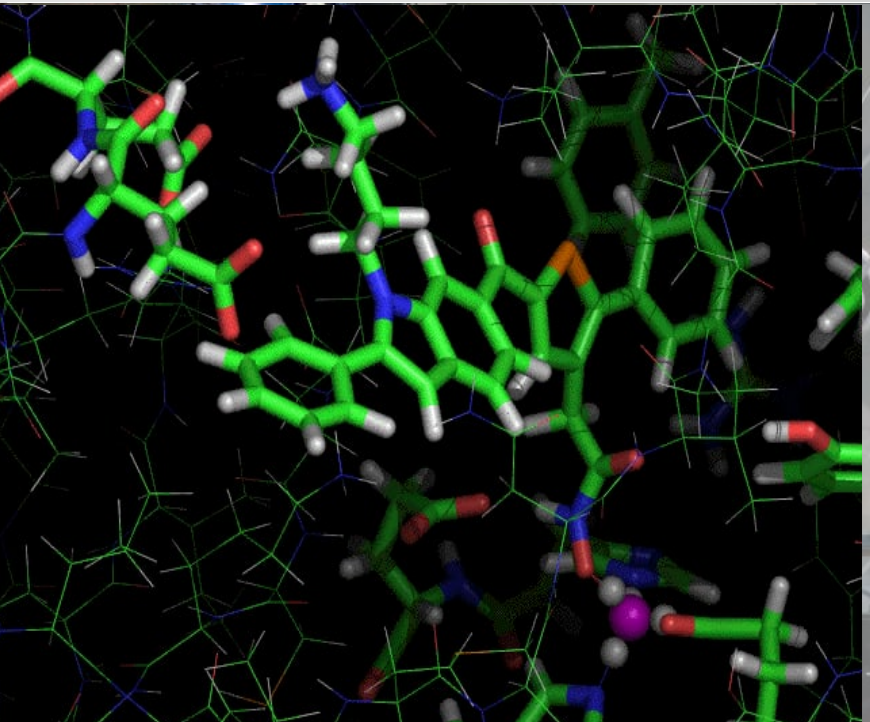
DL_POLY 4 – Gramicidin Simulation

Performance *Relative to the Hawk SKL 6148 2.4 GHz (1 Node)*

[Node to Node]

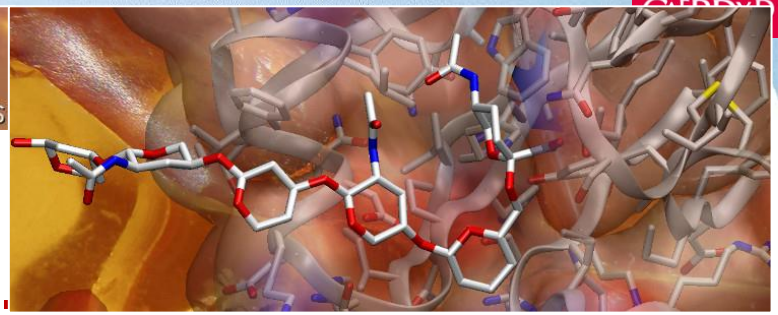


Performance of Computational Chemistry and Ocean Modelling Codes



Molecular Simulation: 2. AMBER

Molecular Simulation II. AMBER



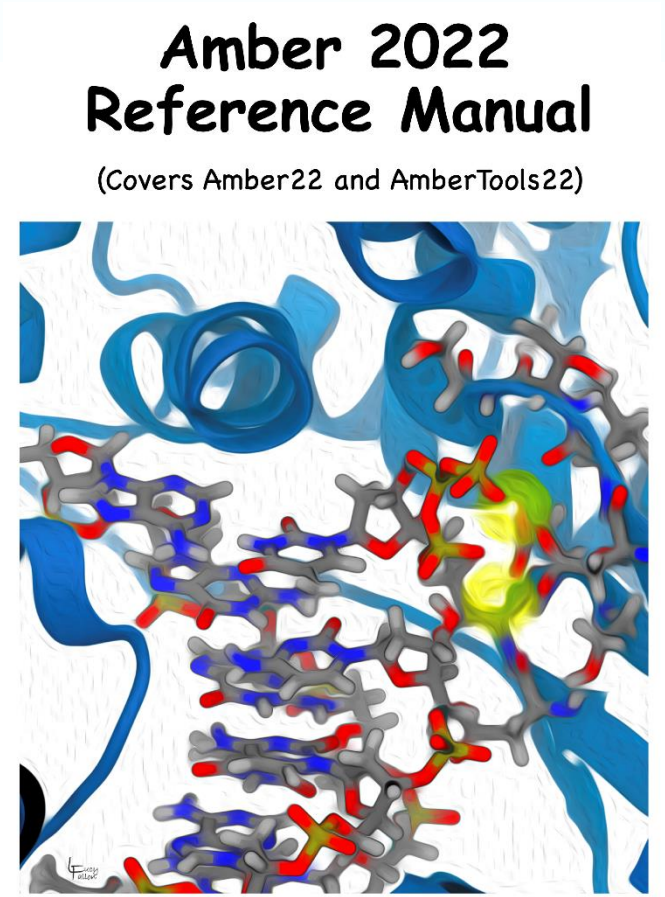
- AMBER18 and AMBER22 used:
  PMEMD & GPU accelerated PMEMD.
- **M01 Benchmark**
  - Major Urinary Protein (MUP) + IBM ligand (21,736 atoms)
- **M06 Benchmark**
  - Cluster of six MUPs (134,013 atoms)
- **M27 Benchmark**
  - **Cluster of 27 MUPs (657,585 atoms)**
- **M45 Benchmark**
  - **Cluster of 45 MUPs (932,751 atoms)**

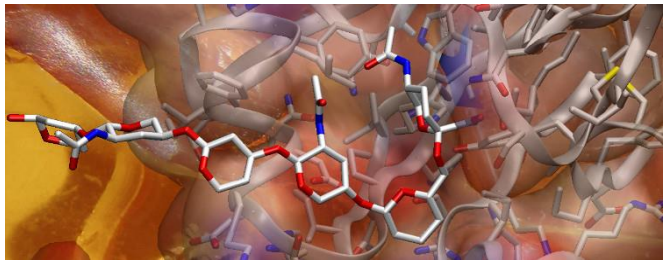
*All test cases run 30,000 steps * 2fs = 60ps simulation time. Periodic boundary conditions, constant pressure, T=300K. Position data written every 500 steps.*

R. Salomon-Ferrer, D.A. Case, R.C. Walker. An overview of the Amber biomolecular simulation package. WIREs Comput. Mol. Sci. 3, 198-210 (2013).

D.A. Case, T.E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. Woods. The Amber biomolecular simulation programs. J. Computat. Chem. 26, 1668-1688 (2005).

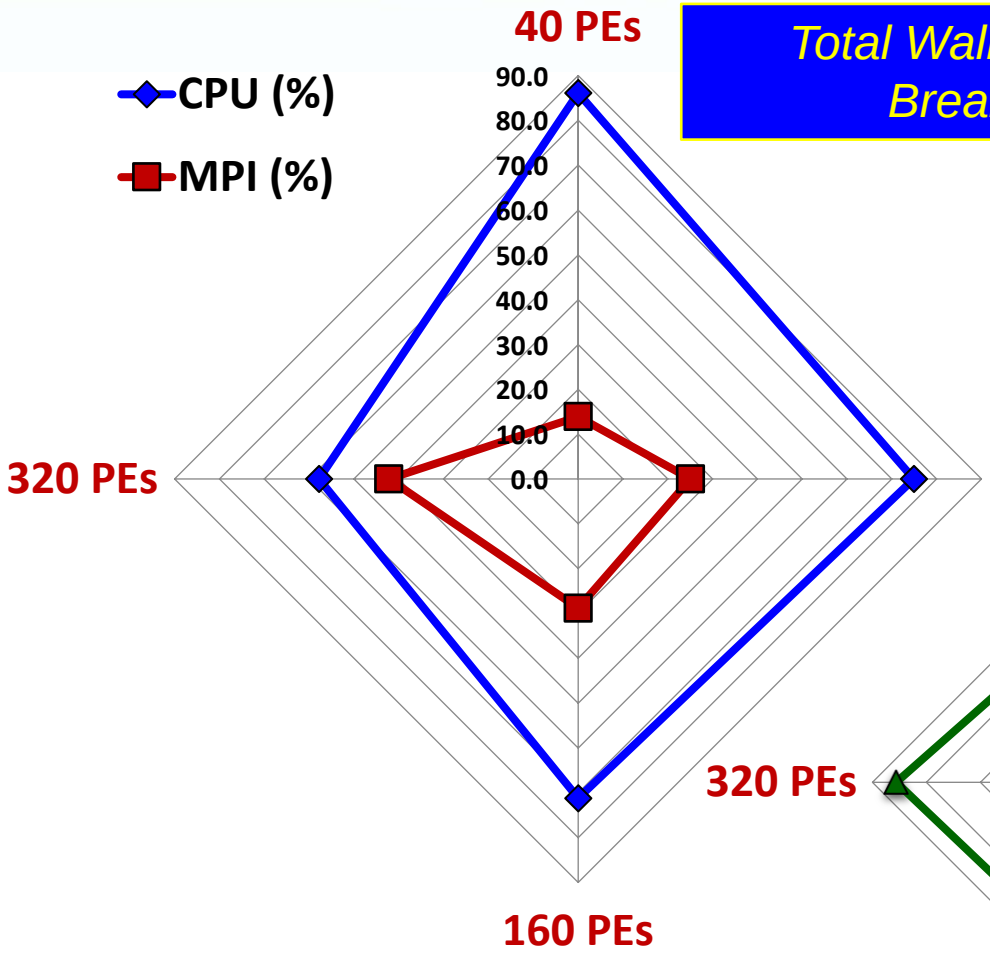
Molecular Simulation II. AMBER





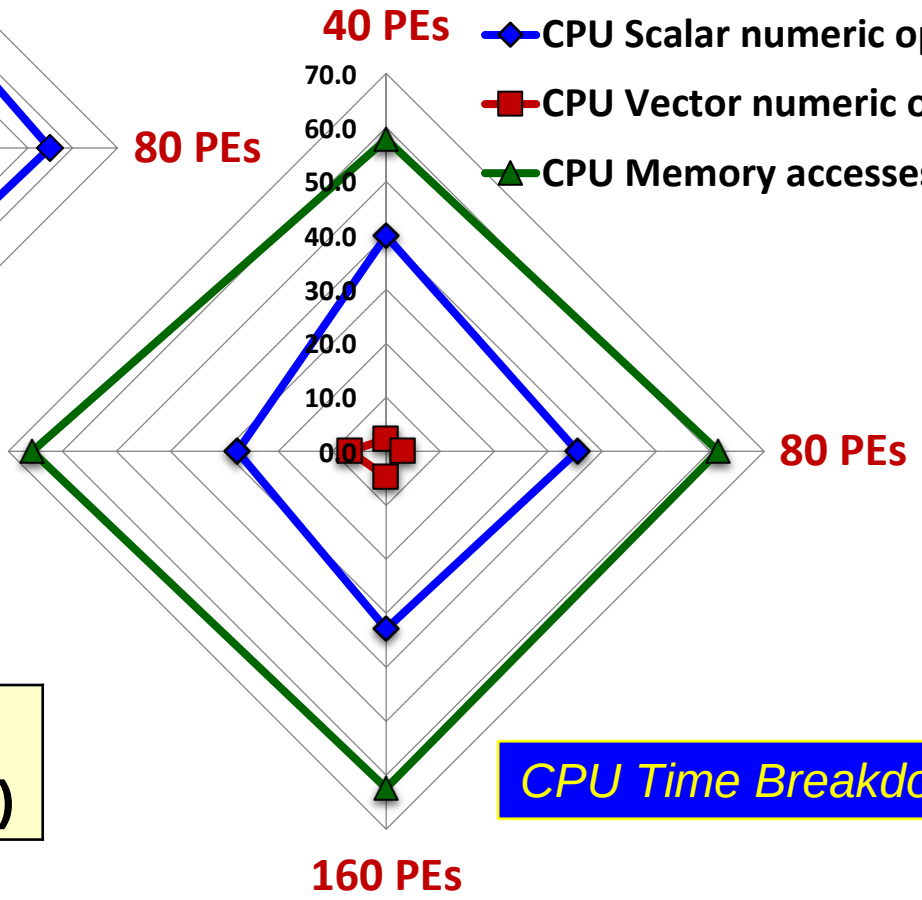
- **AMBER22** released (on April 27, 2022).
- The Amber22 package builds on AmberTools22 by adding the pmemd program, which resembles the sander (MD) code in AmberTools, but provides better performance on multiple CPUs, and dramatic speed improvements on GPUs.
- **AMBER18** (released in 2018) also used in this study. In practice we find **also identical performance of the two code releases** when running the M01, M06, M27 and M45 performance test cases.
- Presentation limited to the **M27 and M45** test cases for M01 and M06 are now too small for meaningful analysis

AMBER22 – M45 Benchmark Performance Report



Total Wallclock Time Breakdown

Performance Data (40-320 PEs)

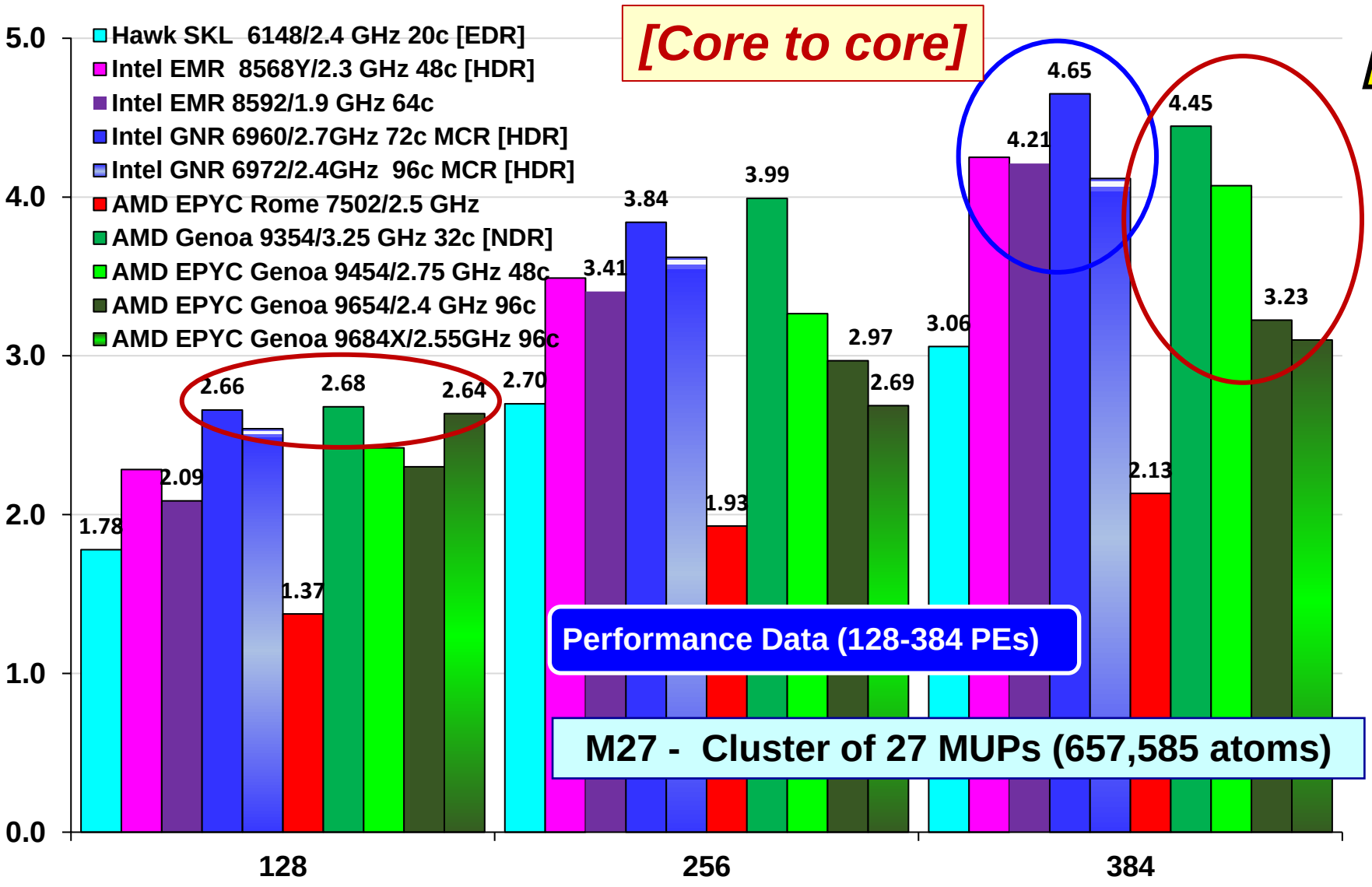


CPU Time Breakdown

Cluster of 45 Major Urinary Proteins (MUPs, 932,751 atoms)

AMBER22 - M27 Performance Analysis

Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*

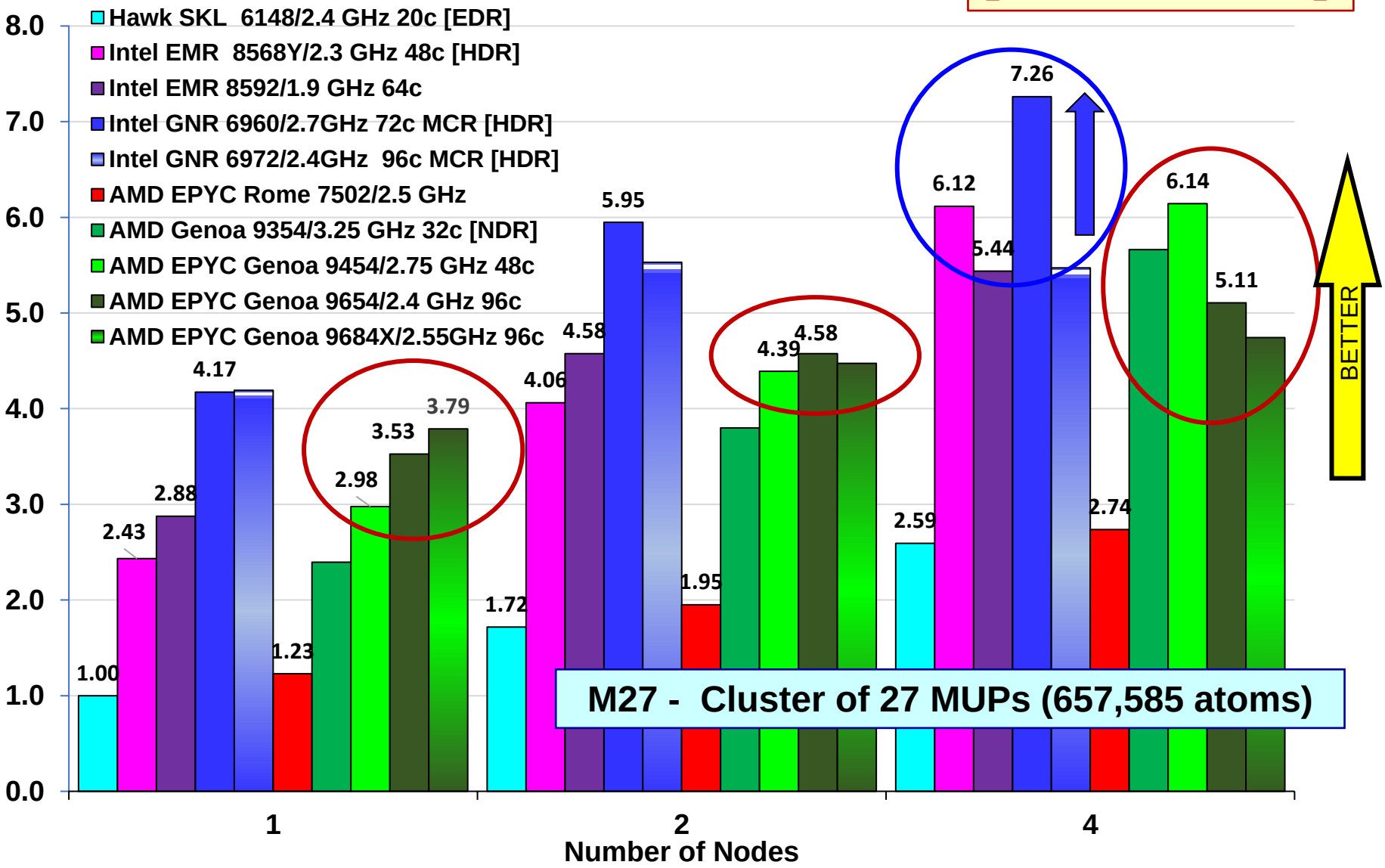


AMBER22 - M27 Performance Analysis

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Performance *Relative to the Hawk SKL 6148 2.4 GHz (40 PEs)*

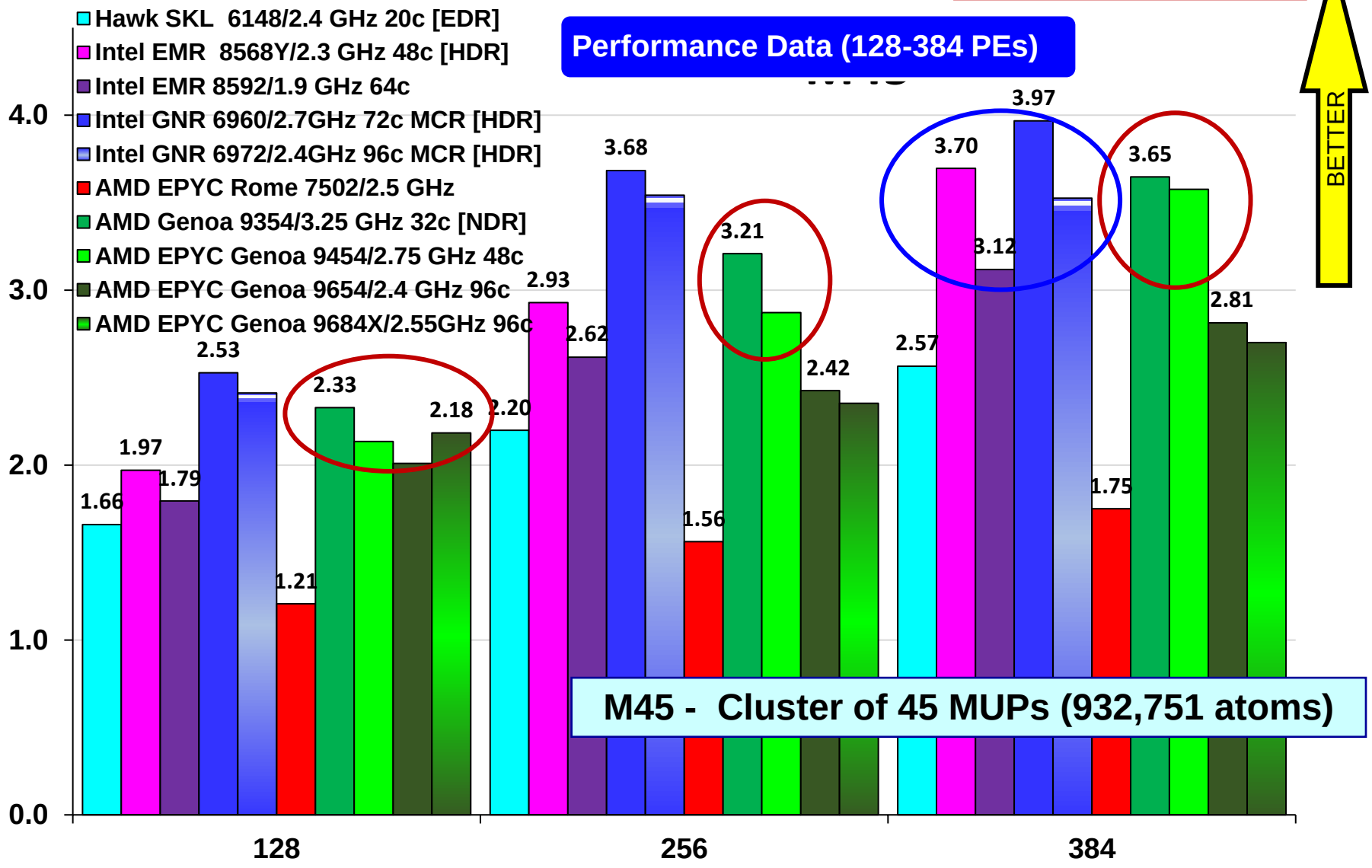
[Node to Node]



AMBER22 - M45 Performance Analysis

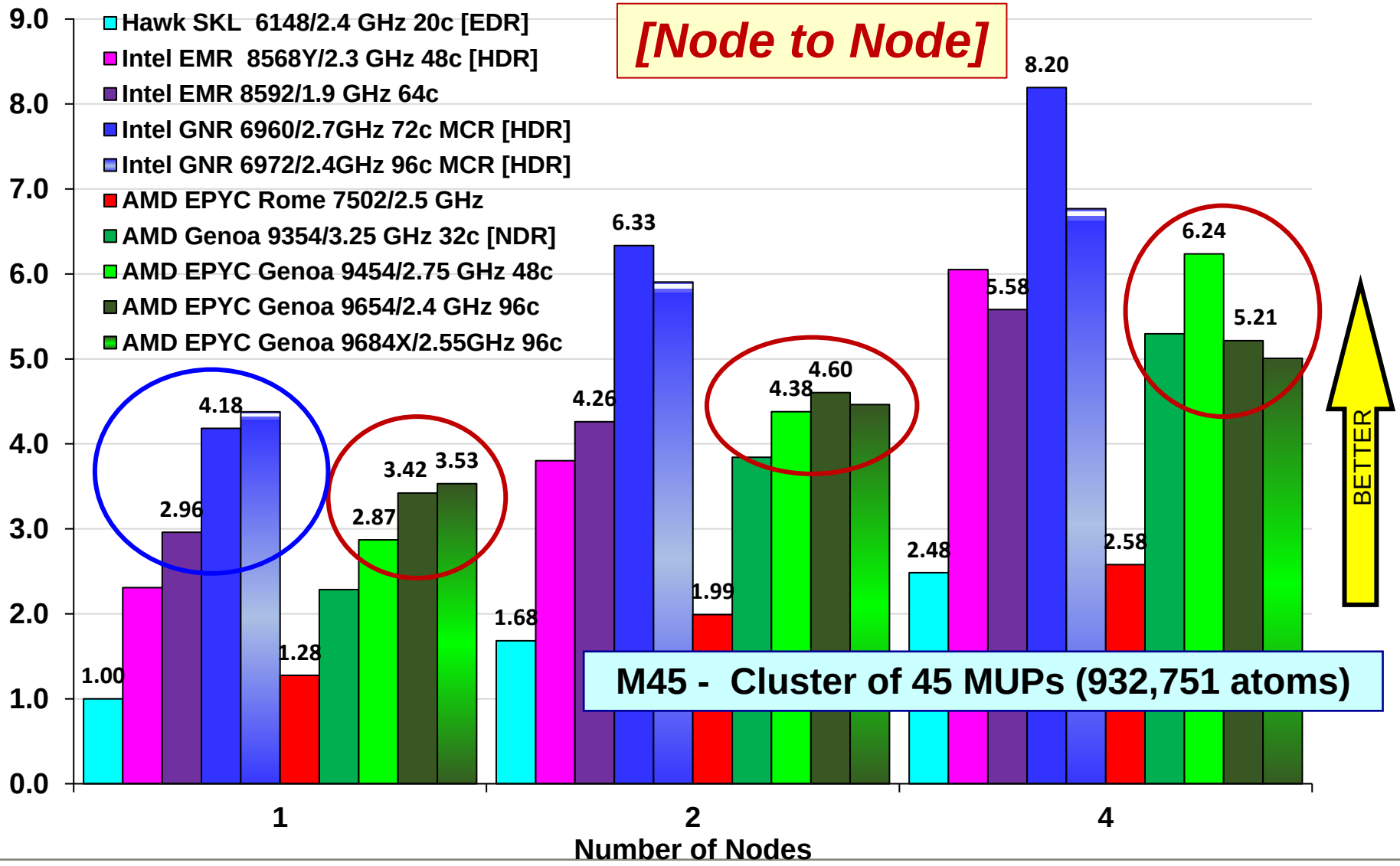
Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*

[Core to core]

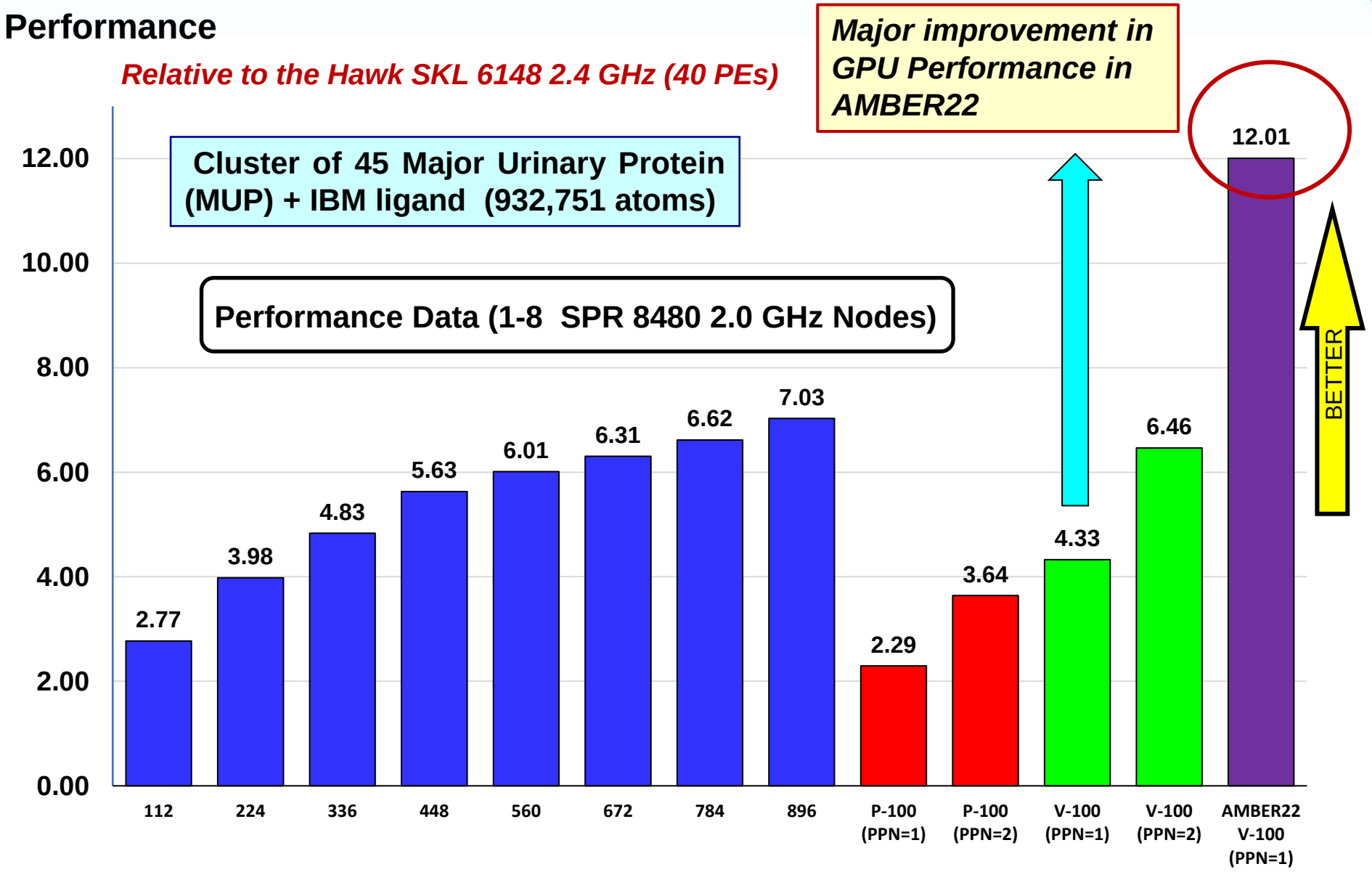


AMBER22 - M45 Performance Analysis

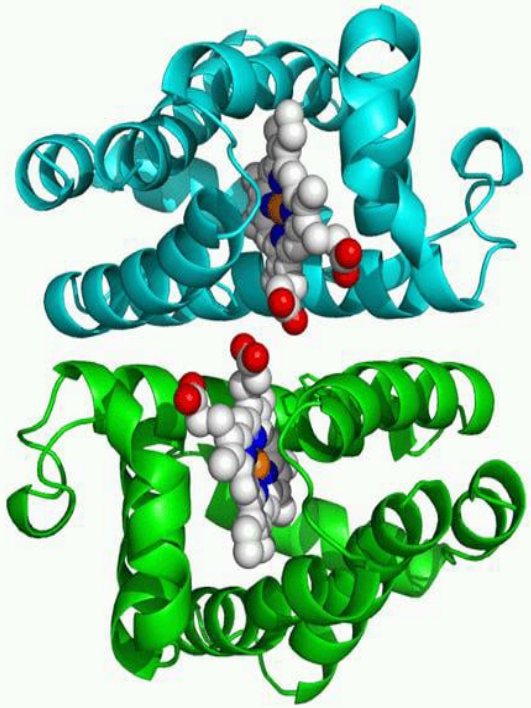
Performance *Relative to the Hawk SKL 6148 2.4 GHz (40 PEs)*

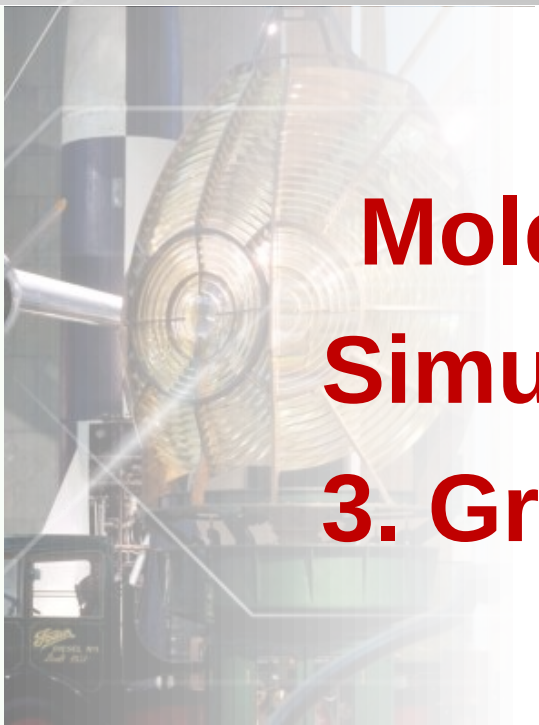


AMBER – GPU Performance M45 Simulation



Performance of Computational Chemistry Codes





Molecular Simulation: 3. Gromacs

Molecular Simulation - GROMACS

Archer Rank: 7

GROMACS (GROningen MAchine for Chemical Simulations) is a molecular dynamics package designed for simulations of proteins, lipids and nucleic acids [University of Groningen] .

Versions under Test:

Version 4.6.1 – 5 March 2013

Version 5.0.7 – 14 October 2015

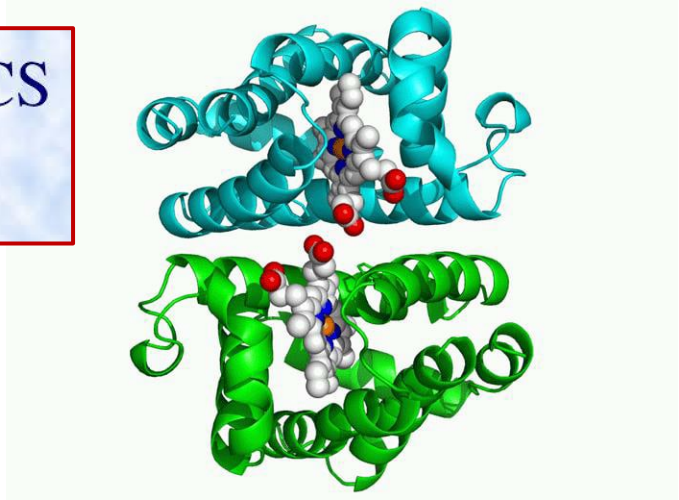
Version 2016.3 – 14 March 2017

Version 2018.2 – 14 June 2018

Version 2019.6 – 28 February 2020

Version 2020.1 – 3 March 2020

Version 2023.1 – 21 April 2023



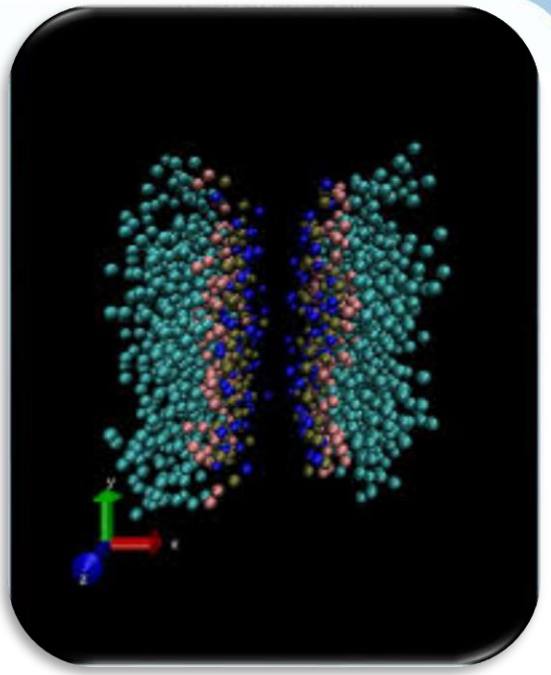
- **Berk Hess et al. "*GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*". *Journal of Chemical Theory and Computation* 4 (3): 435–447.**

http://manual.gromacs.org/documentation/

GROMACS Benchmark Cases

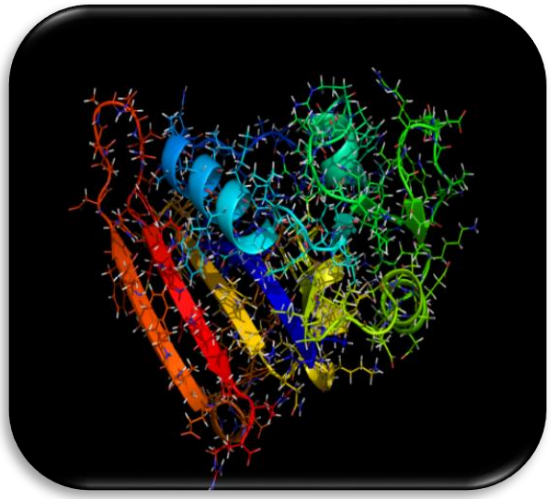
Ion channel system

- The 142k particle ion channel system is the membrane protein GluCl - a pentameric chloride channel embedded in a DOPC membrane and solvated in TIP3P water, using the Amber ff99SB-ILDN force field. This system is a **challenging** parallelization case due to the small size, but was one of the **wanted target sizes** for biomolecular simulations



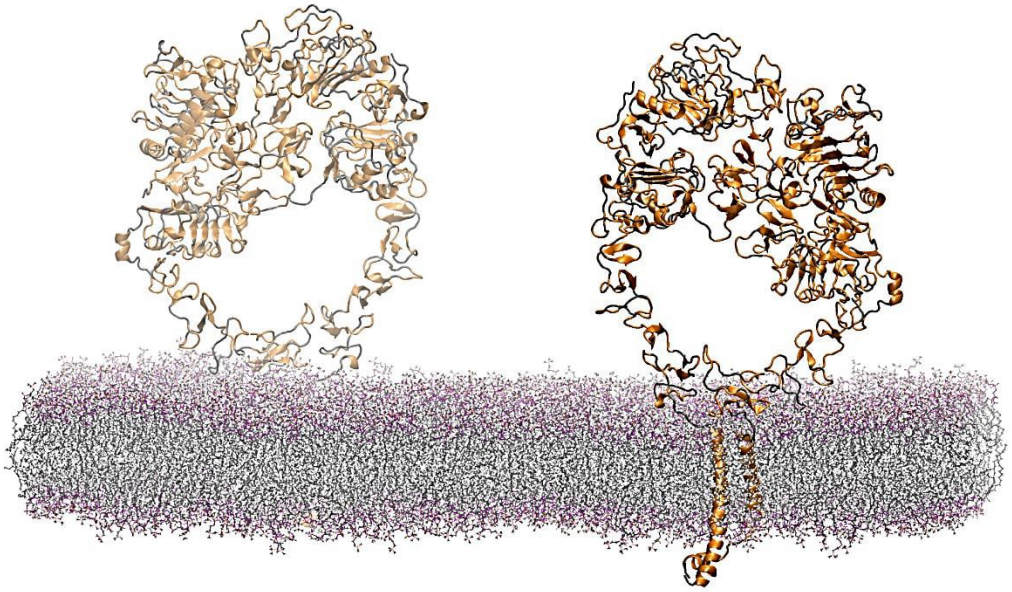
Lignocellulose

- Gromacs Test Case B from the UEA Benchmark Suite. A model of cellulose and lignocellulosic biomass in an aqueous solution. This system of 3.3M atoms is inhomogeneous, and uses **reaction-field electrostatics** instead of PME and therefore should scale well.



GROMACS Benchmark Case II

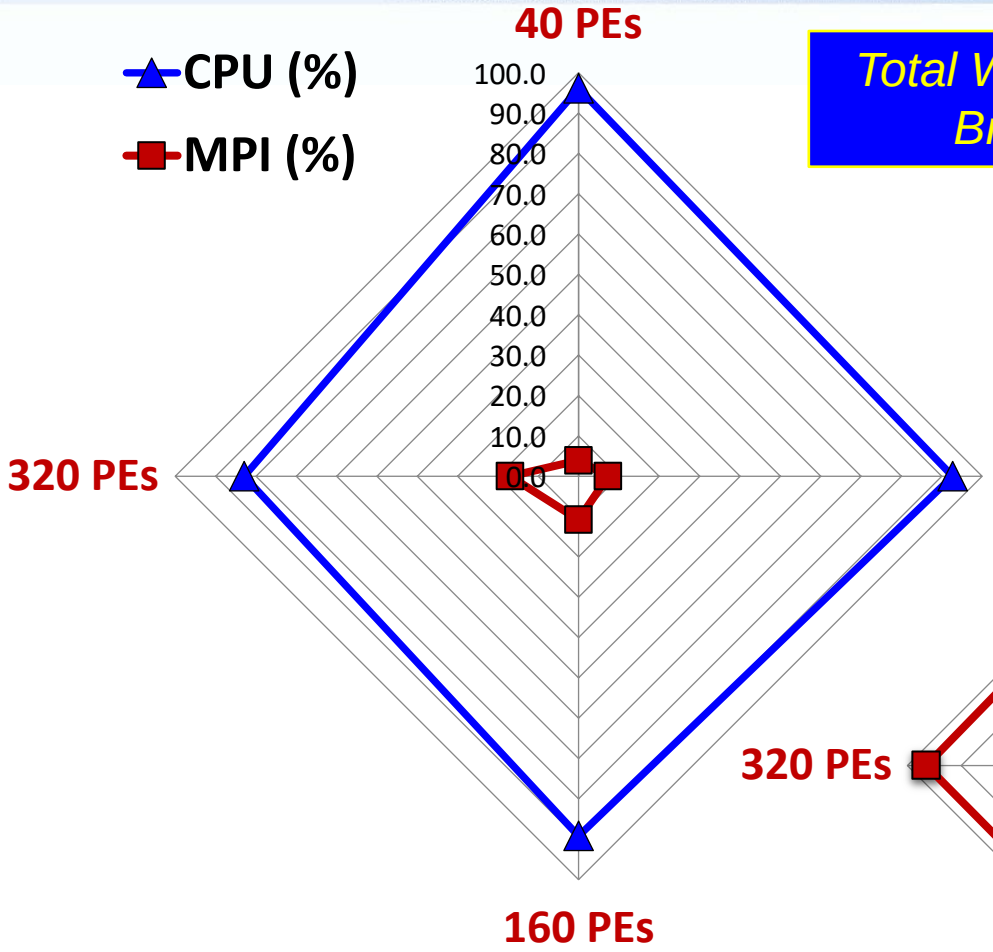
The HECBioSim Benchmarks



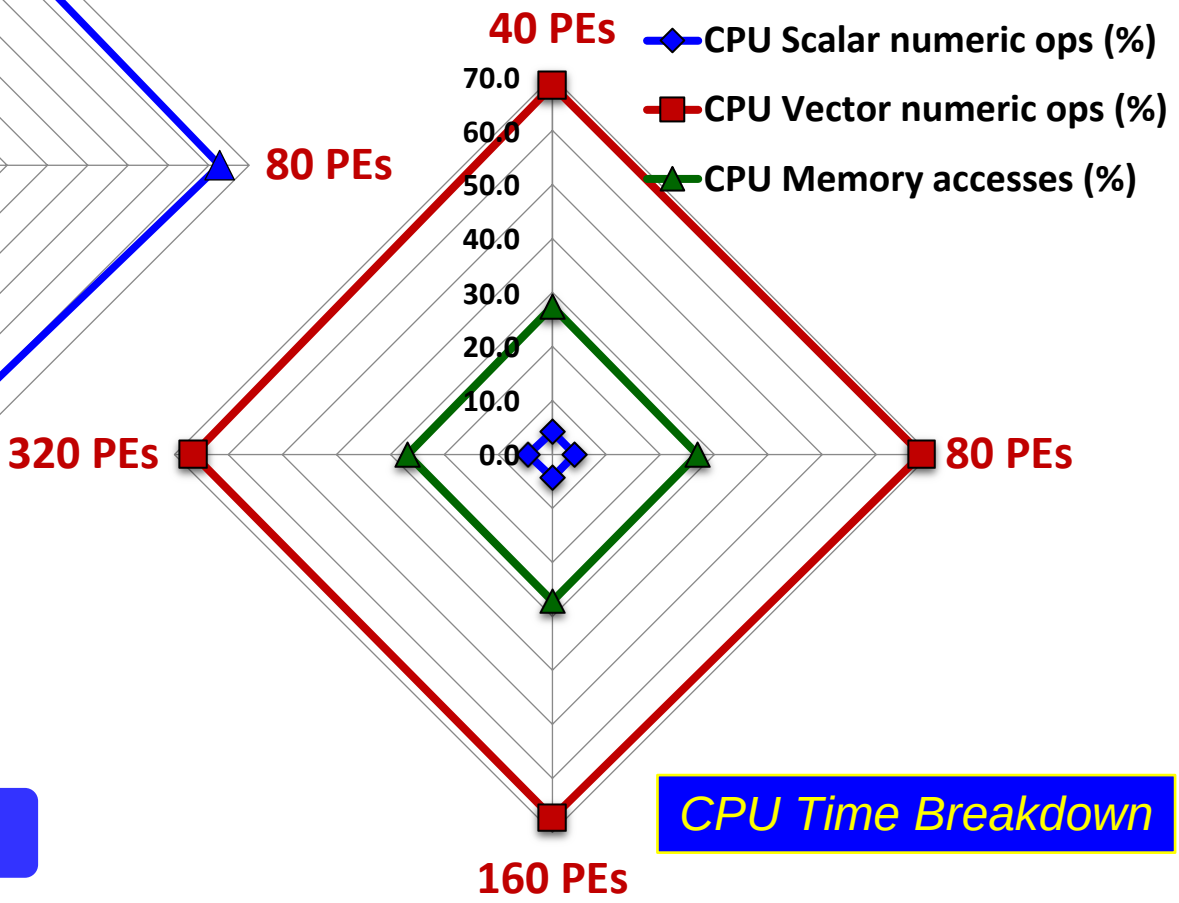
- PME simulation for 1.4M atom system - A Pair of Human Epidermal Growth Factor Receptor (hEGFR) Dimers of 1IVO and 1NQL
- Total number of atoms = **1,403,182**
- Protein atoms = 43,498 Lipid atoms = 235,304 Water atoms = 1,123,392 Ions = 986

https://www.hecbiosim.ac.uk/benchmarks

GROMACS – HECBioSim Performance Report



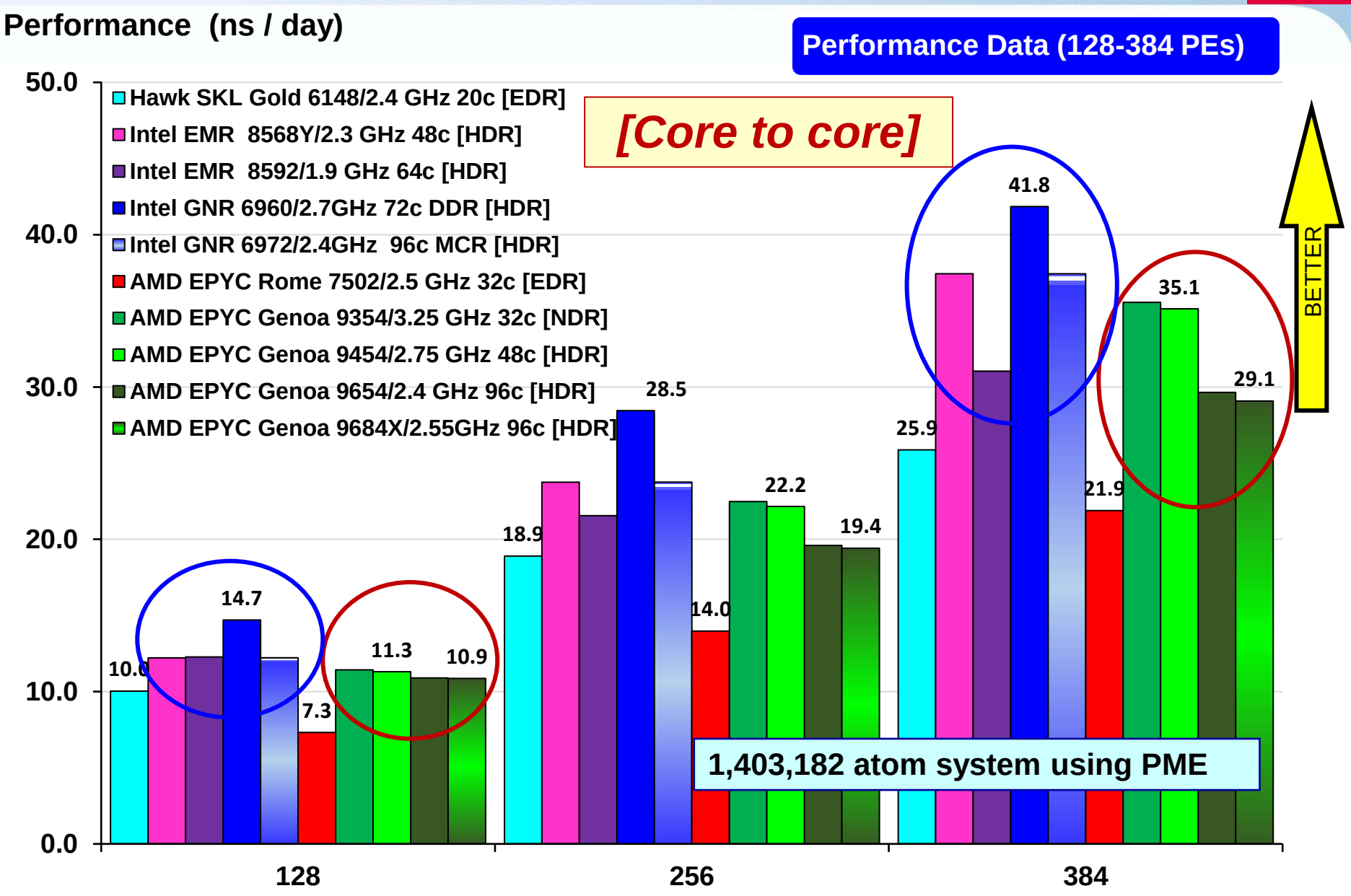
Total Wallclock Time Breakdown



CPU Time Breakdown

Performance Data (40-320 PEs)

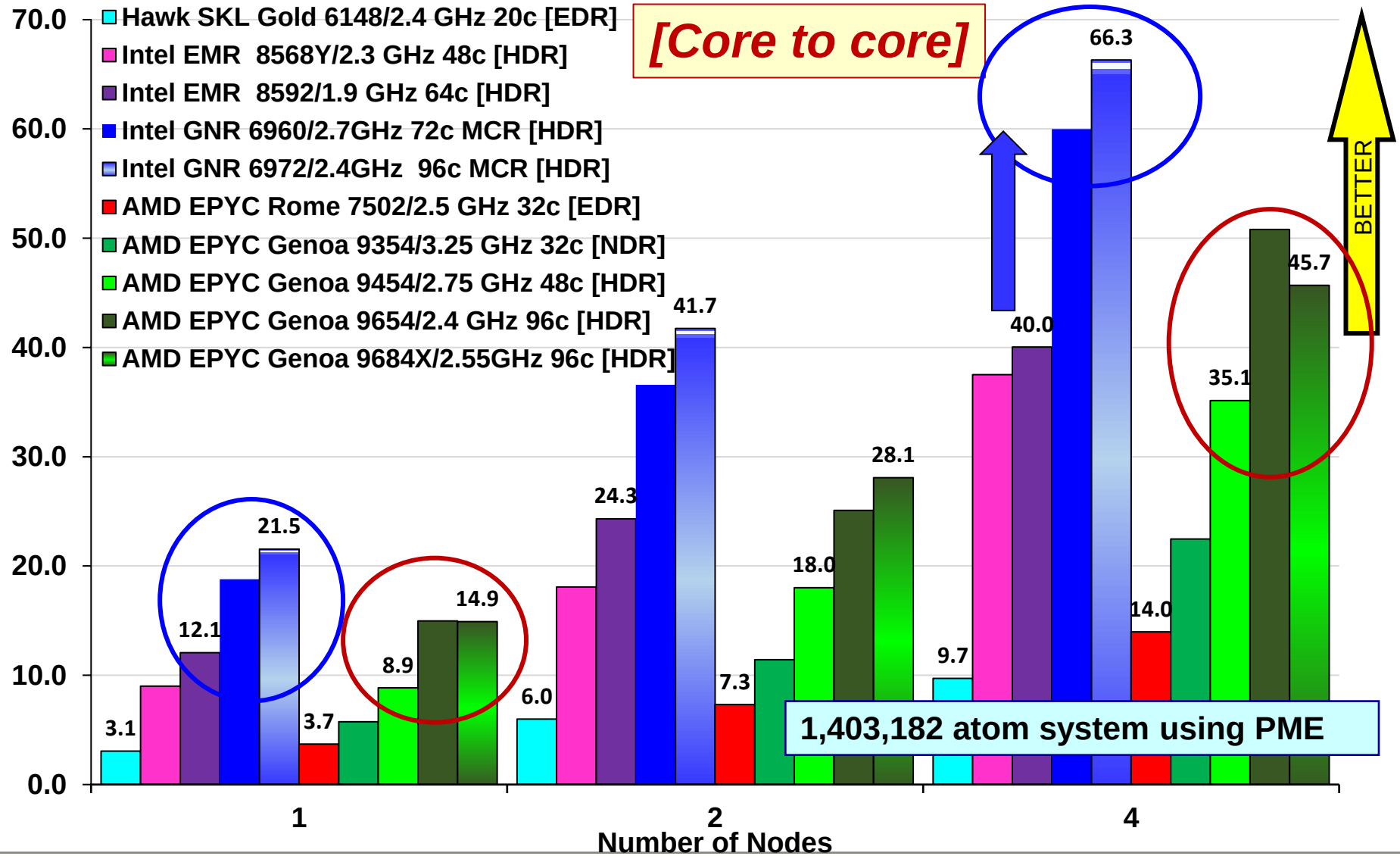
GROMACS – HECBioSim 1.4M Atom System



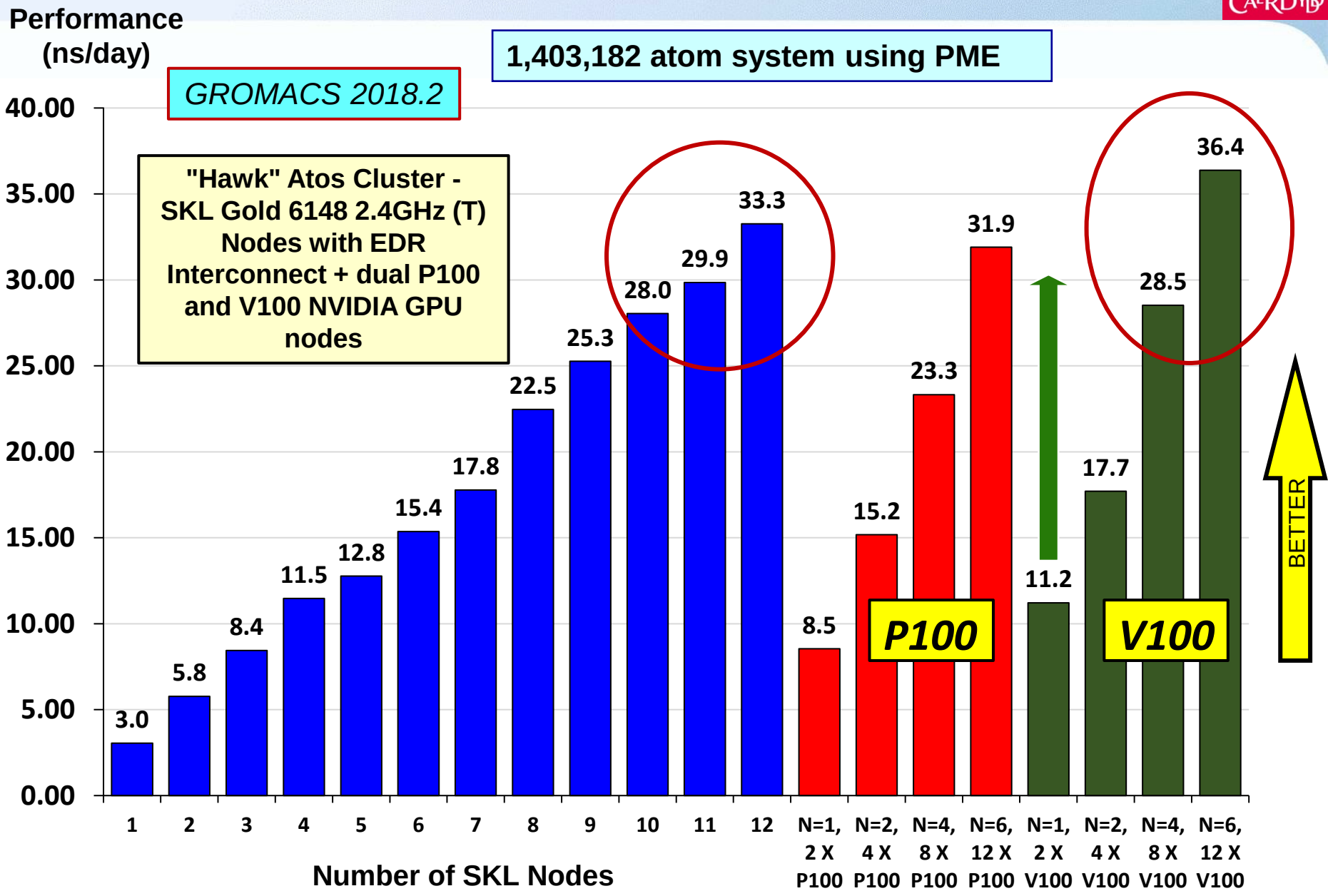
GROMACS – HECBioSim 1.4M Atom System

Performance (ns / day)

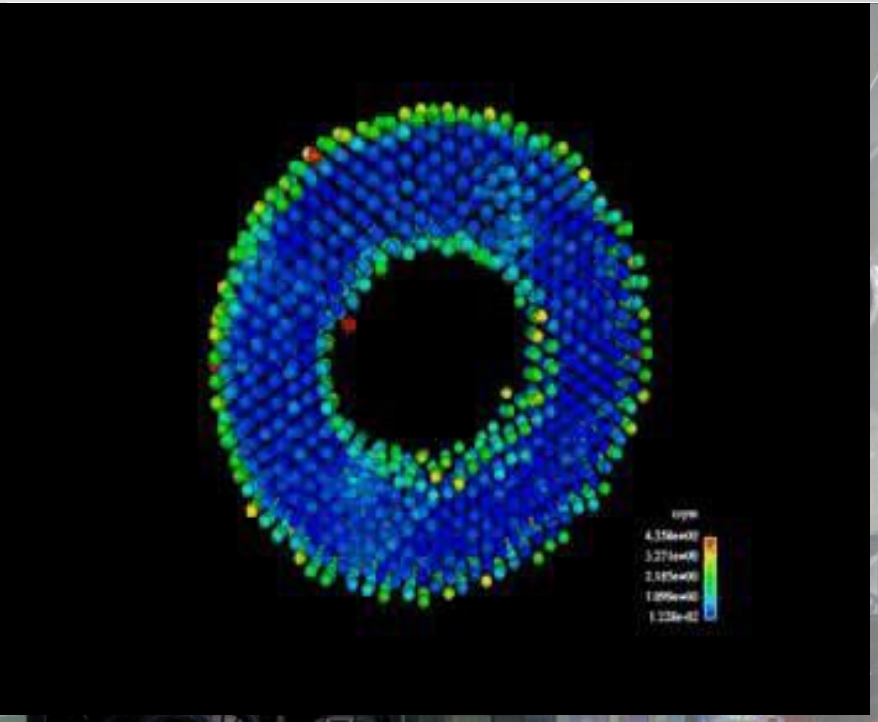
Performance Data (128-384 PEs)



GROMACS – GPU Performance: HECBioSim Simulation



Performance of Computational Chemistry and Ocean Modelling Codes



Advanced Materials Software:

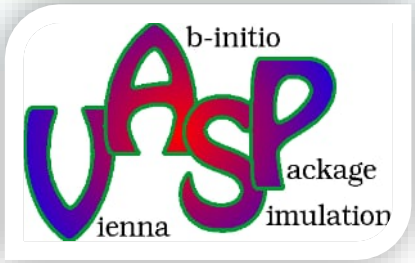
4. VASP

Advanced Materials Software

Computational Materials

- **VASP** – performs ab-initio QM molecular dynamics (MD) simulations using **pseudopotentials** or the projector-augmented wave method and a plane wave basis set.
- **Quantum Espresso** – an integrated suite of Open-Source computer codes for electronic-structure calculations and materials modelling at the nanoscale. It is based on density-functional theory (**DFT**), plane waves, and **pseudopotentials**
- **CASTEP** – a full-featured materials modelling code based on a first-principles QM description of electrons and nuclei. Uses robust methods of a **plane-wave basis set and pseudopotentials**.
- **CP2K** is a program to perform atomistic and molecular simulations of solid state, liquid, molecular, and biological systems. It provides a framework for different methods such as e.g., **DFT** using a mixed Gaussian & plane waves approach (GPW) and classical pair and many-body potentials.
- **ONETEP** (Order-N Electronic Total Energy Package) is a linear-scaling code for quantum-mechanical calculations based on **DFT**.

VASP – Vienna *Ab-initio* Simulation Package



Archer Rank: 1

VASP (**6.3**) performs ab-initio QM molecular dynamics (MD) simulations using pseudopotentials or the projector-augmented wave method and a plane wave basis set.

Benchmark	Details
MFI Zeolite	Zeolite ($Si_{96}O_{192}$), 2 k-points, FFT grid: (65, 65, 43); 181,675 points
Pd-O complex	Palladium-Oxygen complex ($Pd_{75}O_{12}$), 10 k-points, FFT grid: (31, 49, 45), 68,355 points

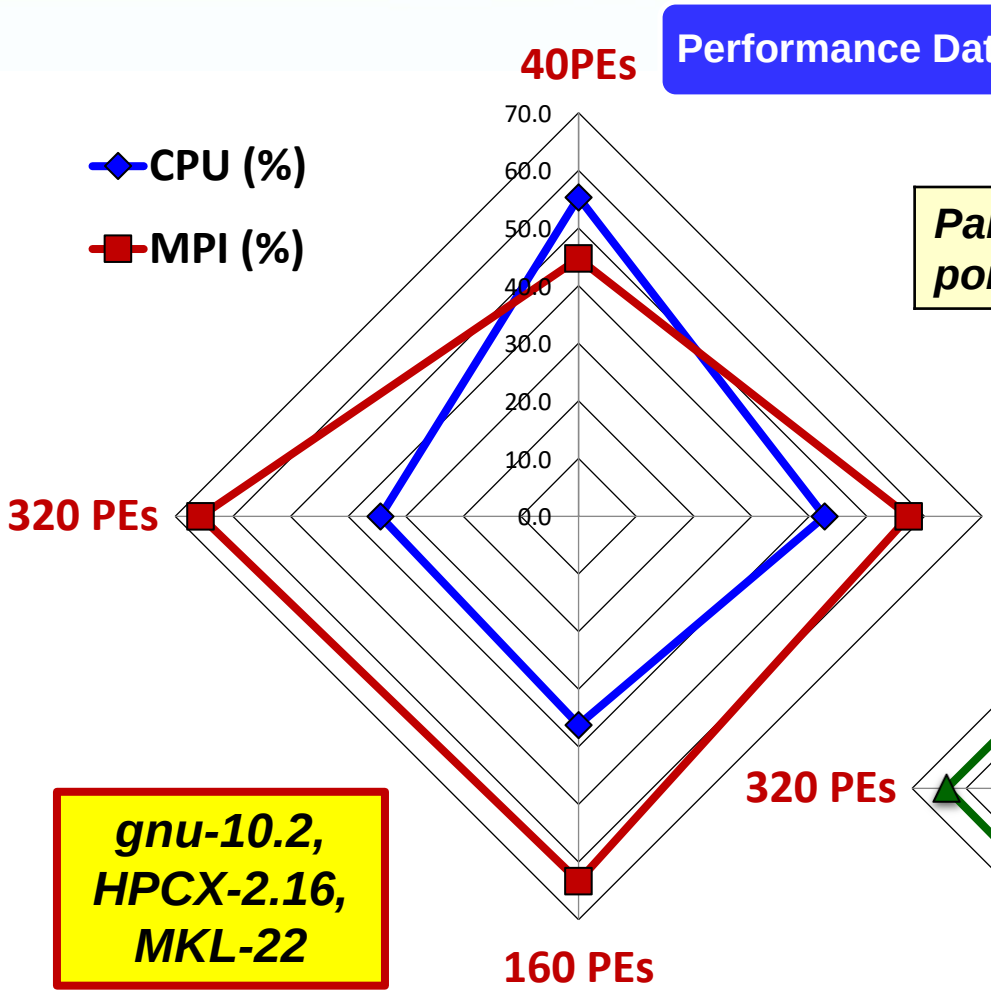
Pd-O Benchmark

- Pd-O complex – $Pd_{75}O_{12}$, 5X4 3-layer supercell running a single point calculation and a planewave cut off of 400eV. Uses the RMM-DIIS algorithm for the SCF and is calculated in real space.
- 10 k-points; maximum number of plane-waves: 34,470
- FFT grid; NGX=31, NGY=49, NGZ=45, giving a total of 68,355 points

Zeolite Benchmark

- Zeolite with the MFI structure unit cell running a single point calculation and a planewave cut off of 400eV using the PBE functional
- 2 k-points; maximum number of plane-waves: 96,834
- FFT grid; NGX=65, NGY=65, NGZ=43, giving a total of 181,675 points

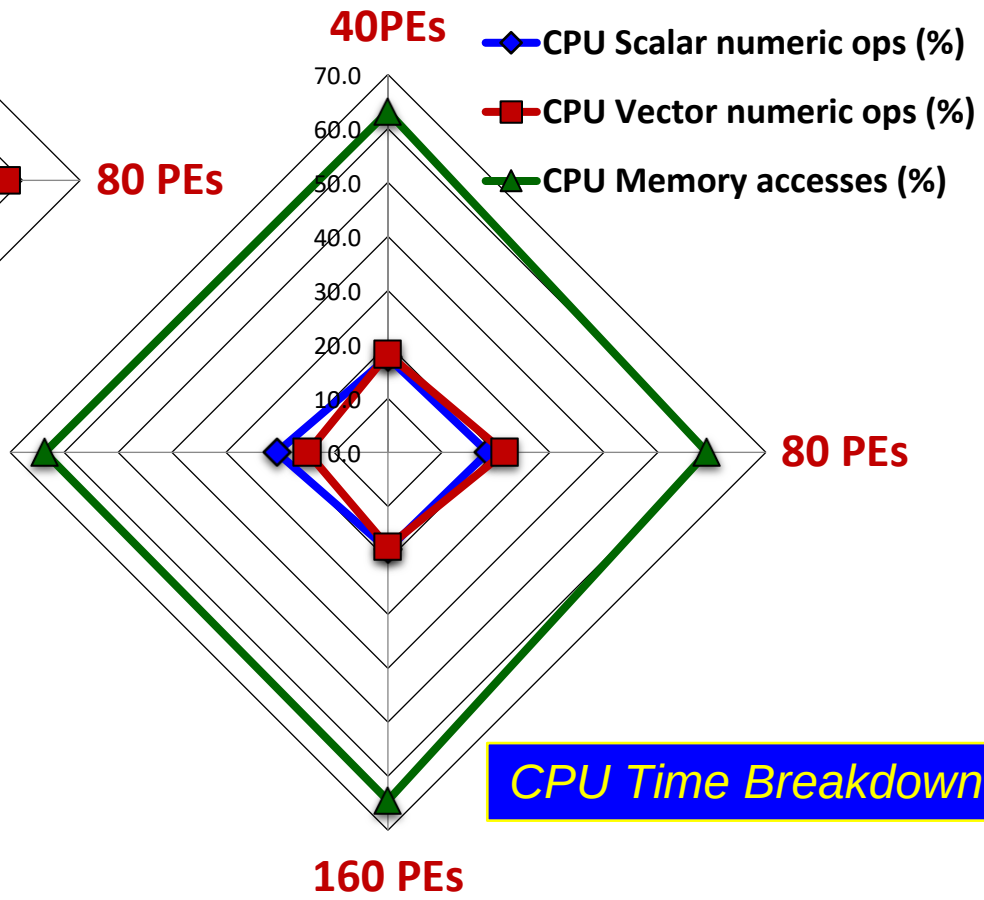
VASP – Pd-O Benchmark Performance Report



Palladium-Oxygen complex ($Pd_{75}O_{12}$), 10 k-points, FFT grid: (31, 49, 45), 68,355 points

gnu-10.2, HPCX-2.16, MKL-22

Total Wallclock Time Breakdown



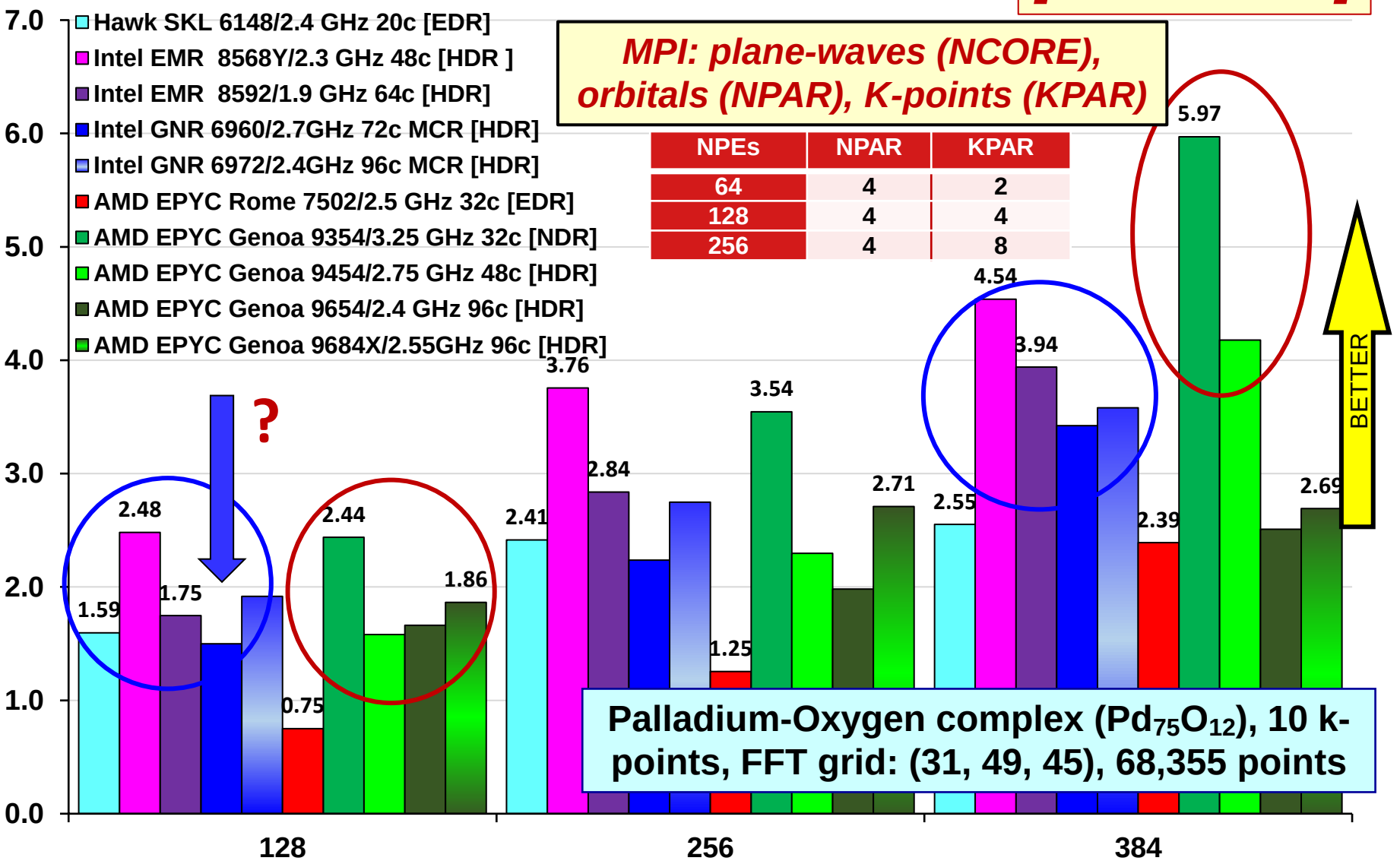
CPU Time Breakdown

VASP 6.3 – Pd-O Benchmark - Parallelisation on k-points

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*

[Core to core]



MPI: plane-waves (NCORE), orbitals (NPAR), K-points (KPAR)

NPEs	NPAR	KPAR
64	4	2
128	4	4
256	4	8

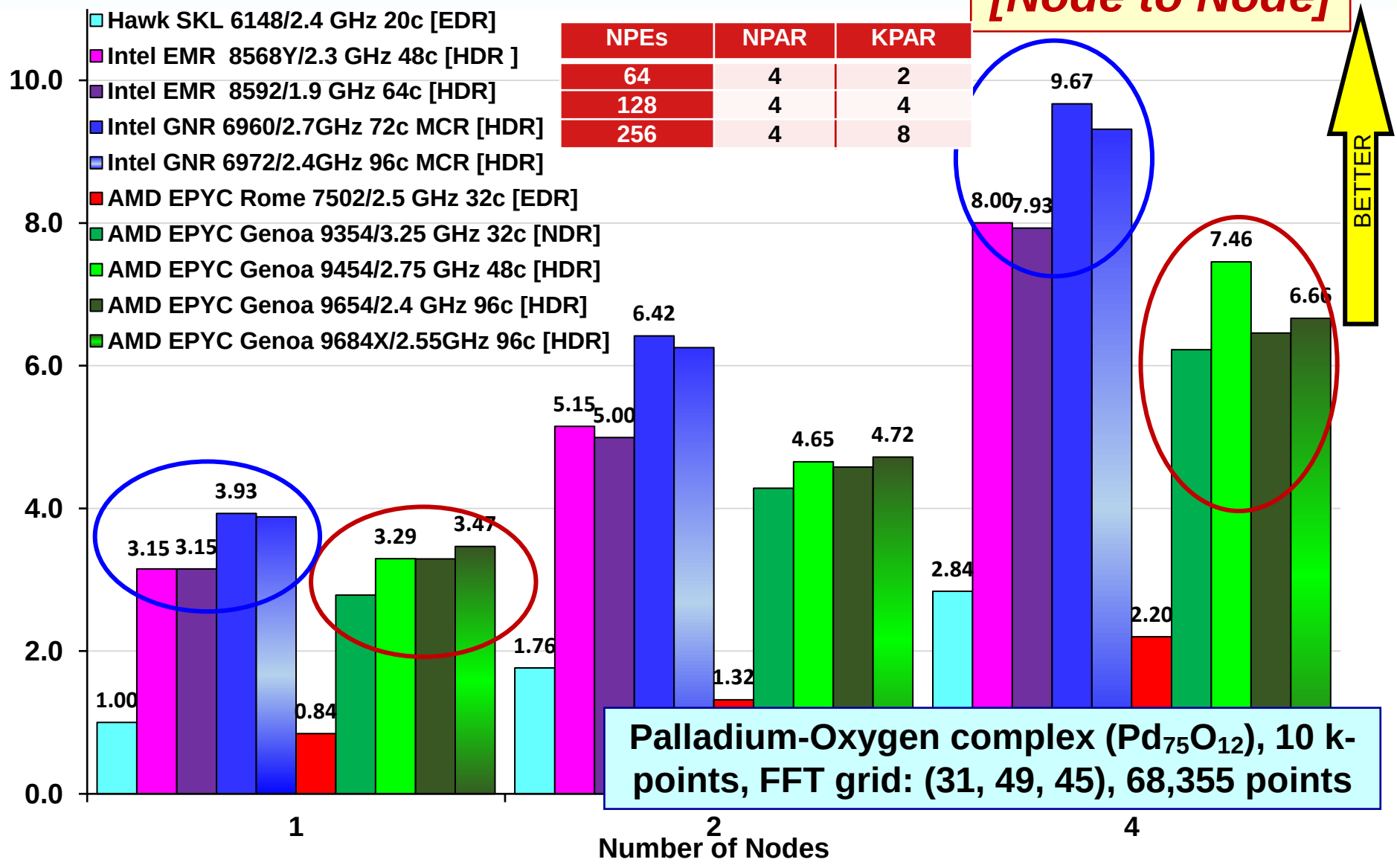
Palladium-Oxygen complex ($Pd_{75}O_{12}$), 10 k-points, FFT grid: (31, 49, 45), 68,355 points

VASP 6.3 – Pd-O Benchmark - Parallelisation on k-points

Performance *Relative to the Hawk SKL 6148 2.4 GHz (1 Node)*

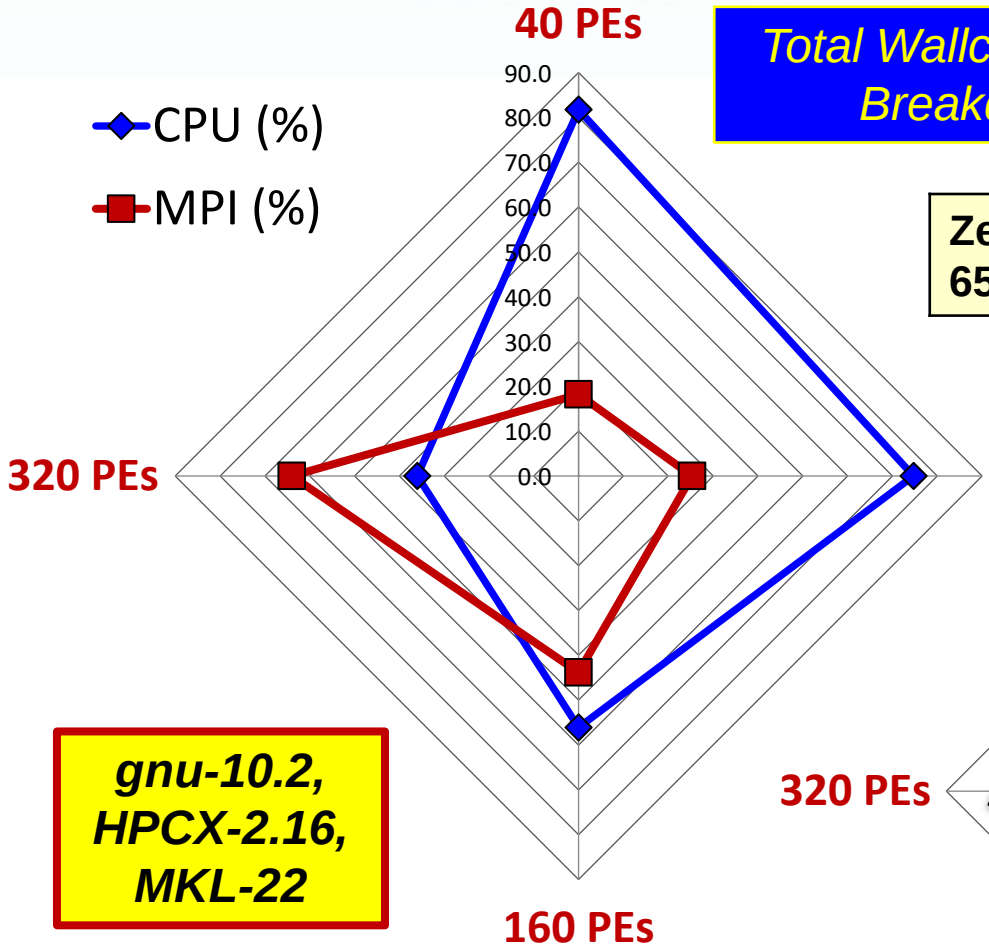
[Node to Node]

NPEs	NPAR	KPAR
64	4	2
128	4	4
256	4	8



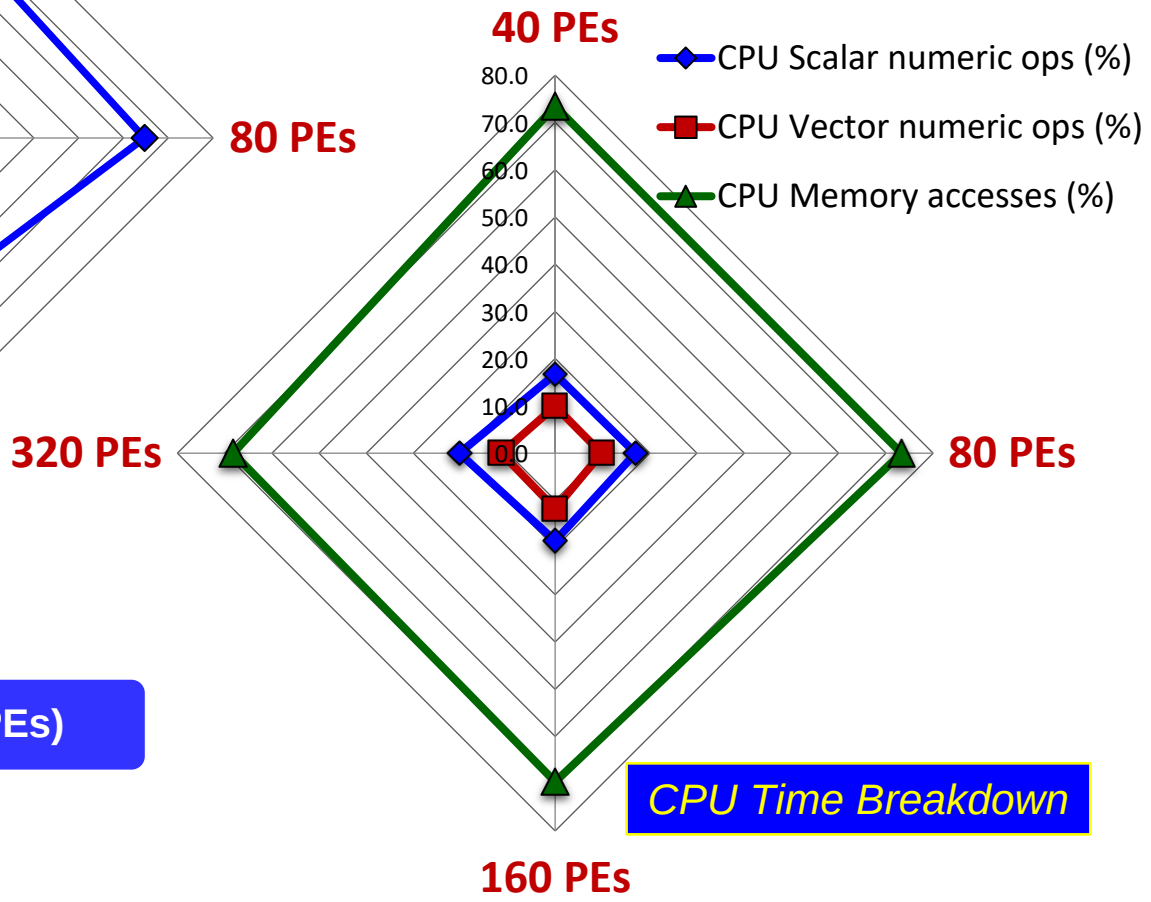
Palladium-Oxygen complex ($Pd_{75}O_{12}$), 10 k-points, FFT grid: (31, 49, 45), 68,355 points

VASP – Zeolite Cluster Performance Report



Total Wallclock Time Breakdown

Zeolite ($Si_{96}O_{192}$), 2 k-points, FFT grid: (65, 65, 43); 181,675 points



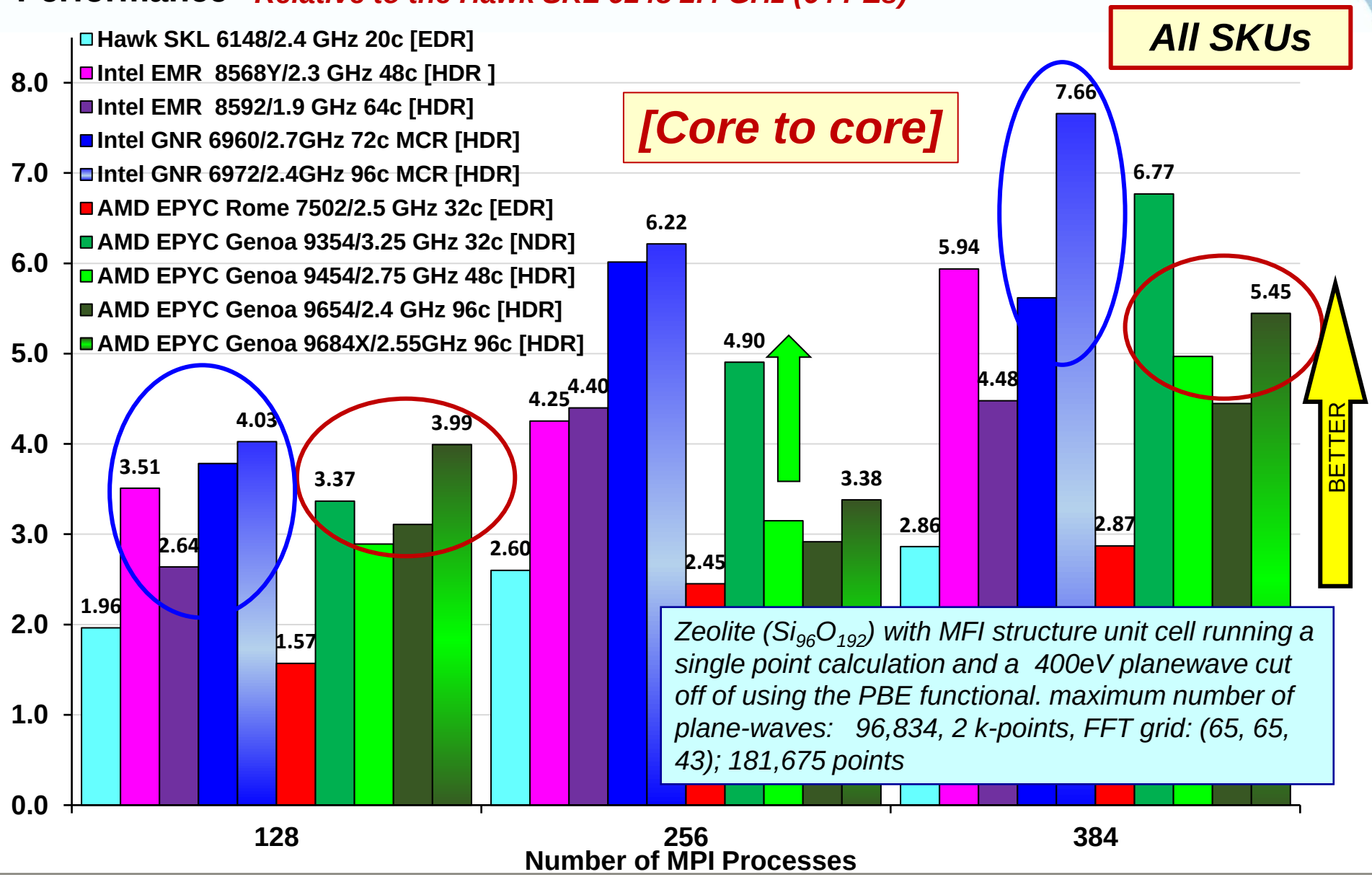
CPU Time Breakdown

gnu-10.2, HPCX-2.16, MKL-22

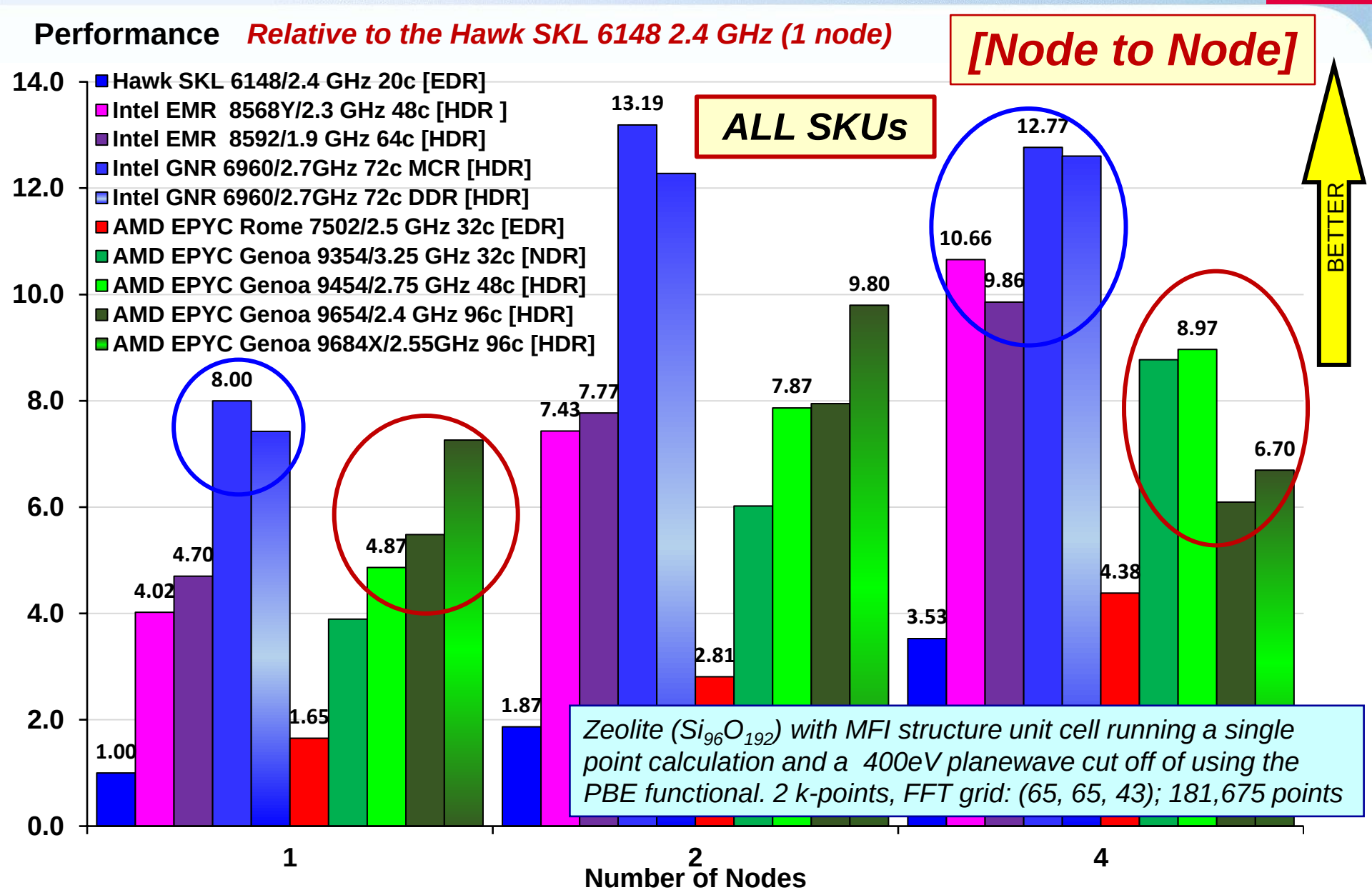
Performance Data (40-320 PEs)

VASP 6.3 – Zeolite Benchmark - Parallelisation on k-points

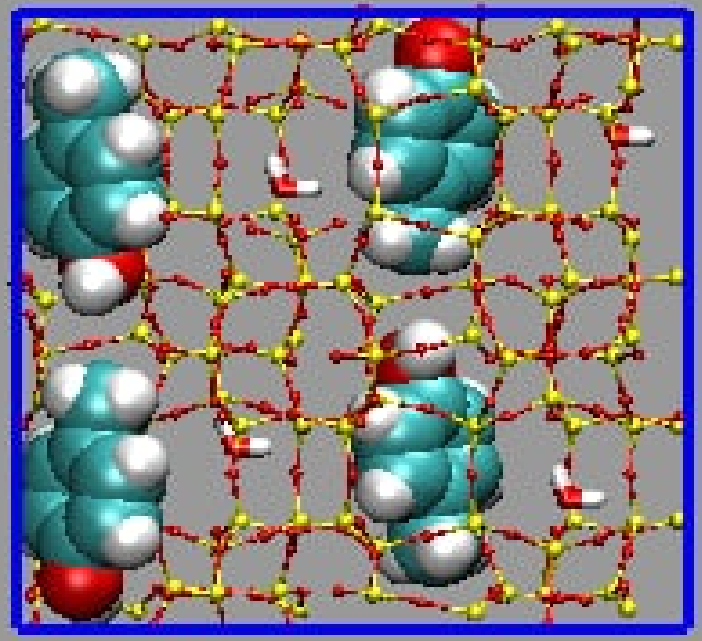
Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*



VASP 6.3 – Zeolite Benchmark - Parallelisation on k-points



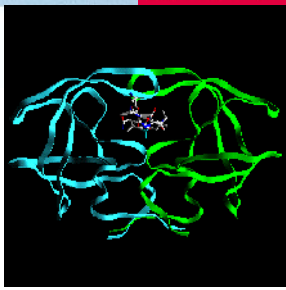
Performance of Computational Chemistry and Ocean Modelling Codes



Electronic Structure

5. GAMESS -UK

The MPI/ScaLAPACK Implementation of the GAMESS-UK SCF/DFT module

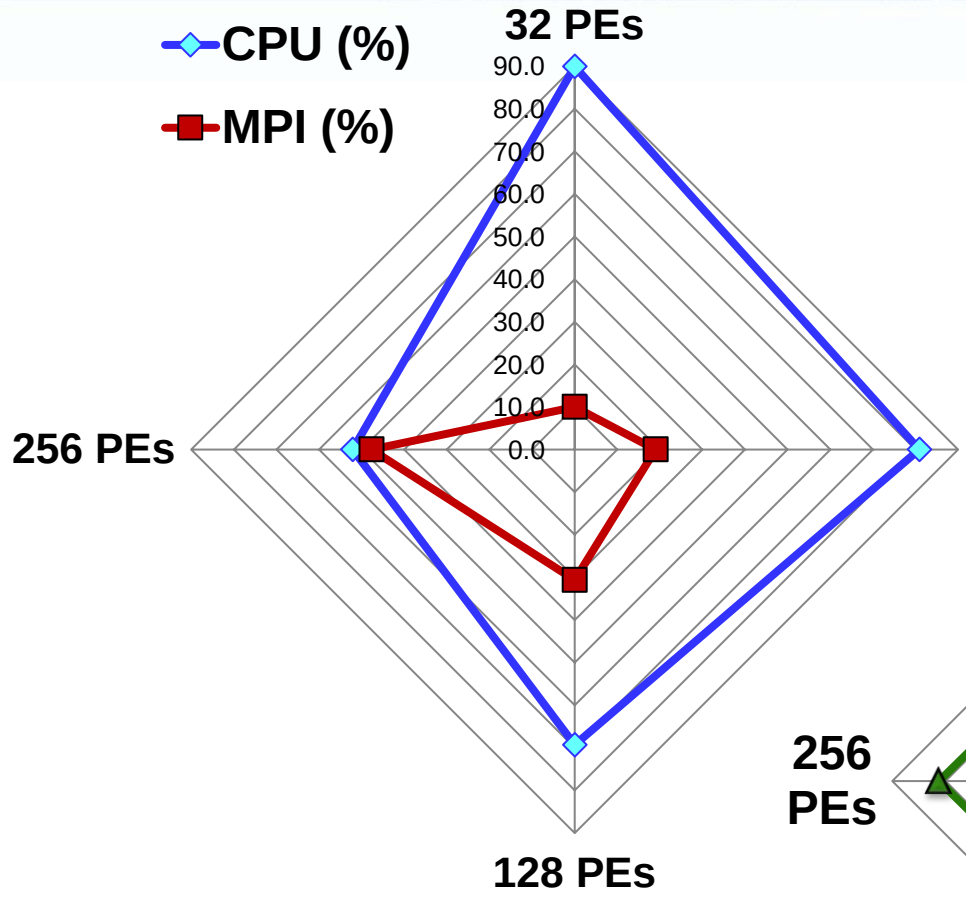


- Pragmatic approach to the replicated data constraints:
- MPI-based tools (such as ScaLAPACK) used in place of Global Arrays
- All data structures except those required for the Fock matrix build are fully distributed (F, P)

- Partially distributed model chosen because, in the absence of efficient one-sided communications it is difficult to efficiently load balance a distributed Fock matrix build.
- Obvious drawback - some large replicated data structures are required.
  - These are kept to a minimum. For a closed shell HF or DFT calculation only **2 replicated matrices** are required, 1 × Fock and 1 × Density (doubled for UHF).

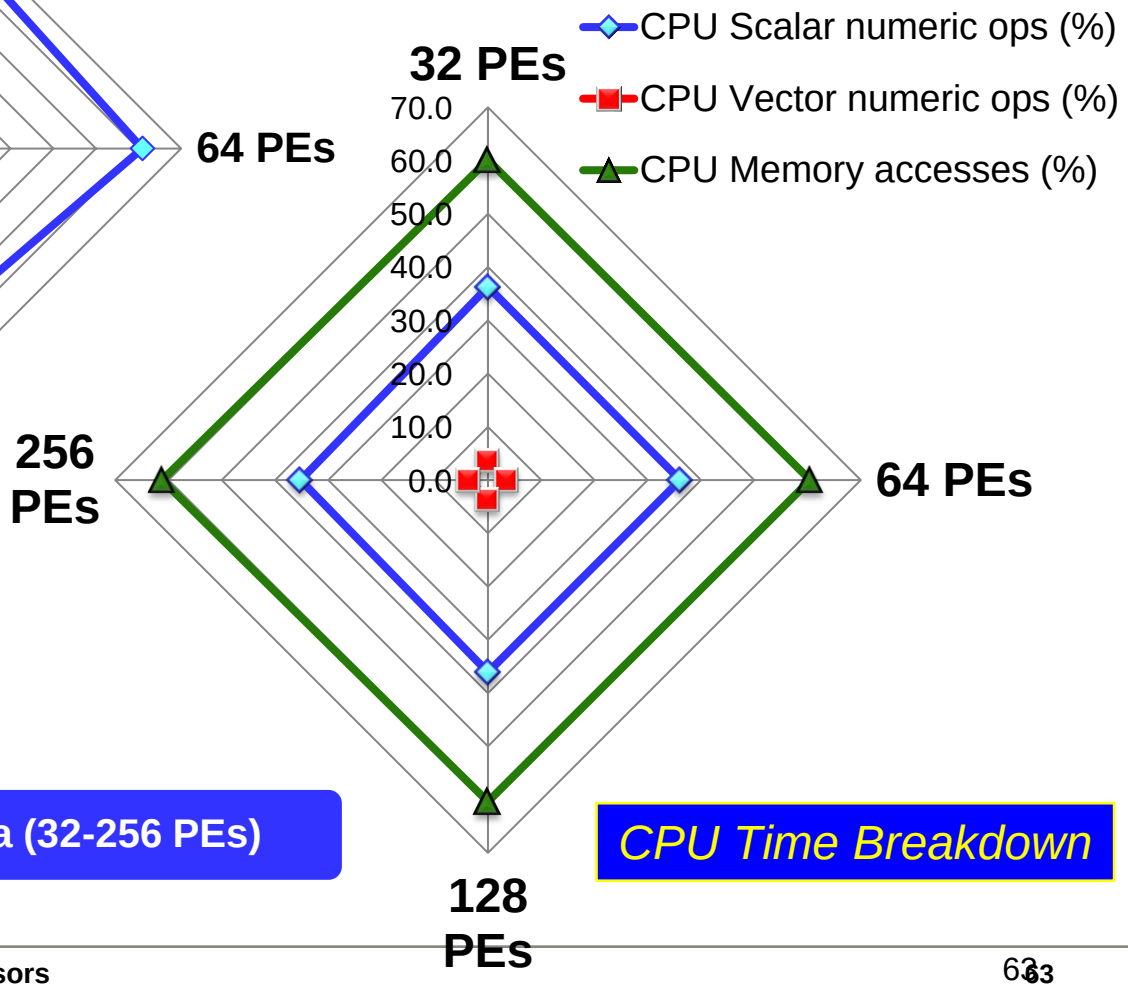
"The GAMESS-UK electronic structure package: algorithms, developments and applications"
M.F. Guest, I. J. Bush, H.J.J. van Dam, P. Sherwood, J.M.H. Thomas, J.H. van Lenthe, R.W.A Havenith, J. Kendrick, Mol. Phys. 103, No. 6-8, 2005, 719-747.

GAMESS-UK.MPI DFT – DFT Performance Report



Cyclosporin 6-31G** basis (1855 GTOs); DFT B3LYP

Total Wallclock Time Breakdown



Performance Data (32-256 PEs)

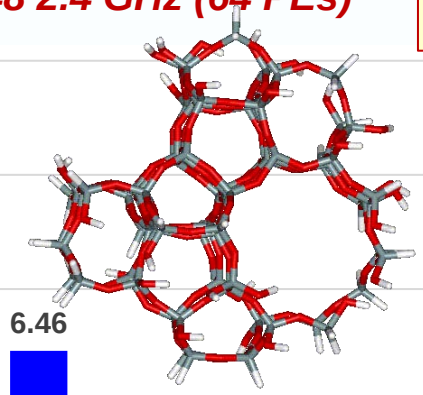
CPU Time Breakdown

GAMESS-UK Performance - Zeolite Y cluster

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Performance *Relative to the Hawk SKL 6148 2.4 GHz (64 PEs)*

[Core to core]



- Hawk SKL Gold 6148/2.4 GHz 20c [EDR]
- Intel SPR 8480/2.0 GHz 56c [NDR]
- Intel EMR 8592/1.9 GHz 64c [HDR]
- Intel GNR 6960/2.7GHz 72c MCR [HDR]
- Intel GNR 6972/2.4GHz 96c MCR [HDR]
- AMD EPYC Rome 7502/2.5 GHz 32c [EDR]
- AMD EPYC Genoa 9354/3.25 GHz 32c [NDR]
- AMD EPYC Genoa 9654/2.4 GHz 96c [HDR]
- AMD EPYC Genoa 9684X/2.55GHz 96c [HDR]

9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0

1.84 2.81 3.64 3.23 3.20 2.96 6.46 5.66 4.31 2.95 6.85 8.26 7.24 6.44

BETTER

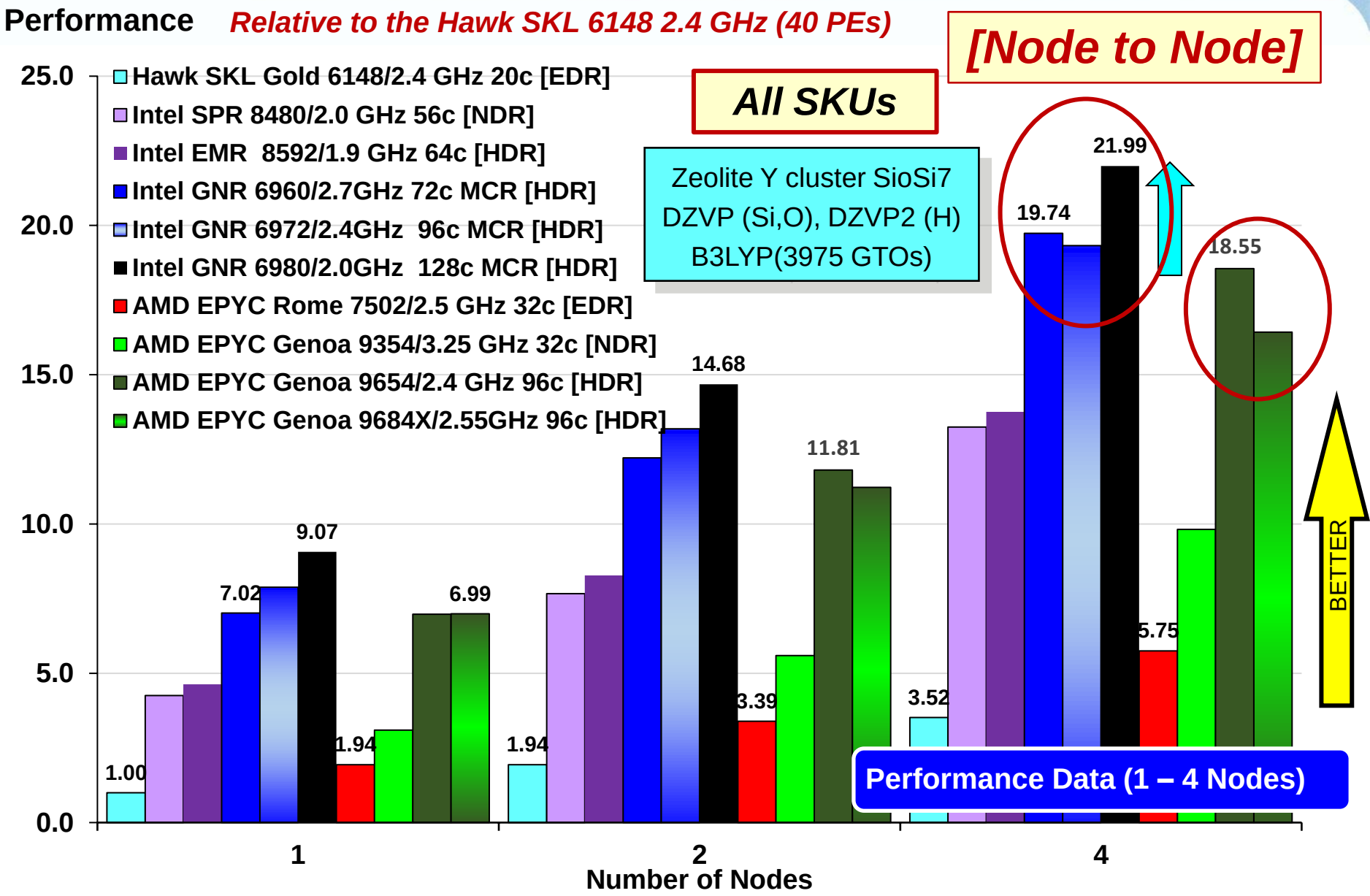
Performance Data (128 - 384 PEs)

Zeolite Y cluster SioSi7 DZVP (Si,O), DZVP2 (H) B3LYP(3975 GTOs)

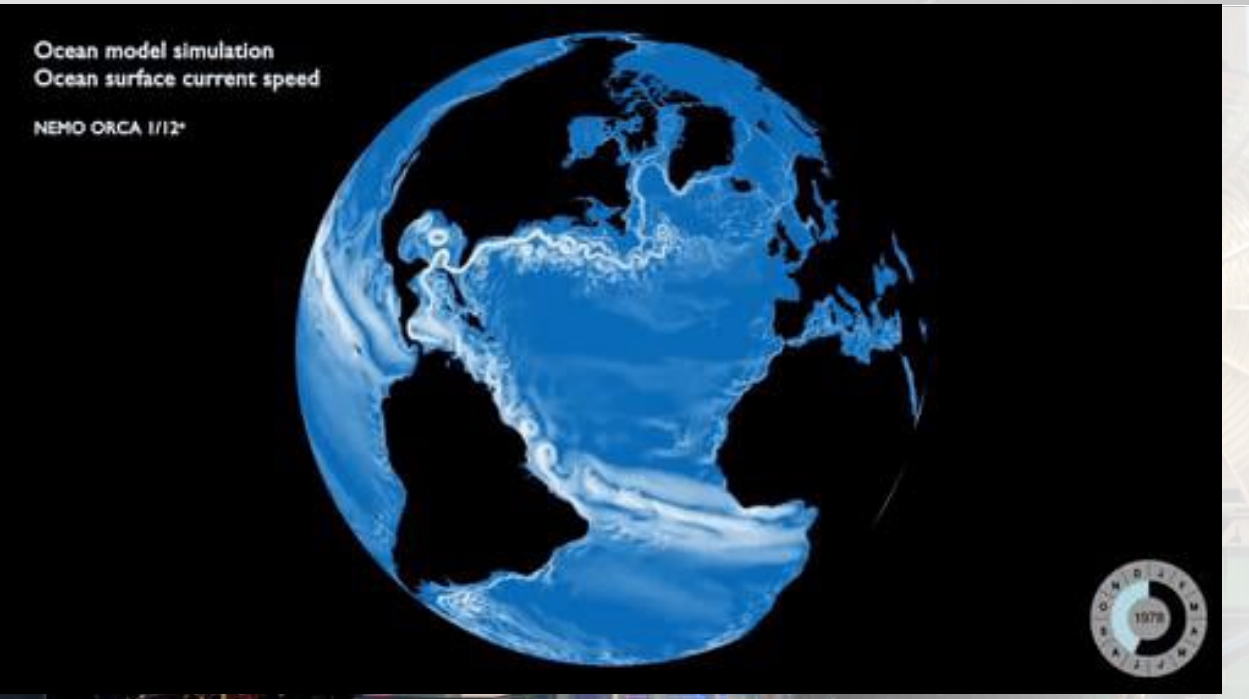
128 256 384

Number of MPI Processes

GAMESS-UK Performance - Zeolite Y cluster



Performance of Computational Chemistry and Ocean Modelling Codes



Ocean Modelling: 6. NEMO

Performance of Ocean Modelling Codes

- Assistance provided to **The Marine Systems Modelling Group at Plymouth Marine Laboratory.**
- At the heart of much of the group's work are two numerical models of the ocean's circulation:

  The NEMO Community Ocean Model

  A prognostic, primitive equation ocean circulation model for studying problems relating to both the global ocean and marginal seas. Uses a ***structured* model grid.**

  The Finite Volume Community Ocean Model (FVCOM)

  A prognostic, primitive equation ocean circulation model for (mainly) studying problems relating to estuarine and coastal environments. **Uses an *unstructured* model grid.**

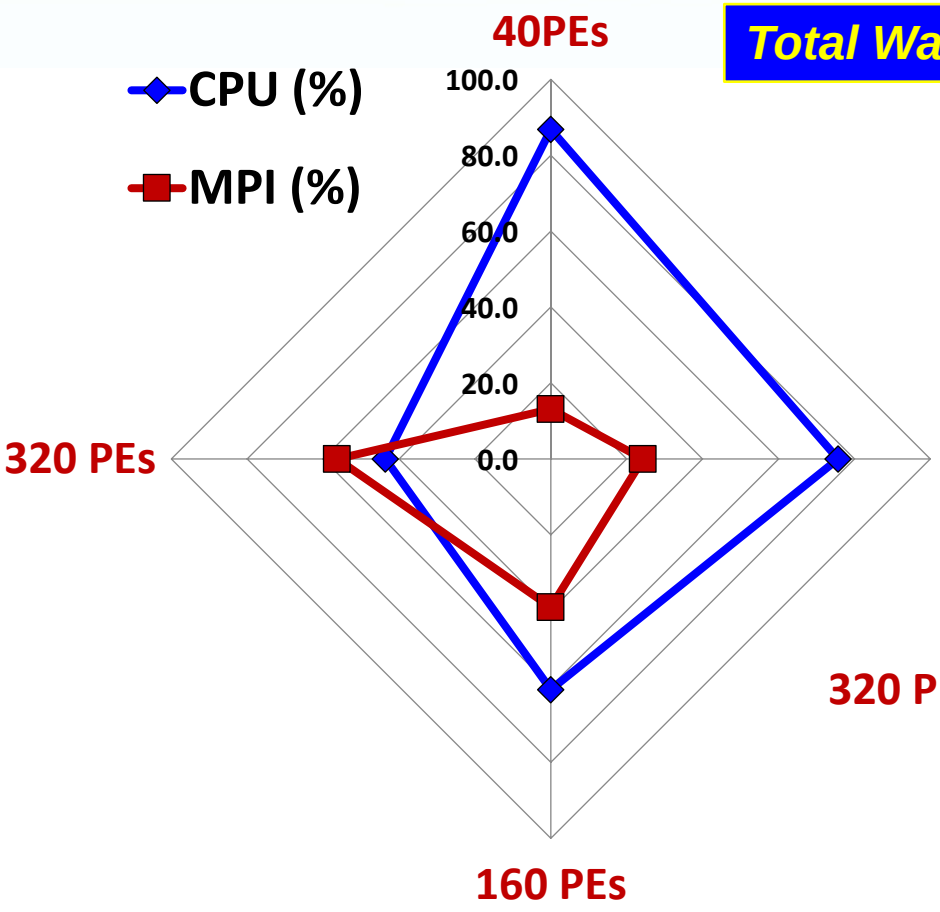
- Both models are often run with a **biogeochemical model called ERSEM -** significantly increases the compute & memory requirements.
- To be run efficiently, both models require a CPU based HPC system

The NEMO-ERSEM Benchmark

- NEMO, "Nucleus for European Modelling of the Ocean" is a modelling framework for research activities and forecasting services in ocean and climate sciences, developed by a European consortium. (https://www.nemo-ocean.eu)
- NEMO is a **memory-bandwidth limited code** where performance can be improved by part-populating nodes.
- ERSEM, “European Regional Seas Ecosystem Model” is a bio-geochemical and ecosystem mode, developed at PML (https://github.com/pmlmodelling/ersem)
- **Benchmark Case:** NEMO-FABM-ERSEM on the AMM7 (Atlantic Margin Model) domain covering the NW European shelf at ca. 7 km resolution. Four elements to the code (a) **XIOS**: an I/O library, (b) **ERSEM:** Biogeochemical model code, (c) **FABM:** Interface between ERSEM and NEMO and (d) **NEMO**.
- Compilation requires **parallel netcdf and hdf5 libraries**. Several cores are allocated to the I/O server XIOS, with remainder allocated to NEMO:

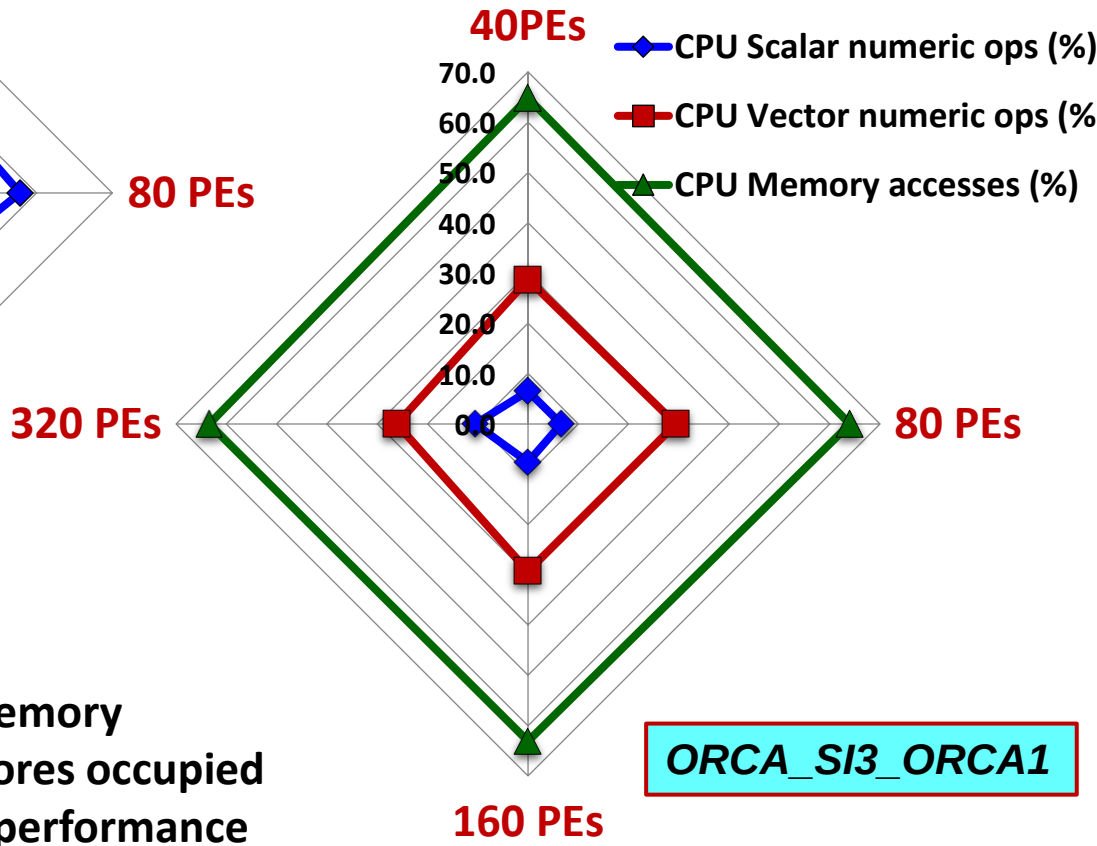
mpirun -n *$XIOSCORES* $code_xios : -n *$OCEANCORES* $code_nemo

NEMO – ORCA_SI3 Model Performance Report



Total Wallclock Time Breakdown

horizontal resolutions of 1-degree



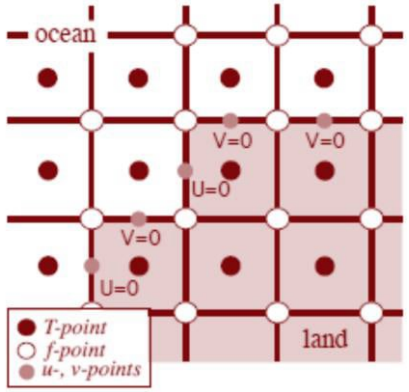
NEMO performance is dominated by memory bandwidth – running with 50% of the cores occupied on each Hawk node typically improves performance by ca. 1.6 for a fixed number of MPI processes.

ORCA_SI3_ORCA1

CPU Time Breakdown

Primary Requirements for PML

NEMO* - Data parallelism through domain decomposition

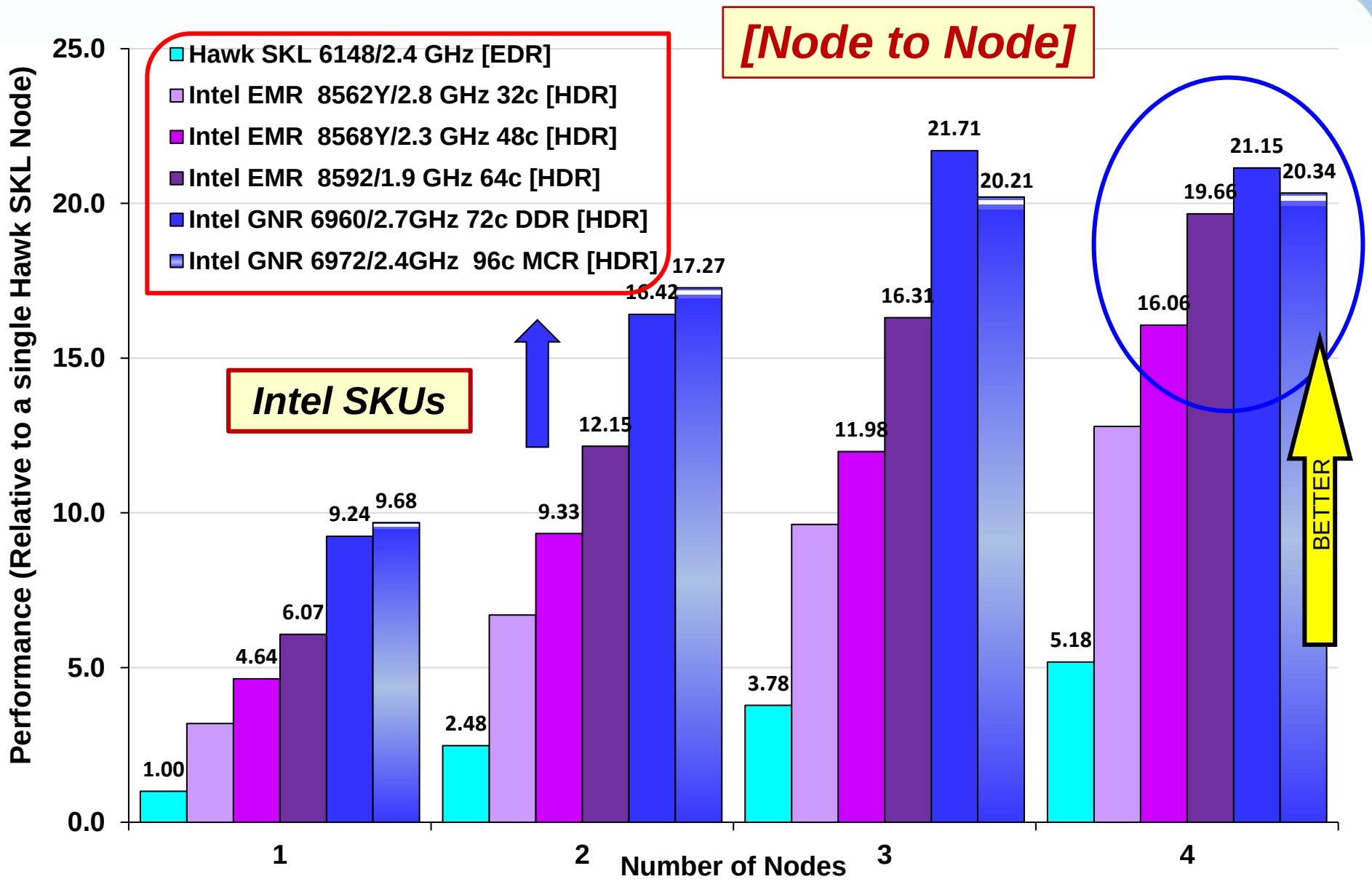


Example: The North Atlantic Ocean

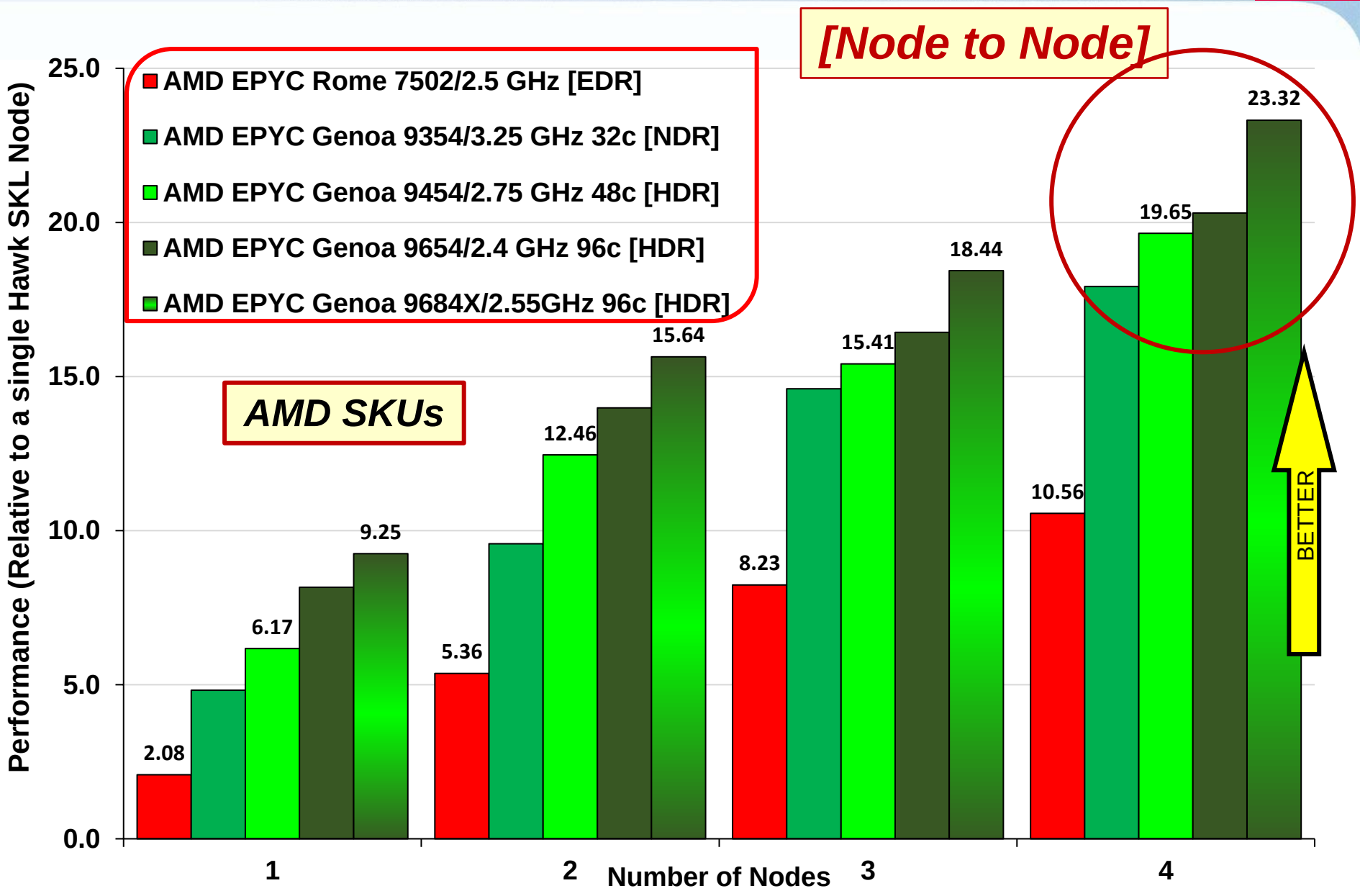
- 773 x 1236 horizontal grid points, multiplied by ‘k’ depth levels.
- Full horizontal domain split into 9 x 20 sub-domains.
- Each subdomain is handled by a separate core during parallel runs.
- MPI for handling communication between subdomains.
- Known memory B/W issues – avoid full node occupancy

***FVCOM** employs a similar approach to parallelism, albeit based upon an unstructured, triangular horizontal mesh.

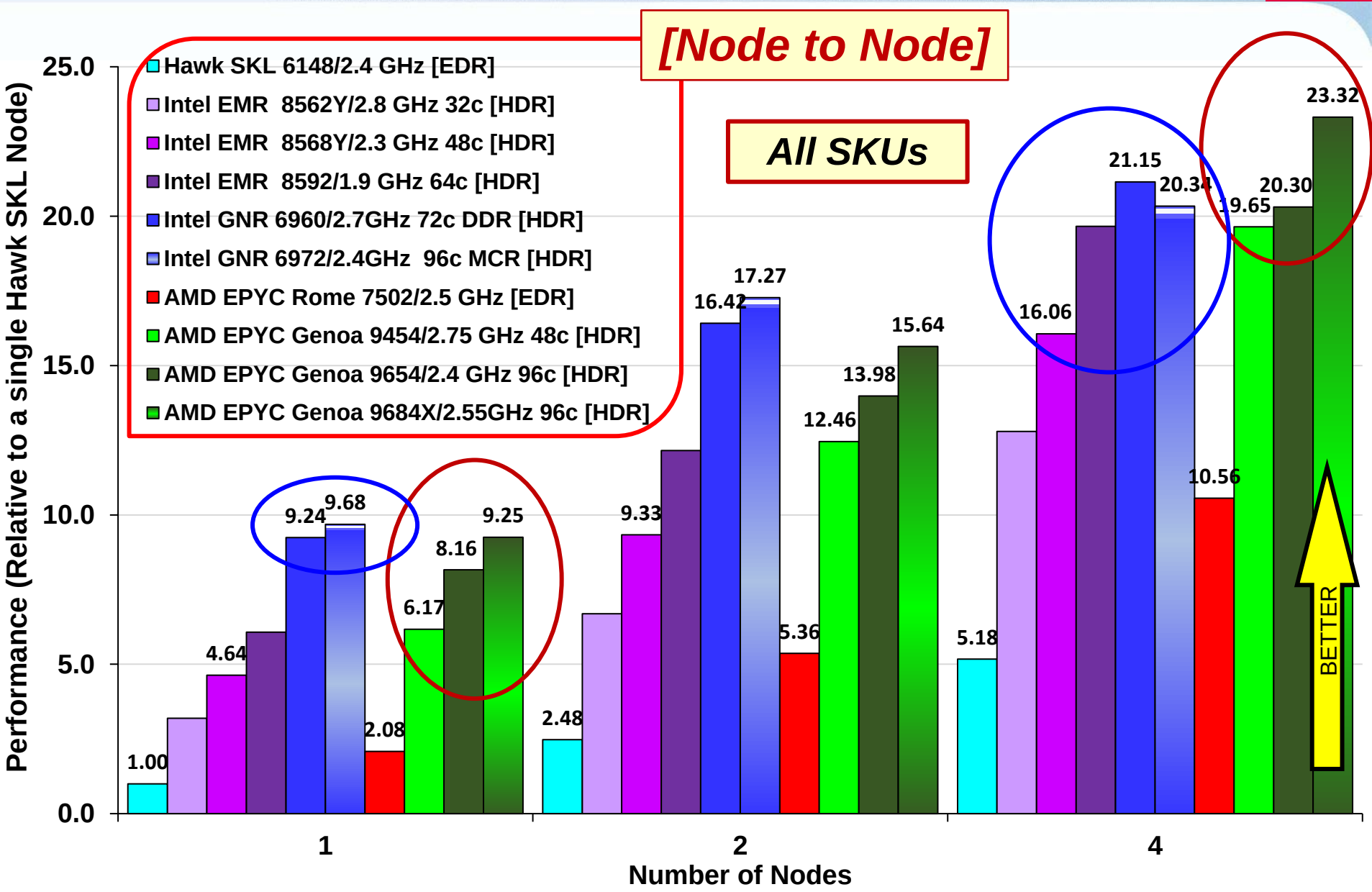
NEMO-FABM-ERSEM (AMM7) – Node Performance



NEMO-FABM-ERSEM (AMM7) – Node Performance



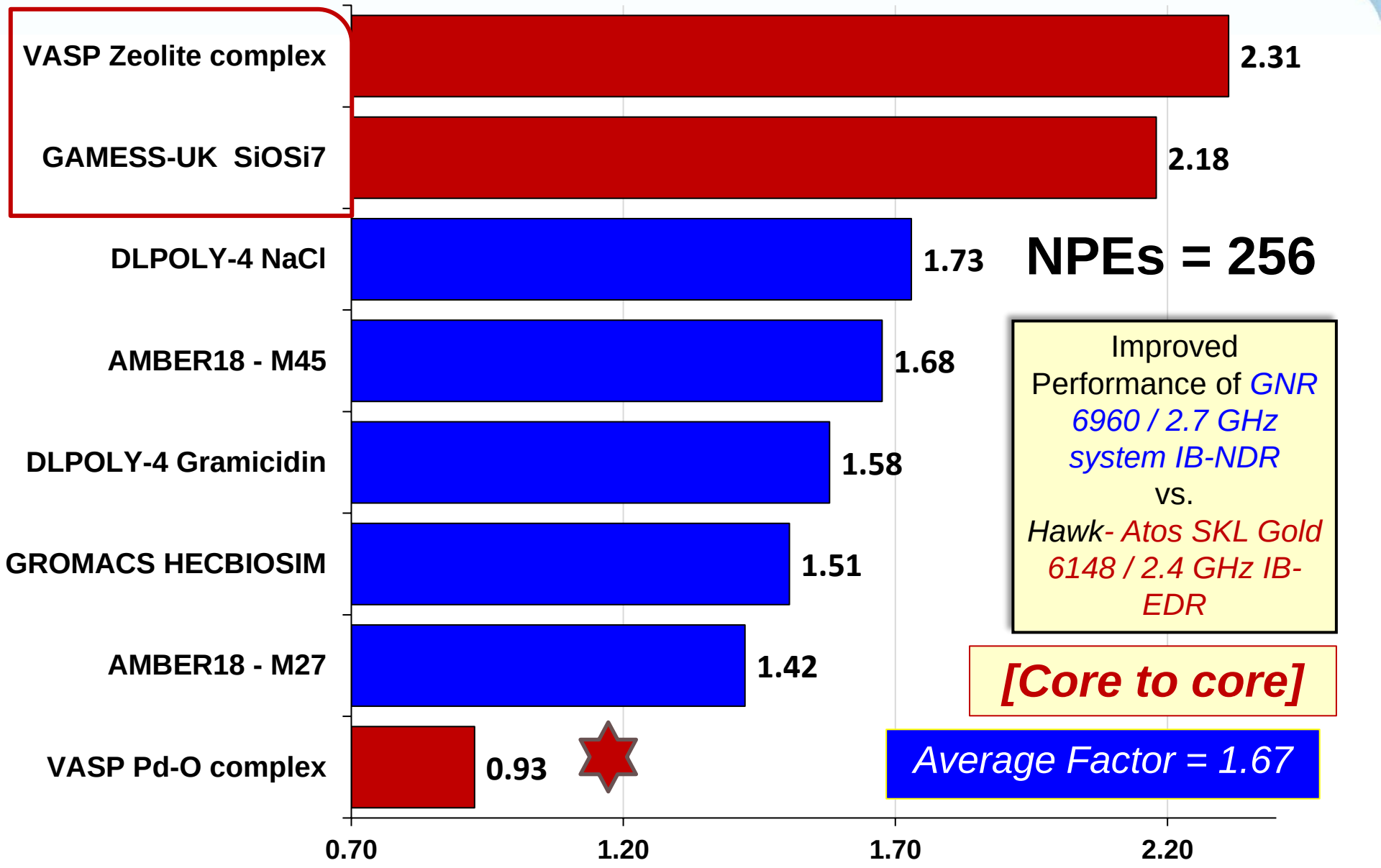
NEMO-FABM-ERSEM (AMM7) – Node Performance



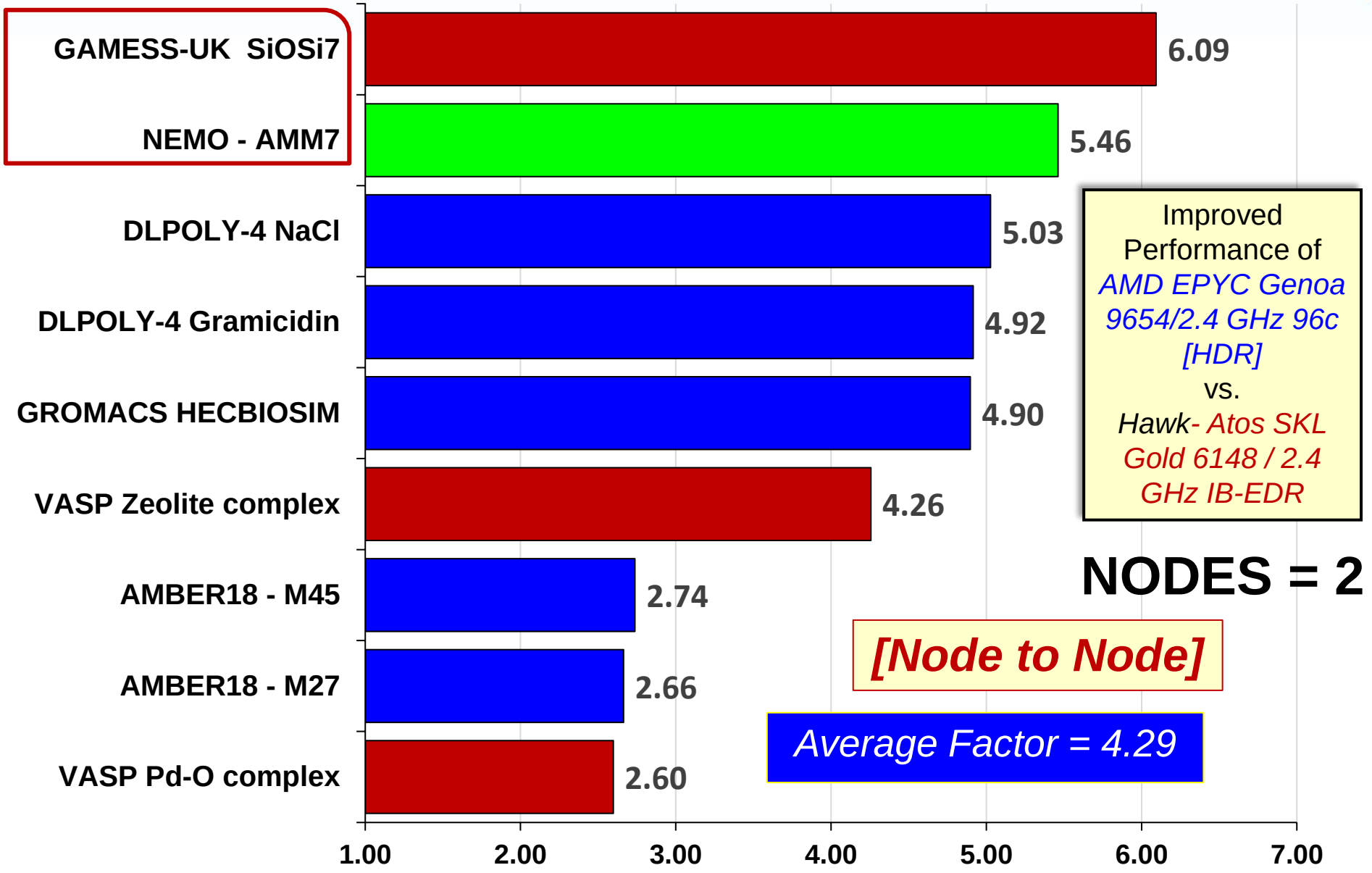
Performance of Computational Chemistry and Ocean Modelling Codes

Relative Performance as a Function of Processor Family

Granite Rapids 6960 2.7 GHz NDR vs. SKL 6148 2.4 GHz EDR



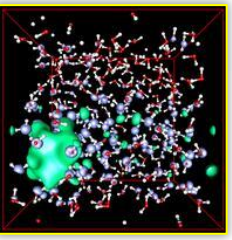
AMD Genoa 9654 2.4GHz NDR vs. SKL 6148 2.4 GHz EDR



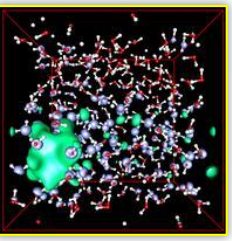
Target Codes and Data Sets – 256 PEs

2023

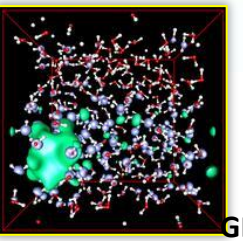
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD



Target Codes and Data Sets – 256 PEs

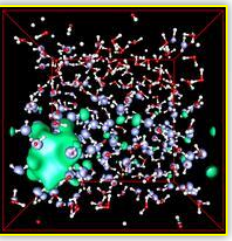


Target Codes and Data Sets – 2 Nodes

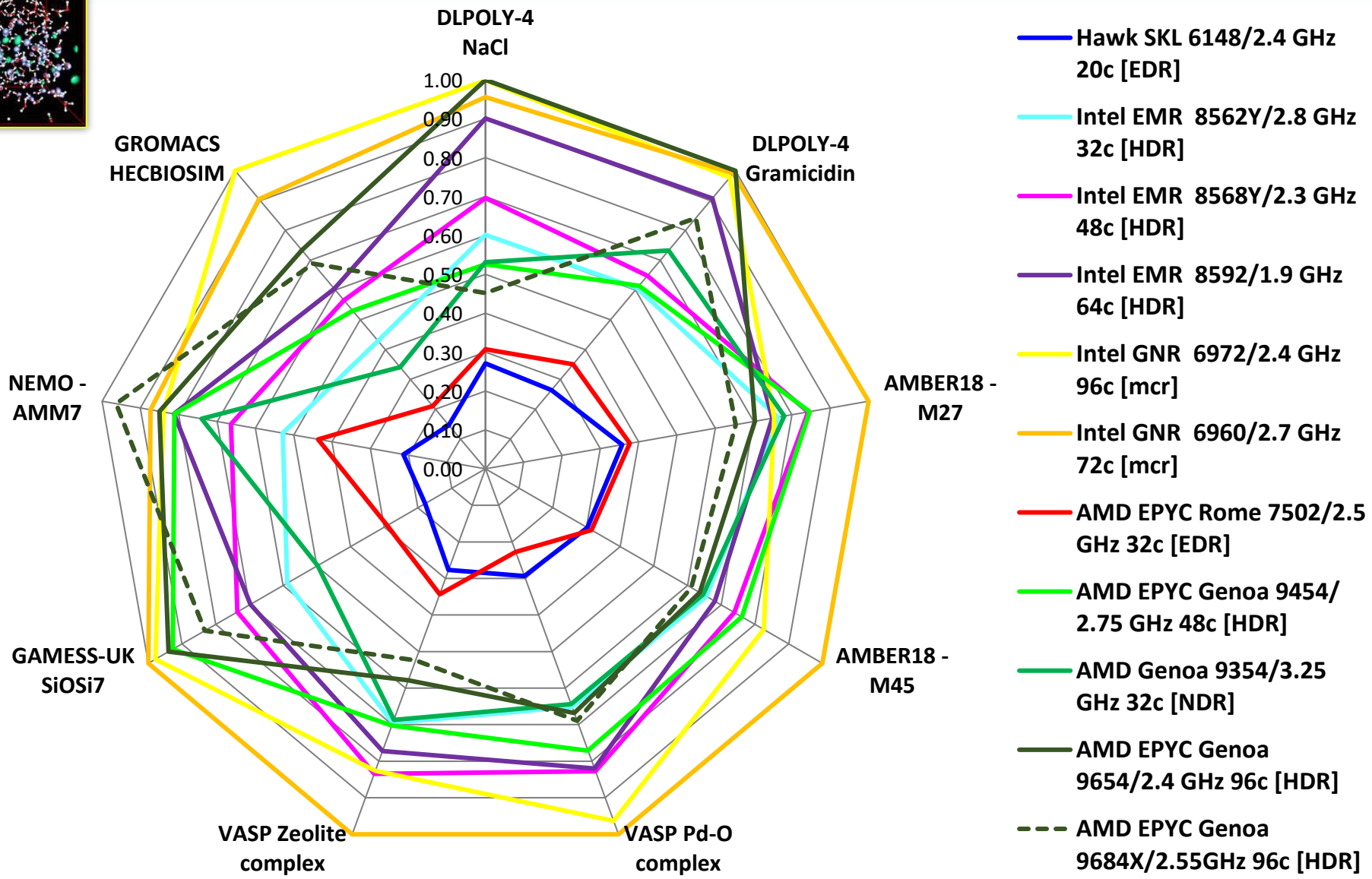


2 Nodes Performance [Applications]

Target Codes and Data Sets – 4 Nodes



4 Nodes Performance [Applications]



Conclusions – Core-to-Core Comparisons

- Remaining issue with the core-to-core performance of the **VASP 6.3 Pd-O simulations** on the Intel GNR systems. Under investigation … as indeed are the "small" differences between **MCR and DDR** simulations across all applications.
- ***Core-to-Core* comparisons** suggest that the **Intel GNR 6960/2.7 GHz 72c [MCR] outperforms** the baseline **SKL Gold 6148/2.4 GHz** (144 cores vs. 40 core nodes) by average improvements factors of **1.69** (128 cores) **and 1.77** (256 cores) **respectively** across all applications.
- ***Core-to-Core* comparisons** further show that the **Intel GNR 6960/2.7 GHz 72c [MCR] outperforms** in most cases the **AMD Genoa 9354 32c 3.25 GHz** [last year's leading system]**.** The exceptions being the AMBER M27 & DLPOLY4 Gramicidin simulations.
- **The Intel GNR 6960/2.7 GHz 72c [MCR] outperforms the GNR 6972/2.4 GHz 96c [MCR]** in most cases [factor ca. **1.1**)**,** the exceptions being the **VASP simulations.**

Conclusions – Core-to-Core Comparisons

- Comparing with the **previous EMR systems**, we find the **Intel GNR 6960/2.7 GHz 72c [MCR] outperforms the previous EMR SKUs** – the EMR **8562Y / 2.8GHz 32c**, **8568Y/2.3GHz 48c** and **8592/1.9GHz 64c** by average improvements factors on 256 cores of **1.01, 1.17 and 1.22** respectively.
- In general, major performance improvements when using the **HPC-X MPI** library on **both Intel and AMD clusters**.
- ***Core-to-Core* comparisons** show that with the exception of **DL_POLY** (NaCl), the **AMD Genoa 9354/3.25 GHz 32c** outperforms the **AMD EPYC Genoa 9654/2.4 GHz 96c** in all applications by a factor of **1.34**.
- Baselined in part across the **V100** NVIDIA GPU performance.

Conclusions – Node-to-Node Comparisons

- Decline in performance when moving from 2-node to 4-node results attributed to either (i) **limited I/O performance** on the **Endeavour cluster**. NEMO performance a clear indicator of this issue, and/or (ii) limited scalability of the chosen **application data sets**.
- Given superior core performance, a ***Node-to-Node comparison*** typical of the performance when running a workload shows the **Intel GNR 6960/2.7 GHz 72c [MCR]** SKUs delivering **far superior performance** compared to the **SKL Gold 6148** (144 cores vs. 40 cores). Average improvements factors of **5.4** (2-node) and **4.0** (4-nodes) across all applications.
- A ***Node-to-Node comparison*** shows the **Intel GNR 6972/2.4 GHz 96c [MCR]** SKUs delivering **far superior performance** compared to the SKL Gold 6148 (192 cores vs. 40 cores). Average improvements factors of **5.6** (2-node) and **3.9** (4-nodes) across all applications.
- The **Intel GNR 6972/2.4 GHz 96c [MCR]** outperforms its predecessor, the **Intel EMR 8592/1.9 GHz 64c,** by factors of **1.50** (2-node) and **1.19** (4-nodes) across all applications.

Conclusions – Node-to-Node Comparisons

- There's little difference in performance between the two GNR SKUs under investigation, with the average performance of the **GNR 6972/2.4 GHz 96c** somewhat greater on **2-nodes**, while the **GNR 6960/2.7 GHz 72c** **is faster on 4-nodes.**
- The **Intel GNR 6972/2.4 GHz 96c [MCR]** outperforms the **AMD EPYC Genoa 9654/2.4 GHz 96c,** by factors of **1.30** (2-node) and **1.19** (4-nodes) across all applications
- Candidate for the leading node-to-node system in last year's analysis, the **AMD EPYC Genoa 9354/3.25 GHz 32c** is now outperformed by factors of **2.09** (2-nodes) and **1.58** (4-nodes) by the **Intel GNR 6972/2.4 GHz 96c [MCR]**

Acknowledgements

- **Joseph Stanfield, Joshua Weage,** Benoit Lodej, Mark Mendez & Dave Coughlin for access to, and assistance with, the variety of AMD EPYC & Intel Xeon SKUs at the Dell HPC & AI Innovation Lab.

❖ **Bogdan Stefan Pop, Fouzhan, Hosseini, Gilles Civario, Thomas Spencer Jones, Adam Roe** and David Clark for access to and assistance with both the Emerald Rapids cluster at the Swindon Lab and Intel's Endeavour Cluster.

❖ **Erwin James and John Swinburne** for implementing the NETCDF and XIOS-5 libraries on the Endeavour cluster for testing NEMO.

❖ **Mathieu Gontier and Benjamin Pajot,** AMD HPC Center of Excellence, for access to, assistance with a variety of AMD Genoa partitions on the Munich **"muncl02"** cluster.

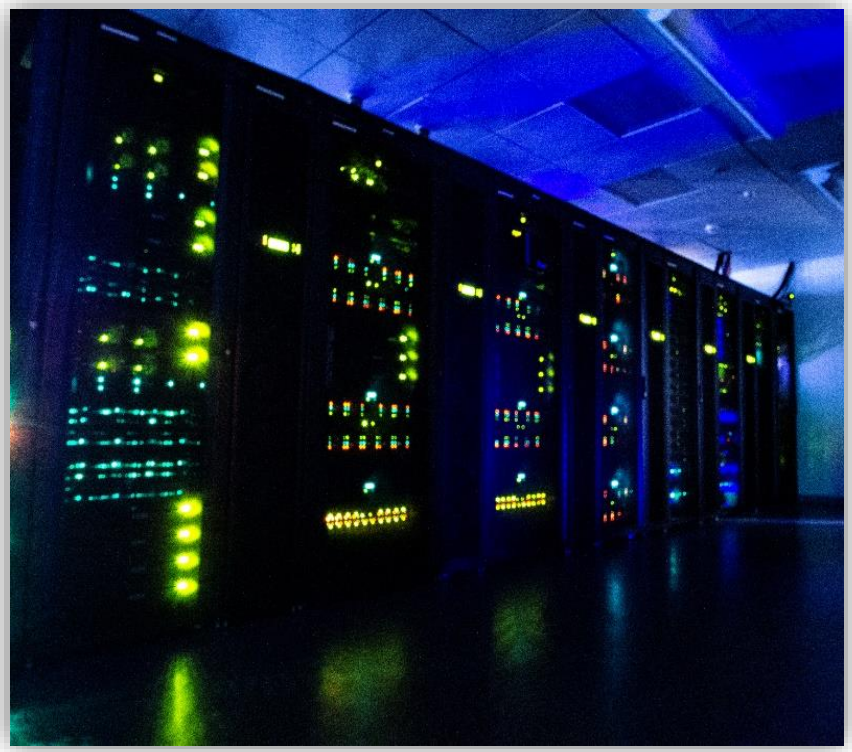
- **Jim Clark, Dale Partridge, Gary Holder and Jerry Blackford** at Plymouth Marine Laboratory for discussions on NEMO performance.

- **Colin Bannister,** for access to and assistance with "Ada", the 7008-core Genoa cluster, at Nottingham University.

Summary

- Focus on systems featuring **processors from Intel** (Emerald and Granite Rapids) and **AMD** (Genoa SKUs) with IB & Cornelis Networks.
  - ❖ Baseline clusters: Skylake (SKL) **Gold 6148/2.4 GHz** and **AMD EPYC Rome 7502 2.5Gz** cluster – “Hawk” – at Cardiff University.
  - ❖ **Three Intel Granite Rapids** clusters – the 72-core **6960** (2.7 GHz), 96-core **6972** (2.4 GHz) and 128-core **6980** (2.0 GHz).
  - ❖ **Three** Intel Xeon **Emerald Rapids** clusters, the 32-core **8562Y** (2.8 GHz), 48-core **8568Y** (2.3GHz) and 64-core **8592** (1.9 GHz).
  - ❖ **Four** AMD **Genoa clusters** featuring the 32-core **9354** (3.25 GHz), 48-core **9454** (2.85 GHz) & the 96-core **9654** (2.4GHz) and **9684X** (2.55GHz) SKUs.
- **Performance** of both synthetic & **end-user applications,** including molecular simulation (**DL_POLY, AMBER & GROMACS MD codes**), materials modelling (**VASP**) & electronic structure (**GAMESS-UK**) + NEMO ocean modelling code.
- **Scalability analysis** by **processing elements (cores)** and by **nodes** (ARM Performance Reports). Baselined against **V100** NVIDIA GPUs.
- **Pricing** – remains a key issue but lies outside the scope of this presentation.

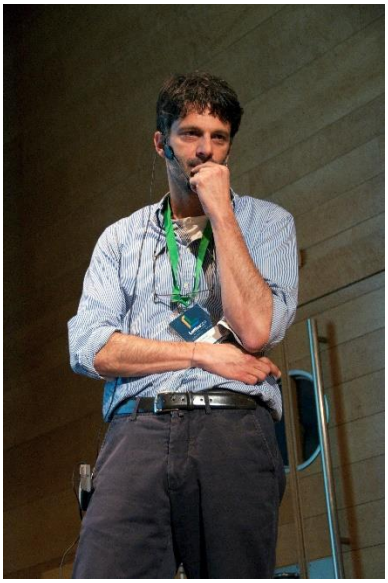
Any Questions?



Martyn Guest *GuestMF@Cardiff.ac.uk*

Jose Munoz Criollo *MunozCriolloJJ@cardiff.ac.uk*

CIUK 2024 Keynote Speaker



Luigi Del Debbio (The University of Edinburgh)

Visions of Computing

Abstract: Computing has witnessed the most exciting growth over the last few decades. As we enter the era of Exascale computers, Computing is having a major impact on our lives across multiple disciplines. 'Visions of Computing' tries to capture some of the ideas, challenges and projects that are currently associated with Computing. Based on my personal research experience, but also on my work in defining the current European landscape in High-Performance Computing, I will present an overview of the opportunities, the complexity and the challenges that we are facing.

Bio: Luigi Del Debbio is a physicist based in Edinburgh, known for their work in theoretical physics, particularly in the areas of quantum field theory and lattice gauge theories. Their research often explores the fundamental aspects of particle physics and the underlying structures of the universe. Del Debbio is actively involved in teaching and mentoring early-career scientists in the field. Their work emphasizes both theoretical insights and computational methods, bridging the gap between abstract theory and practical applications.

Del Debbio has been associated with various European initiatives aimed at enhancing high-performance computing (HPC) capabilities. Their efforts have facilitated international collaborations, enabling researchers across Europe to tackle challenging scientific problems more efficiently. Through their engagement in HPC, Del Debbio not only promotes scientific innovation but also advocates for the importance of computational techniques in modern research, emphasizing their role in addressing fundamental questions in physics and other disciplines.

Visions of Computing

Luigi Del Debbio

The University of Edinburgh

We strive to understand Nature and to solve problems in Science and Engineering through the fusion of models, algorithms, data… and of course computers. In this fusion, we value critical thinking, emotions and ethics. We define our work as ***Computing***.

We believe that Computing is essential for all Humankind.

Petros Koumoutsakos, Harvard

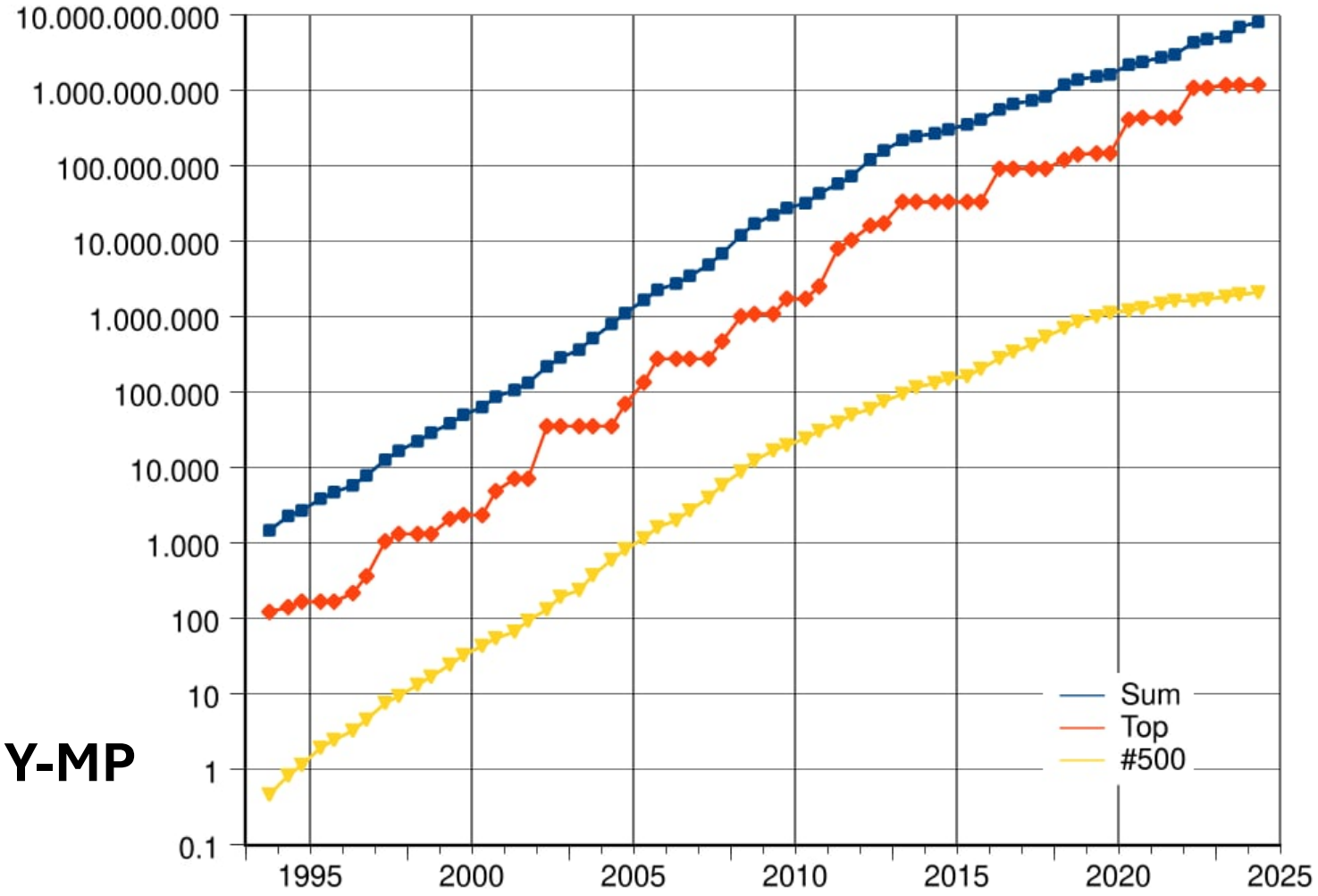
Why *Visions* (pl.)?

Cambridge Dictionary:

- an idea or mental image of something
- a beautiful and impressive sight
- the ability to imagine how a country, society, industry, etc. could develop in the future and to plan for this
- an experience in which you see things that do not exist physically, when your mind is affected powerfully by something such as deep religious thought or drugs or mental illness

Computing power

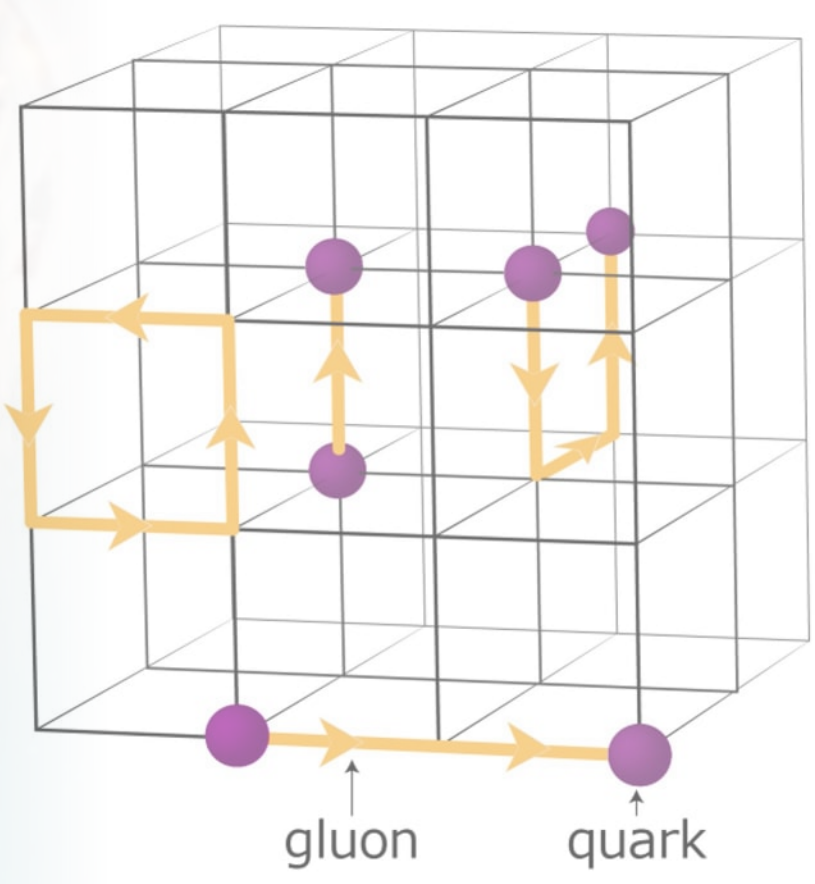
Y-MP



Computing: an on-going revolution

- My personal *visions* of the field, shaped by my own work
    - Research in high-energy theoretical physics
    - Chair of Resource Allocation in EURO-HPC
    - Scientific Advisory Committee of CSCS (Switzerland)
    - Chair of DiRAC Board (UK)
    - SIAB for UK Exascale project
- I hope I can provide a broad overview of the challenges

Lattice QCD

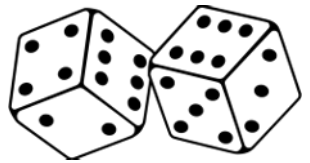


- First principles calculations of correlators

$$\langle O(\phi) \rangle = \frac{1}{Z} \int d\phi \, e^{-S[\phi]} \, O(\phi)$$

- Correlators are computed using Monte Carlo methods

1. Generation of configurations distributed as $e^{-S[\phi]}$
2. Computation of observables $O(\phi)$
3. Statistical analysis $\langle O(\phi) \rangle = \frac{1}{N} \sum_{n=1}^{N} O(\phi_n)$



- Computational bottlenecks are in 1. and 2.

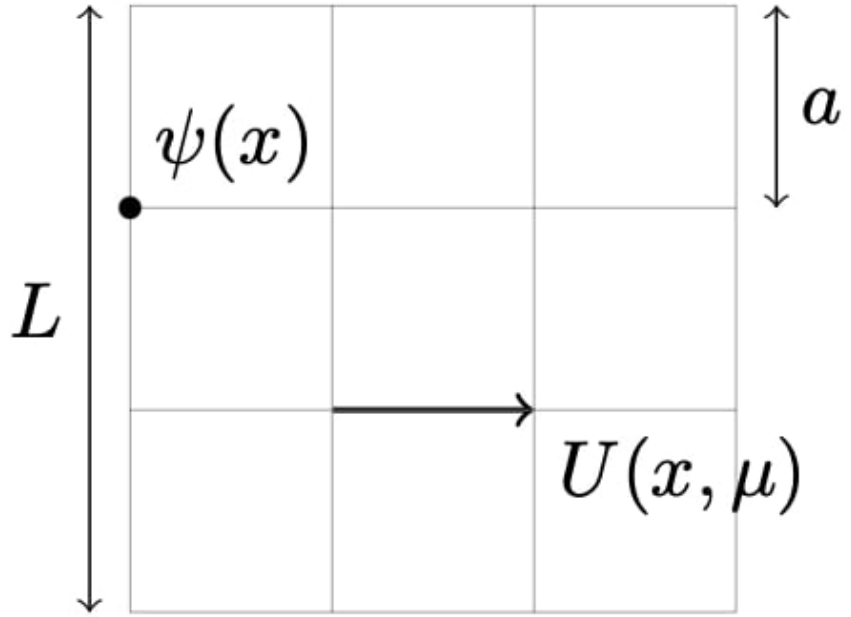
- Inversion of a large, ill-conditioned, sparse matrix $D(x,y)$

$$S[\phi] = \sum_{x,y} \phi(x) D(x,y) \phi(y)$$

D has dimension $\left(\left(\frac{L}{a}\right)^4 \times 4 \times 3 \times 2\right)^2$

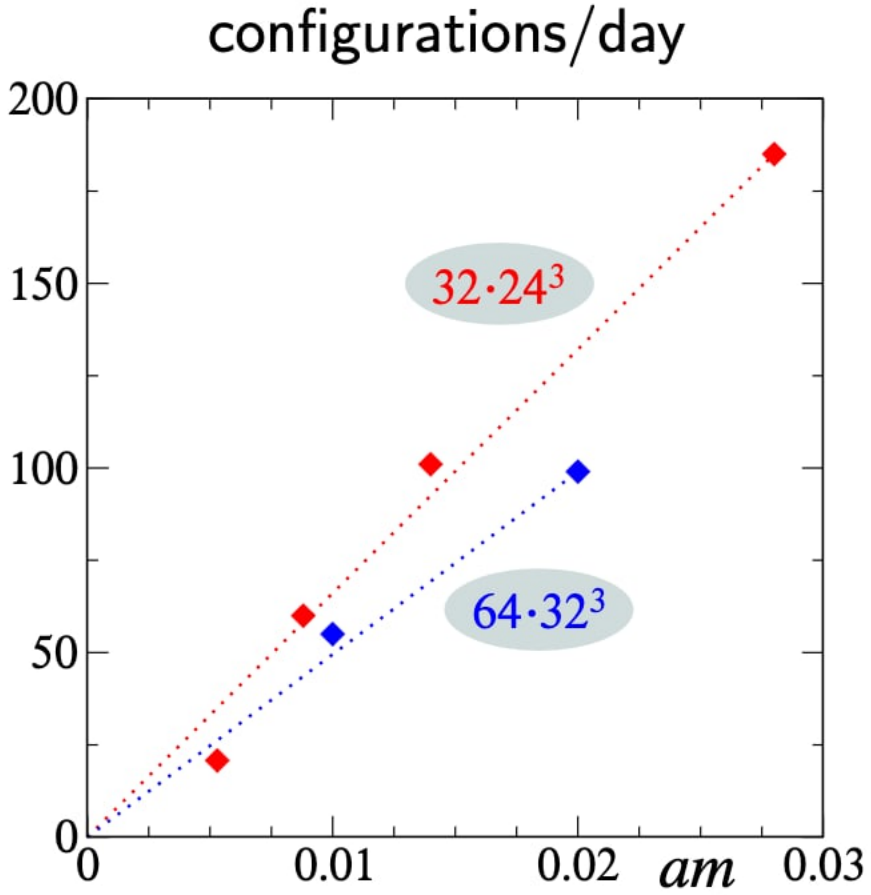
zero eigenvalue for $m \to 0$

scaling as $\mathrm{m}, \mathrm{a} \to 0$

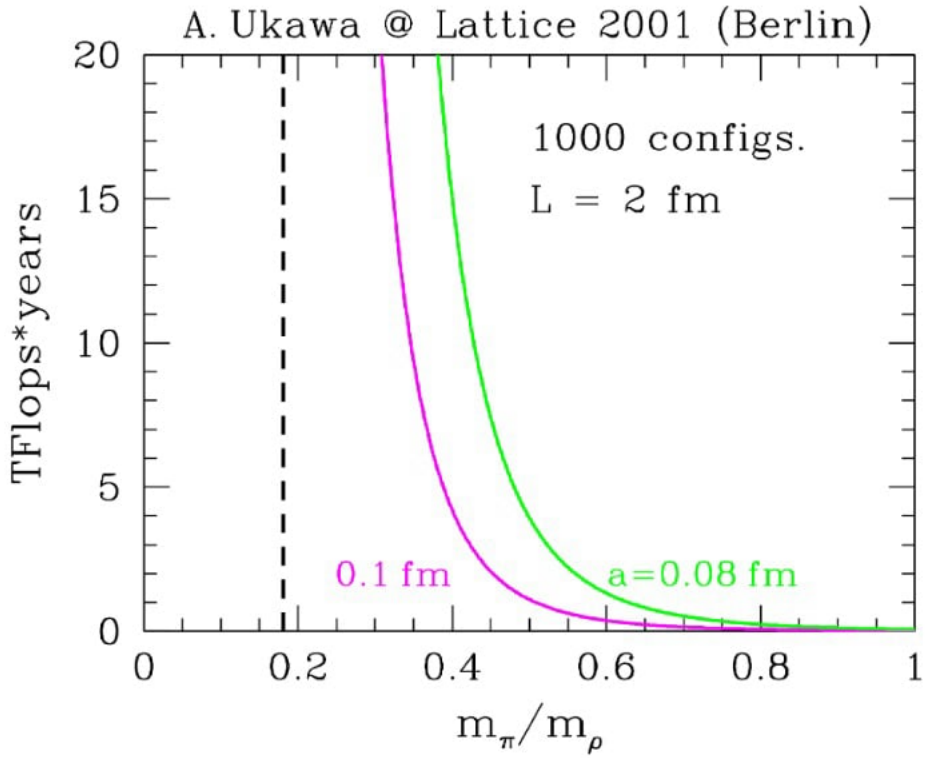


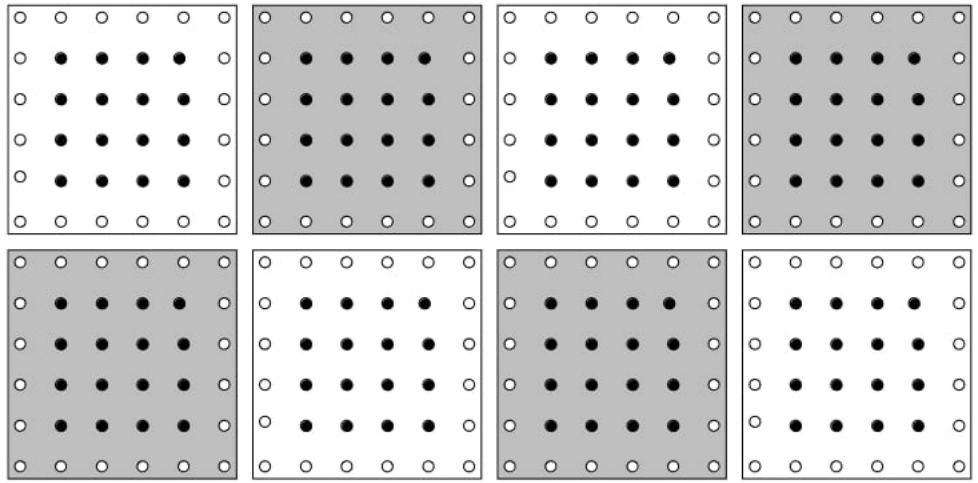
performance driven by algorithmic progress

- preconditioned Dirac operator



[ldd et al 07]





Scaling Lattice QCD beyond 100 GPUs

R. Babich *
Center for Computational Science
Boston University
Boston, MA 02215, USA
rbabich@bu.edu

M. A. Clark *
Harvard-Smithsonian Center for Astrophysics
60 Garden St
Cambridge, MA 02143, USA
mikec@seas.harvard.edu

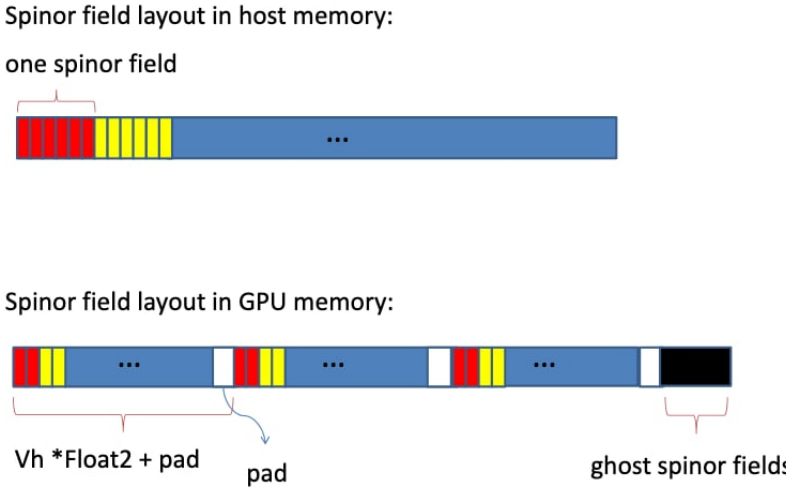
B. Joó *
Thomas Jefferson National Accelerator Facility
Newport News, VA 23606, USA
bjoo@jlab.org

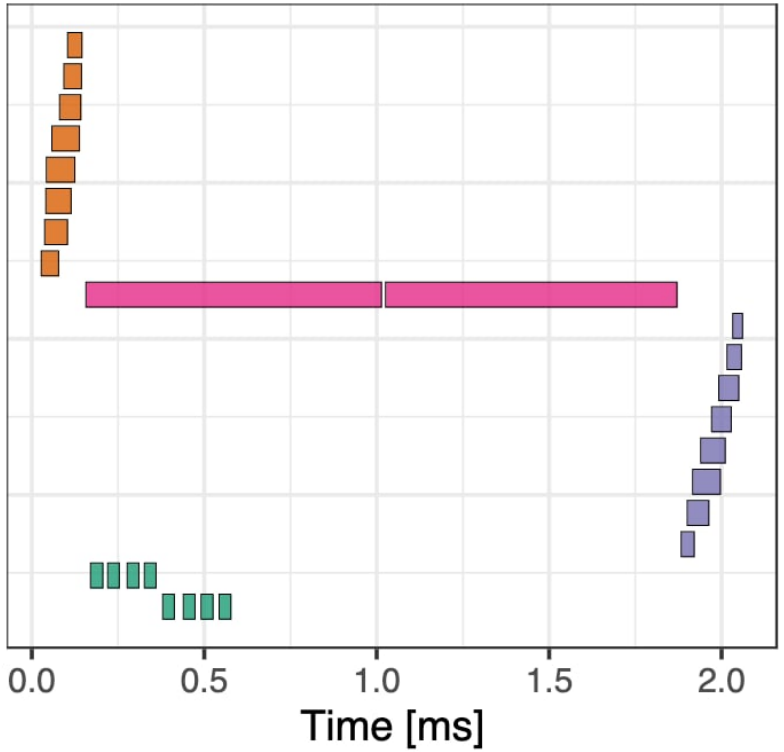
G. Shi *
National Center for Supercomputing Applications
University of Illinois
Urbana, IL 61801
gshi@ncsa.uiuc.edu

R. C. Brower
Center for Computational Science
Boston University
Boston, MA 02215, USA
brower@bu.edu

S. Gottlieb
Department of Physics
Indiana University
Bloomington, IN 47405, USA
sg@indiana.edu

2011





Scaling SU(2) to 1000 GPUs using HiRep

2024

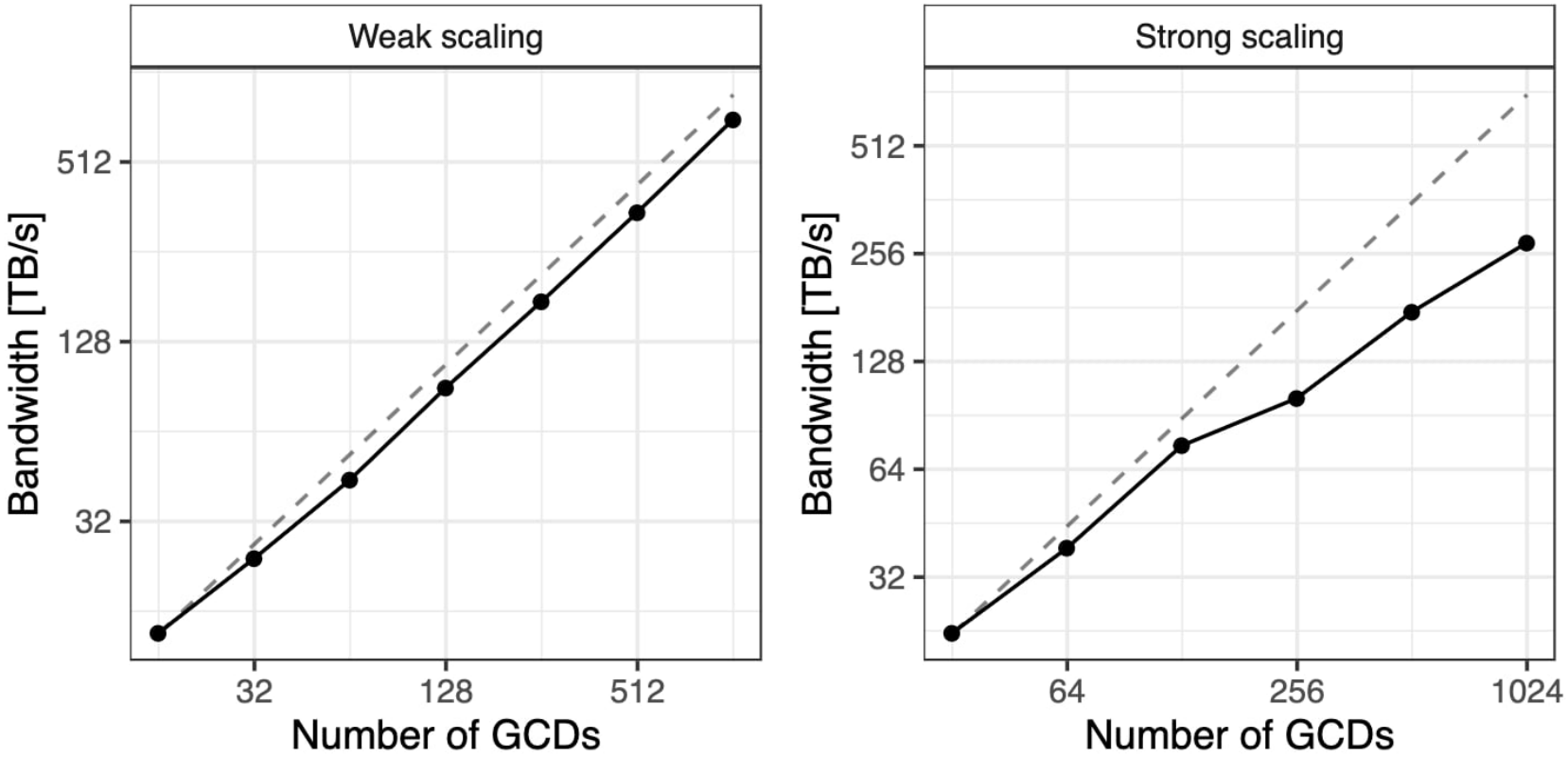
Sofie Martins,[a,b,*] **Erik Kjellgren,**[a] **Emiliano Molinaro,**[a] **Claudio Pica**[a,b] **and Antonio Rago**[a,b]

[a] *University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*

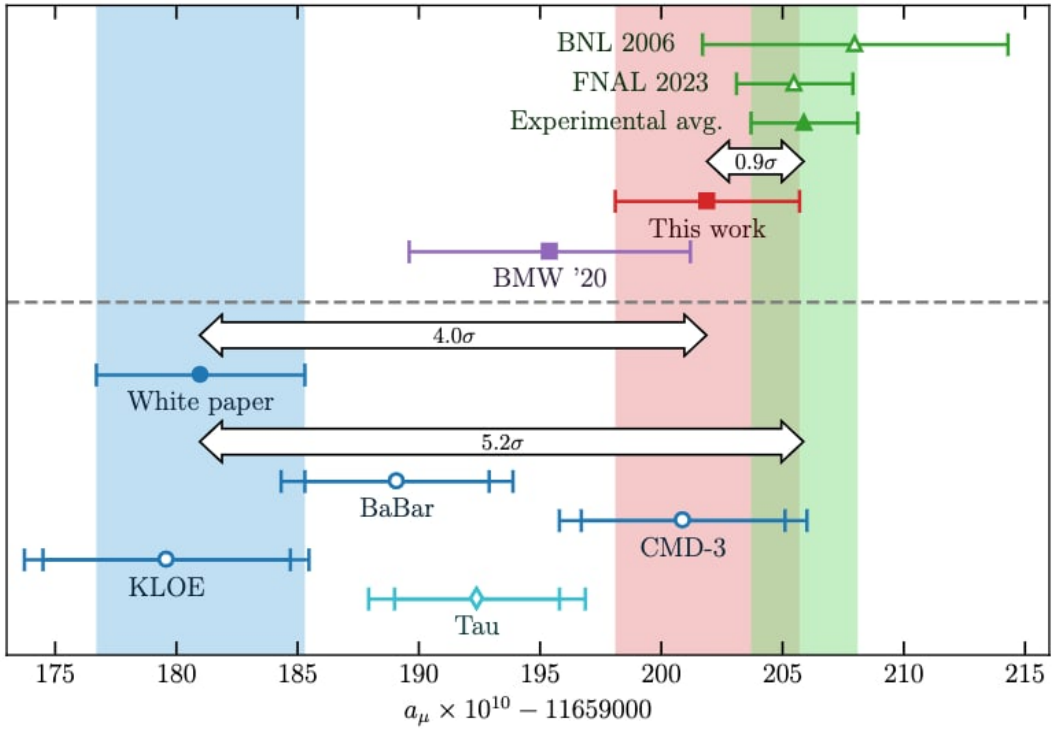
[b] *ħQTC, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*

E-mail: martinss@imada.sdu.dk

- weak and strong scaling



[Martins et al 2024]



[BMW Collaboration 2024]

Apr 7, 2021

SCIENCE

The New York Times

OUT THERE

A Tiny Particle's Wobble Could Upend the Known Laws of Physics

Experiments with particles known as muons suggest that there are forms of matter and energy vital to the nature and evolution of the cosmos that are not yet known to science.

Anomalous Magnetic Moment of the Muon

ML for lattice QCD

- Accelerate algorithms

$$d\phi\, e^{-S[\phi]} \longrightarrow d\chi\, e^{-S'[\chi]}$$

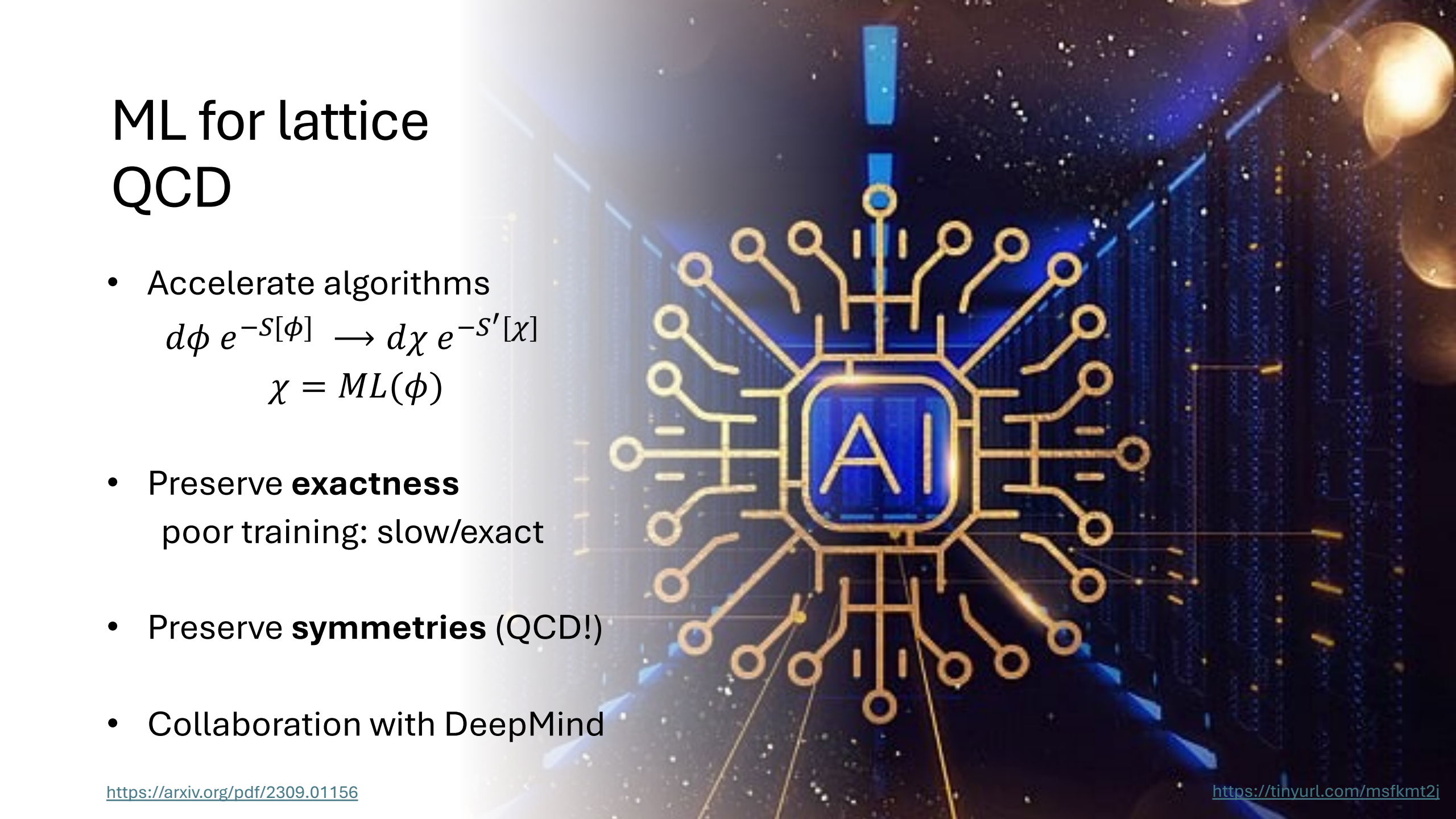
$$\chi = ML(\phi)$$

- Preserve **exactness**

  poor training: slow/exact

- Preserve **symmetries** (QCD!)

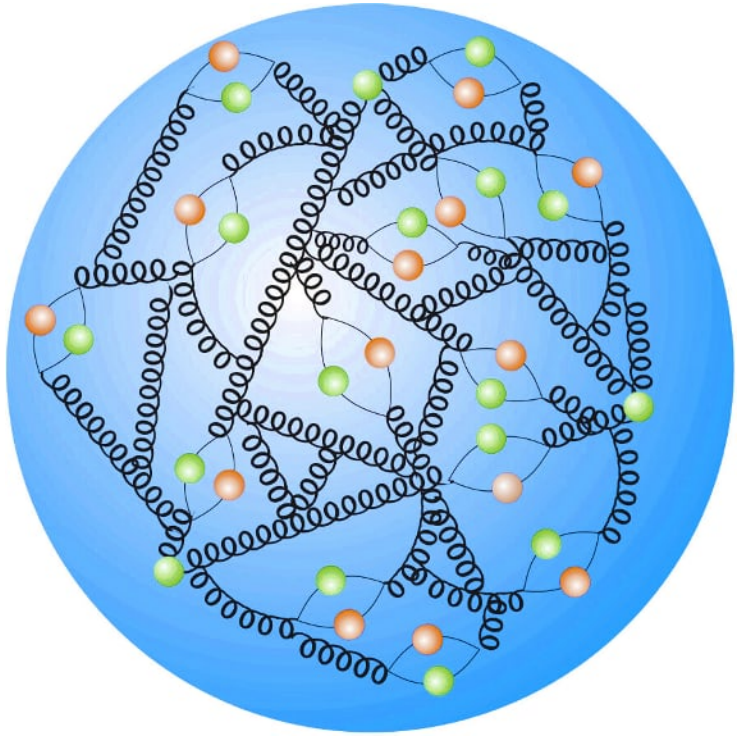
- Collaboration with DeepMind

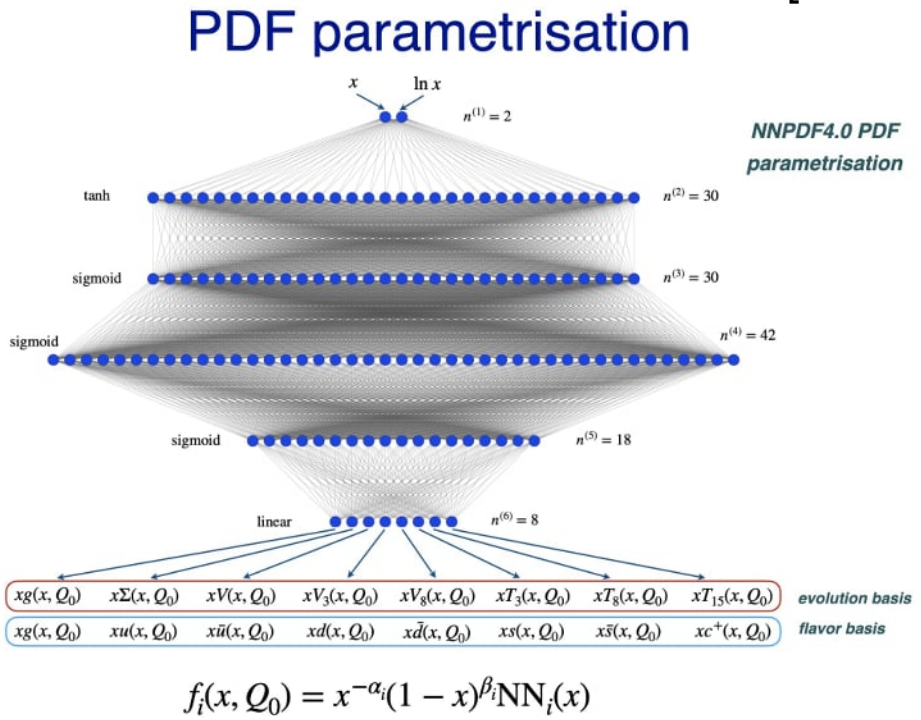


https://arxiv.org/pdf/2309.01156

https://tinyurl.com/msfkmt2j

[ldd et al. 02]





$$f_i(x,Q_0) = x^{-\alpha_i}(1-x)^{\beta_i}\mathrm{NN}_i(x)$$

ML and the proton structure

- ML used to parametrize the **parton** distributions
- one example of inverse problem
- synergies with other areas (ML, lattice QCD, geosciences)

First Global Gyrokinetic Simulations of Multi-Scale Plasma Turbulence in Tokamaks

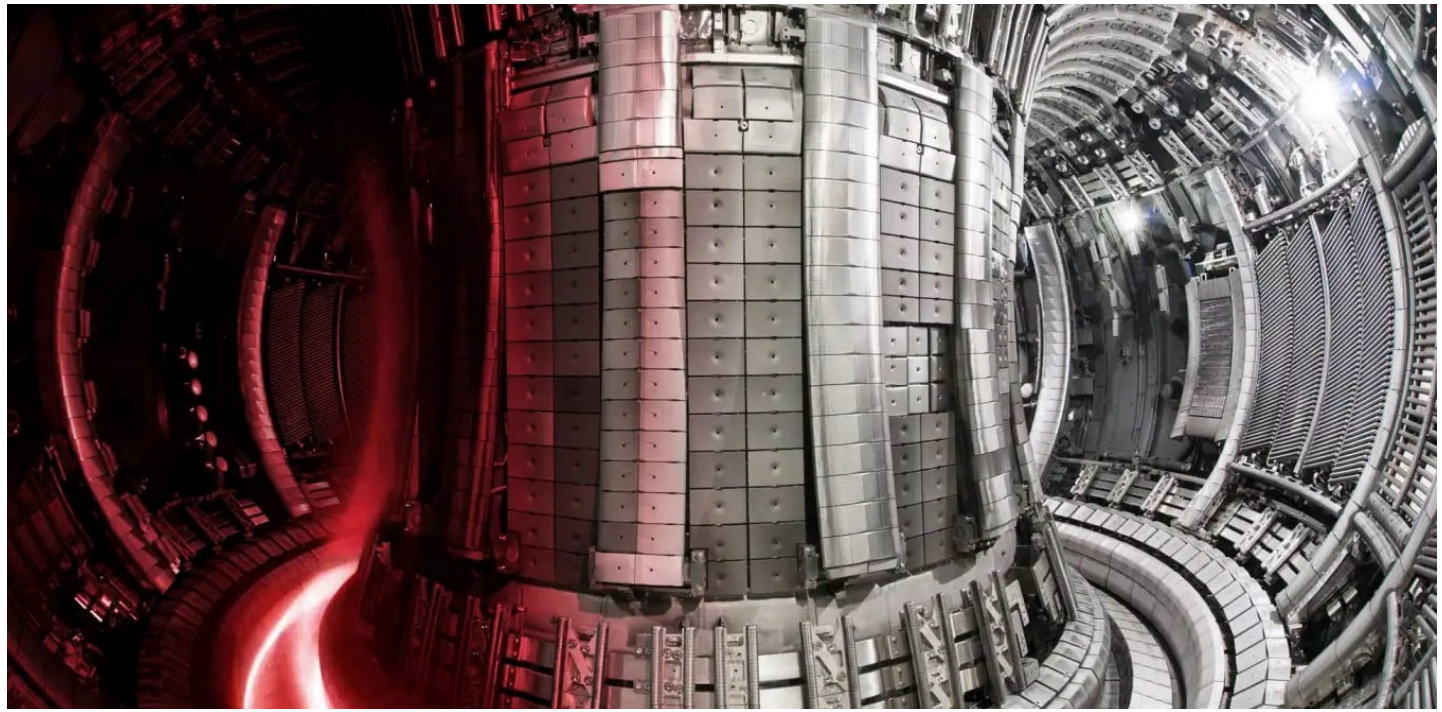
1,520,000
Awarded Resources (in node hours)

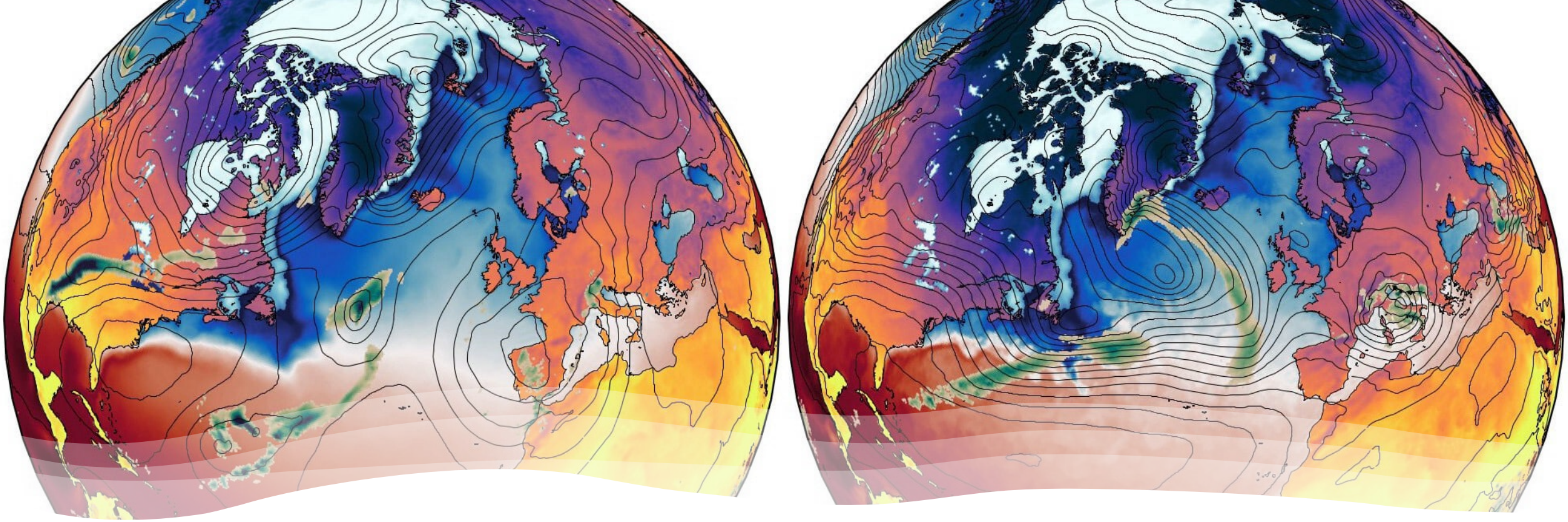
LUMI-G
System Partition

1 May 2024 - 30 April 2025
Allocation Period

Nuclear Fusion

- GENE codebase
- international collaboration
- GPU-enabled
- Potentially huge societal impact





EERIE project: https://eerie-project.eu/

Weather simulations

More EuroHPC projects

Cloud-Circulation Coupling in a Changing Climate (C5)

For the first time in the history of climate modelling, it is becoming possible to run multi-year climate simulations with a global atmospheric model that has a horizontal resolution of a few kilometers.

Forward-Modeling the First Billion Years

The birth of the first stars, galaxies and black holes heralded the beginning of the Cosmic Dawn (CD).

FemtoDose

FLASH radiotherapy (RT) consists in delivering a therapeutic dose over much shorter times than in conventional treatment protocols.

Large Scale and High Resolution CFD Simulations for Wave Energy Farms

The project aims to perform the first simulation of a wave energy farm using high-fidelity numerical modelling to anticipate the effects of arraying wave energy converters (WECs) within the same sea stretch.

The continuum limit of lattice QCD for high-precision tests of the standard model of particle physics from domain wall fermions

Exploring the standard model of particle physics and finding new physics beyond is in many cases limited by the lack of high-precision knowledge of low-energy QCD effects.

Universal Machine Learning Potential for Complex Metal Alloys

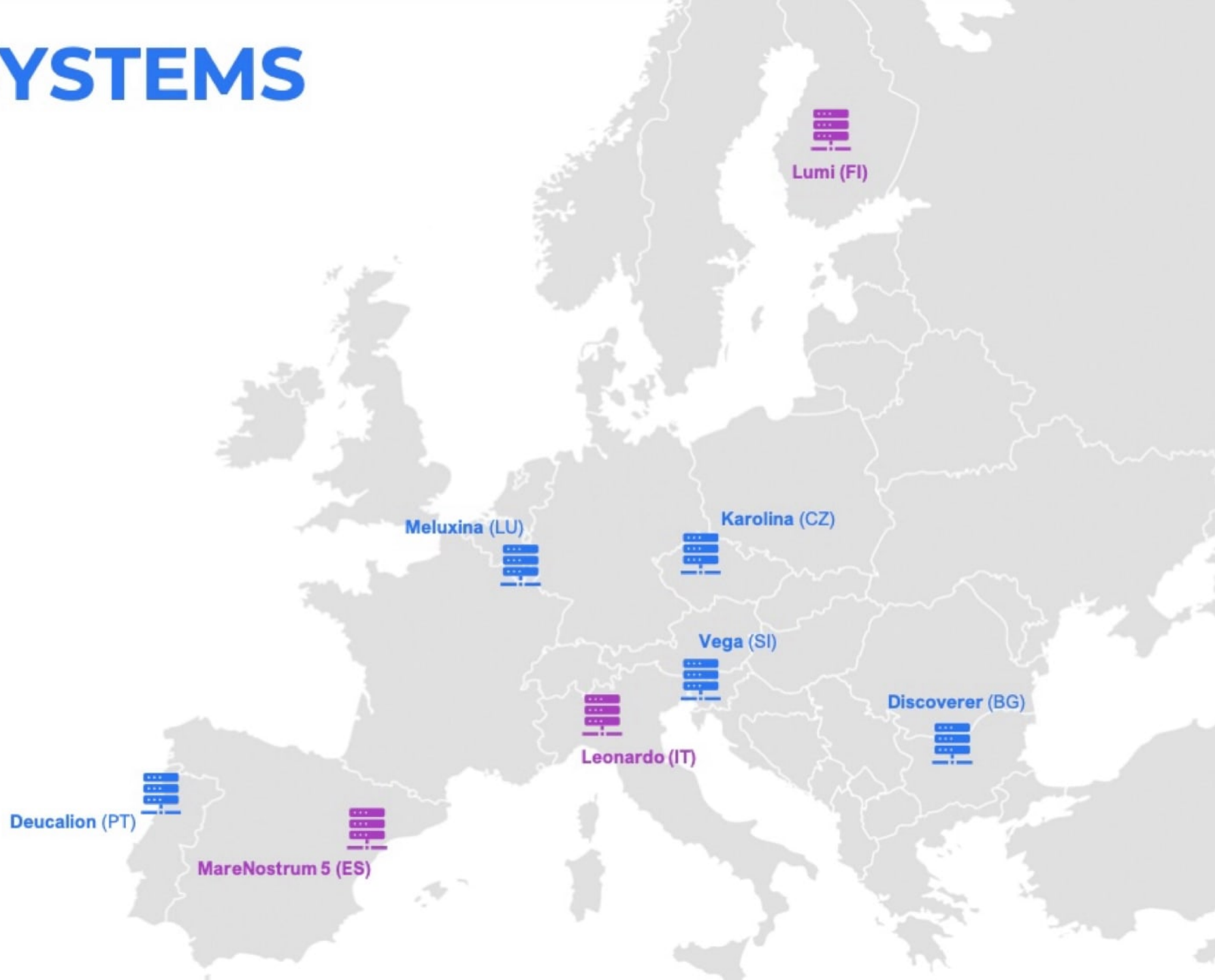
Atomistic modeling can provide mechanistic insights and improved design principles, but it is limited by the complexity of modern alloys that involve up to a dozen carefully tuned components.

Just a few examples...

EUROHPC SYSTEMS
2019 → 2023

PRE-EXASCALE

PETASCALE



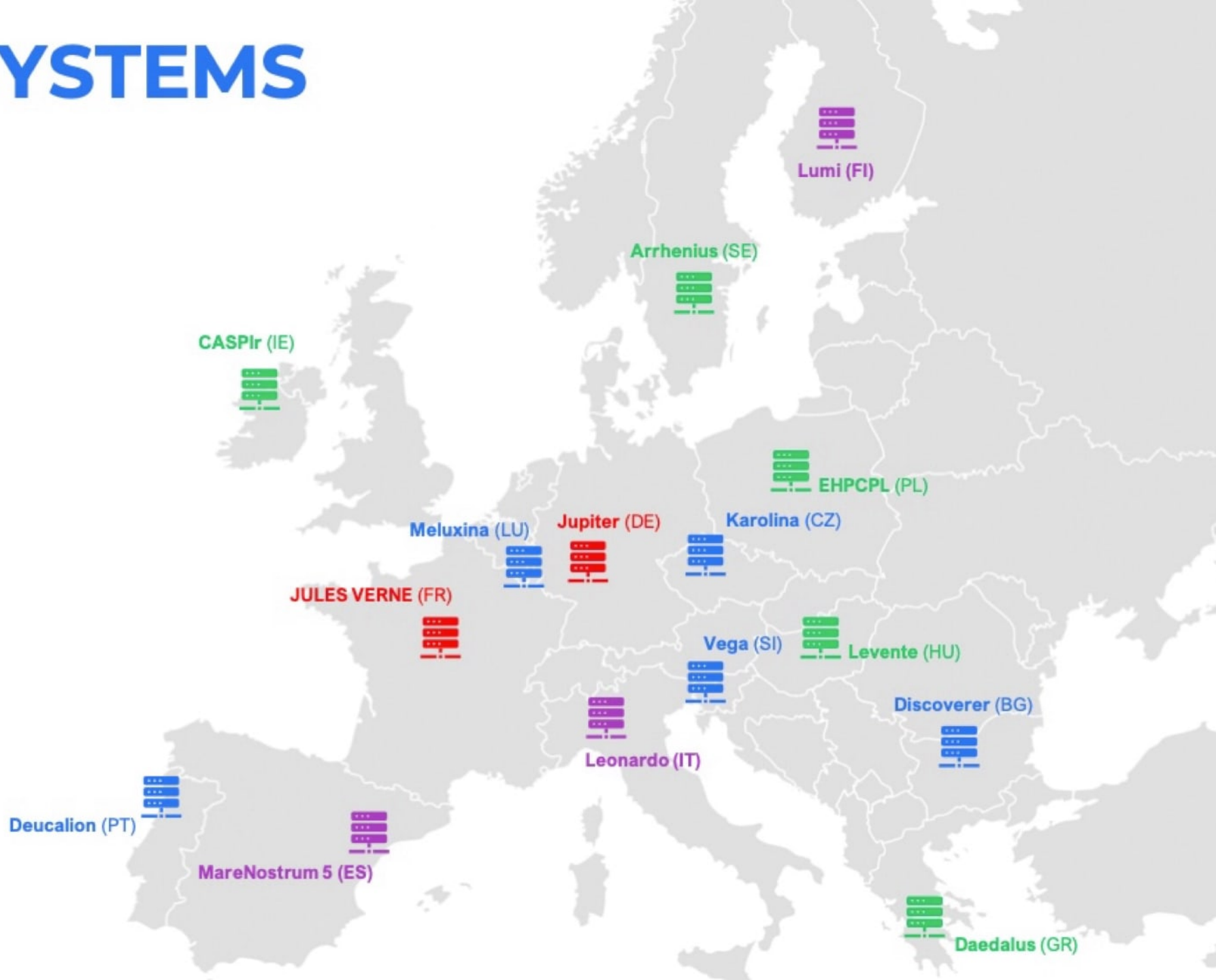
EUROHPC SYSTEMS
2024 → 2026

EXASCALE

PRE-EXASCALE

PETASCALE

MID-RANGE



EuroHPC systems | Pre-exascale

EuroHPC Joint Undertaking

LUMI (CSC)
Kayaani, Finland

Cray EX, Hewlett Packard Enterprise
#5 Top500 (Nov 2023): **379.7** PFlops (LUMI-G)

- 4.976 Nodes (2,928 GPU + 2,048 CPU)
- 11,712 GPUs (AMD MI250X)
- Slingshot Interconnect (200 Gb/s)
- 117 PB Storage (Lustre + Ceph)

AMD platform
- CPU: 64-core AMD EPYC™
- GPU: AMD Instinct™ (MI250X)

Leonardo (CINECA)
Bologna, Italy

Atos BullSequana XH2000
#6 Top500 (Nov 2023): **238.7** PFlops (BOOSTER)

- 4,992 Nodes (3,456 GPU + 1,536 CPU)
- 13,842 GPUs (Nvidia A100)
- Dual-rail Infiniband HDR (200 Gb/s)
- 110 PB Storage (Lustre)

Intel/NVidia platform
- CPU: Intel Sapphire Rapids (56-core), Intel Ice Lake (32-core)
- GPU: Nvidia custom Ampere (A100)

MareNostrum 5 (BSC)
Barcelona, Spain

Atos BullSequana XH3000 / Lenovo Think Systems
#8 Top500 (Nov 2023): **138.2** PFlops (GPU Partition)
#19 Top500 (Nov 2023): **40.1** PFlops (CPU Partition)

- 7,528 Nodes (1,120 GPU + 6,408 CPU)
- 4,480 GPUs (Nvidia H100)
- Quad-rail Infiniband NDR200
- 250 PB Storage (GPFS)

Intel/NVidia platform
- CPU: Intel Sapphire Rapids (56-core), Intel Sapphire Rapids (40-core),
- GPU: Nvidia Hopper (H100)

JUPITER

Project status
- Contract signed Oct 2023
- Kick off meeting 24 October.

Technical specifications
- **ARM** based system
- **First system** with indigenous CPU
- Modular Architecture
  - Booster: **24,000 NVidia GH200**
  - Cluster: **SiPearl Rhea 1**.

Implementation plan
- Test partitions to be available before summer -> **JUPITER Early Access Program (JUREP)** – Enable users to test the new architecture the earliest possible.
- Targeting Top500 ranking **Nov 2024**!

Datacenter status
- Site preparation ongoing
- Modular Data Center contract signed


DCD > Channels > Media > Live Events > Academy > Industry Awards

Europe's first exascale supercomputer will feature 24,000 Nvidia GH200 Superchips

Jupiter will also feature SiPearl's CPU


HPC WIRE

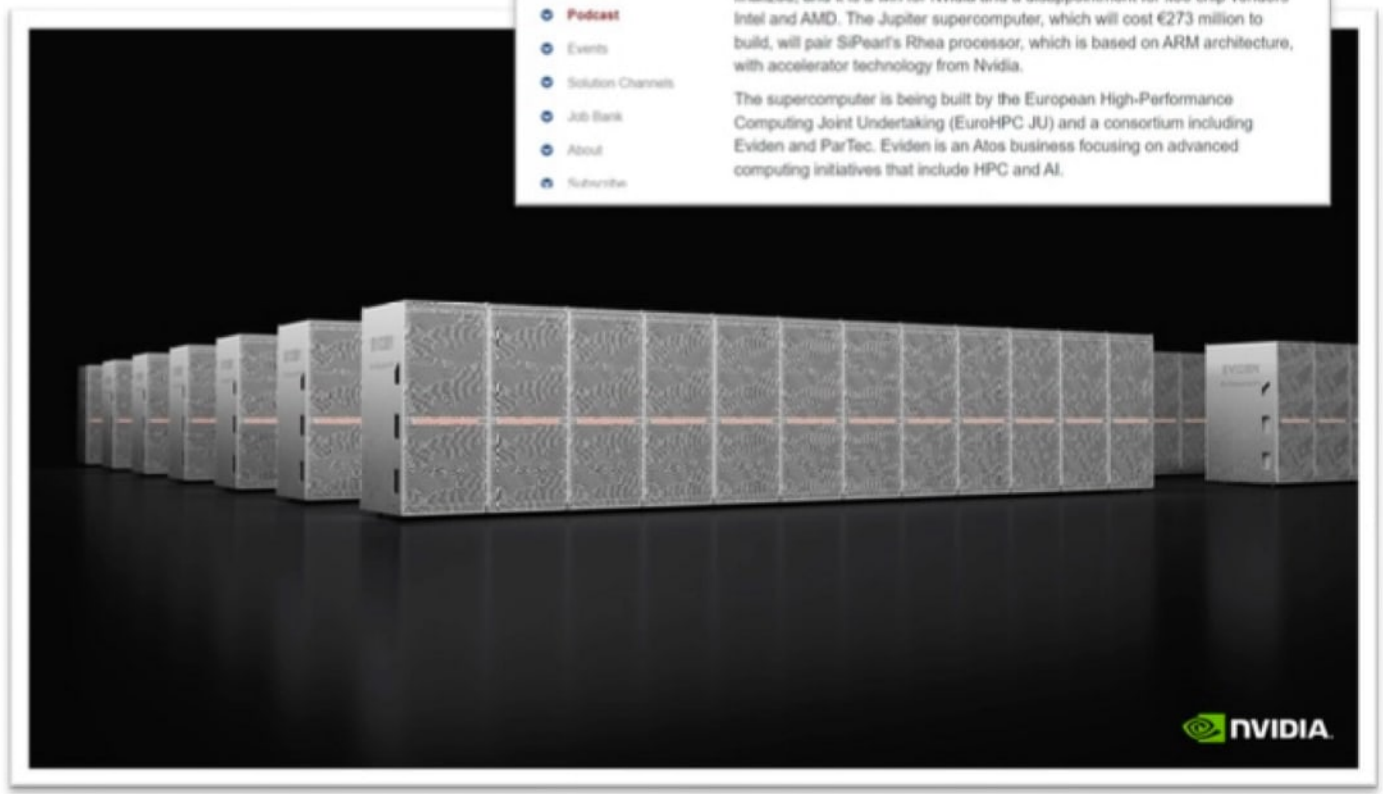
EU Grabs ARM for First ExaFLOP Supercomputer, x86 Misses Out

By Agam Shah

October 4, 2023

The configuration of Europe's first exascale supercomputer, Jupiter, has been finalized, and it is a win for Nvidia and a disappointment for x86 chip vendors Intel and AMD. The Jupiter supercomputer, which will cost €273 million to build, will pair SiPearl's Rhea processor, which is based on ARM architecture, with accelerator technology from Nvidia.

The supercomputer is being built by the European High-Performance Computing Joint Undertaking (EuroHPC JU) and a consortium including Eviden and ParTec. Eviden is an Atos business focusing on advanced computing initiatives that include HPC and AI.



… and ALPS

- GraceHopper nodes: ARM + H100

Computers | **Alps**

Alps

A Research Infrastructure for science with extreme-scale data and computing needs

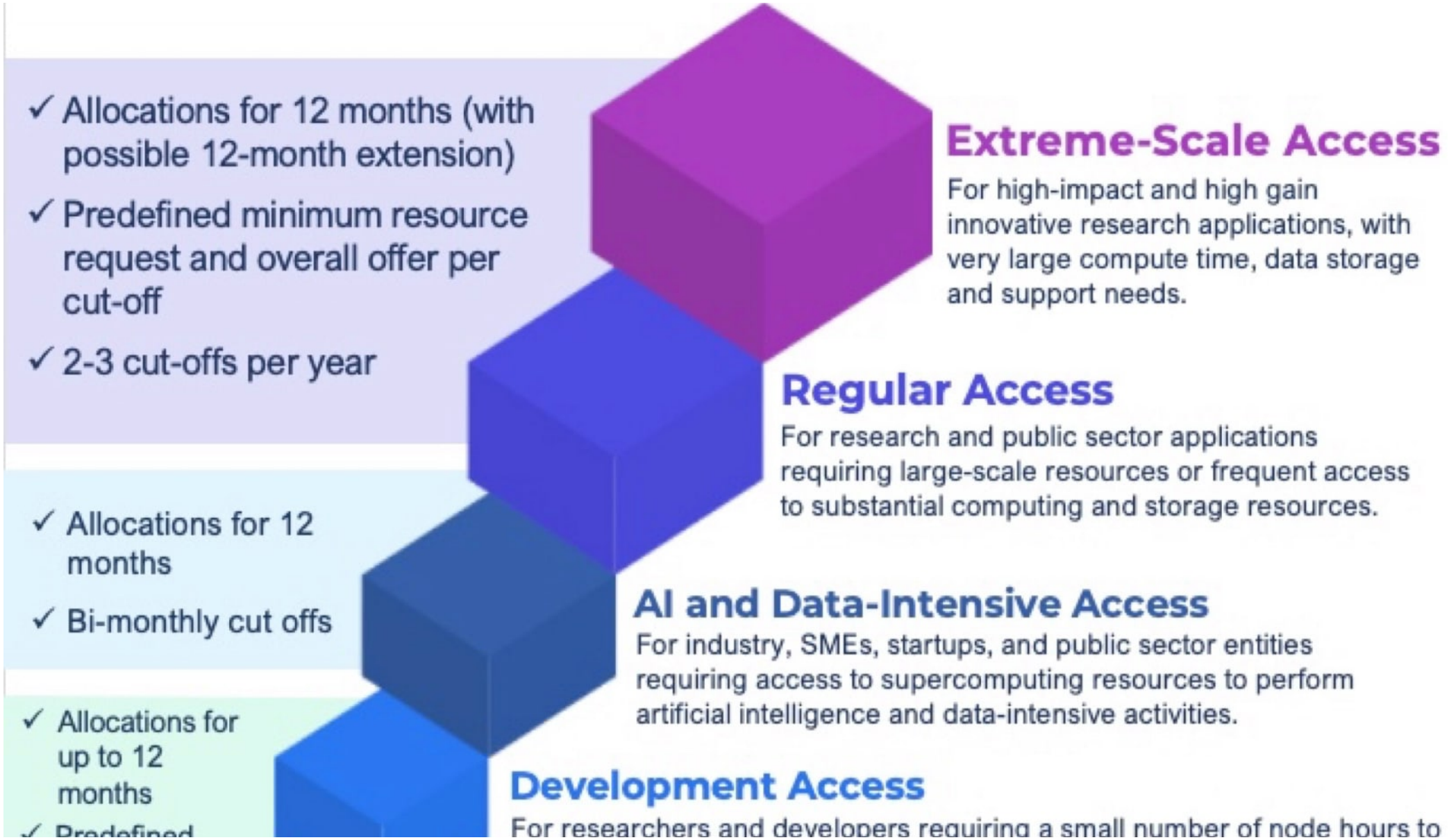
Alps is a general-purpose compute and data Research Infrastructure (RI) open to the broad community of researchers in Switzerland and the rest of the world. Alps will provide a high impact, challenging and innovative RI that will allow Switzerland to advance science and impact society.

Short Summary

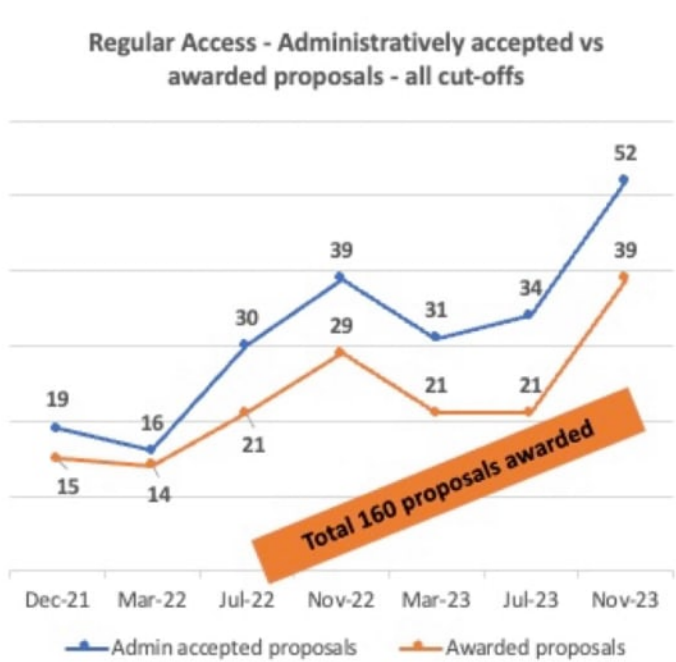
- Heterogeneous ecosystem
- Porting & efficiency are tricky
- … or rather they require a lot of work (and collaboration)!

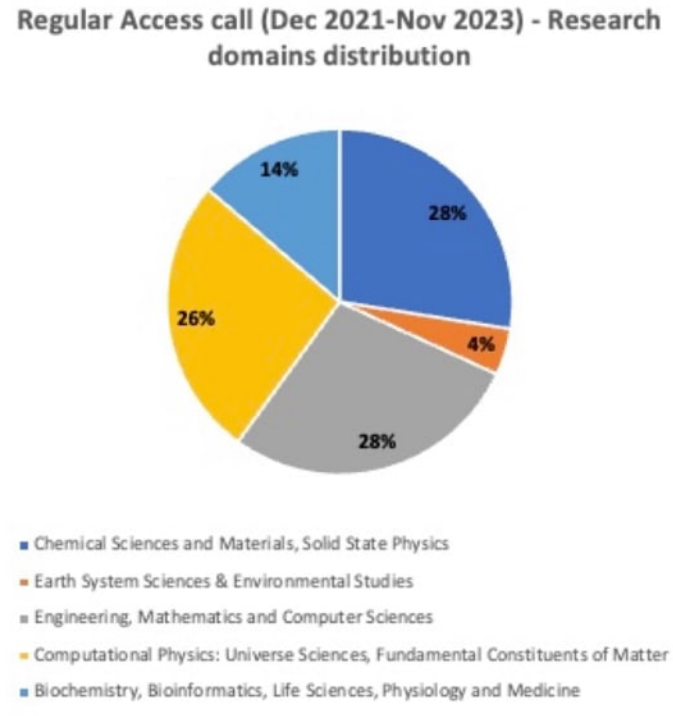
- Need adequate human resources
- Highly-specialised workforce
- Training/careers/retention

Who is using these machines? How do we access them?

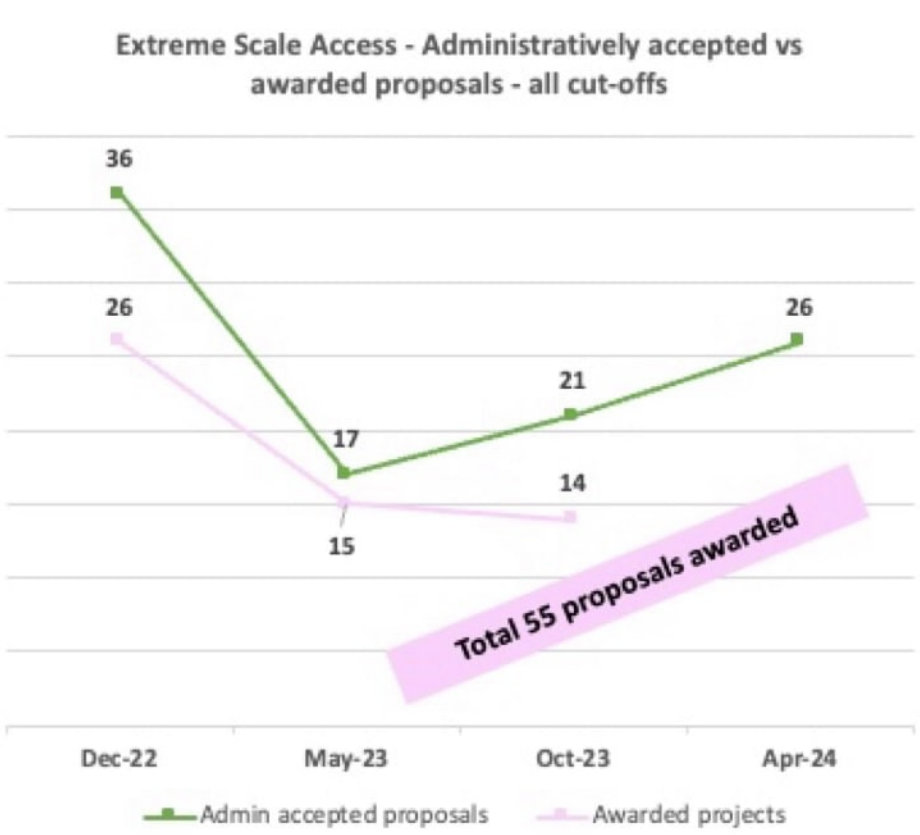

✓ Allocations for 12 months (with possible 12-month extension)
✓ Predefined minimum resource request and overall offer per cut-off
✓ 2-3 cut-offs per year
Extreme-Scale Access
For high-impact and high gain innovative research applications, with very large compute time, data storage and support needs.
✓ Allocations for 12 months
✓ Bi-monthly cut offs
Regular Access
For research and public sector applications requiring large-scale resources or frequent access to substantial computing and storage resources.
✓ Allocations for up to 12 months
AI and Data-Intensive Access
For industry, SMEs, startups, and public sector entities requiring access to supercomputing resources to perform artificial intelligence and data-intensive activities.
Development Access
For researchers and developers requiring a small number of node hours to

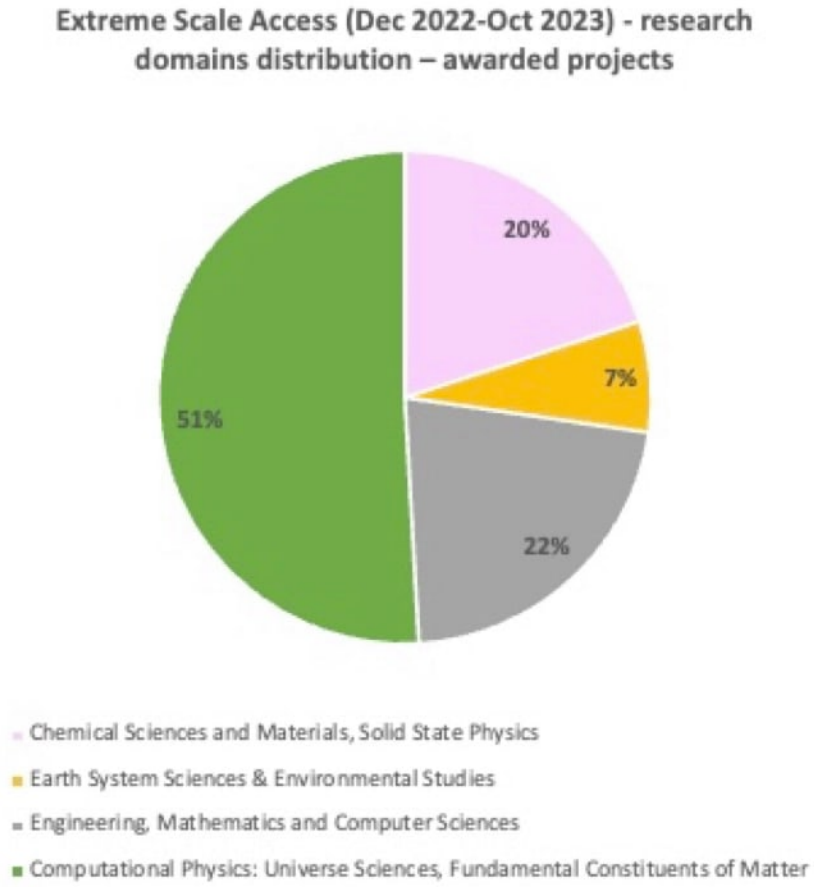
Regular Access statistics – proposal numbers



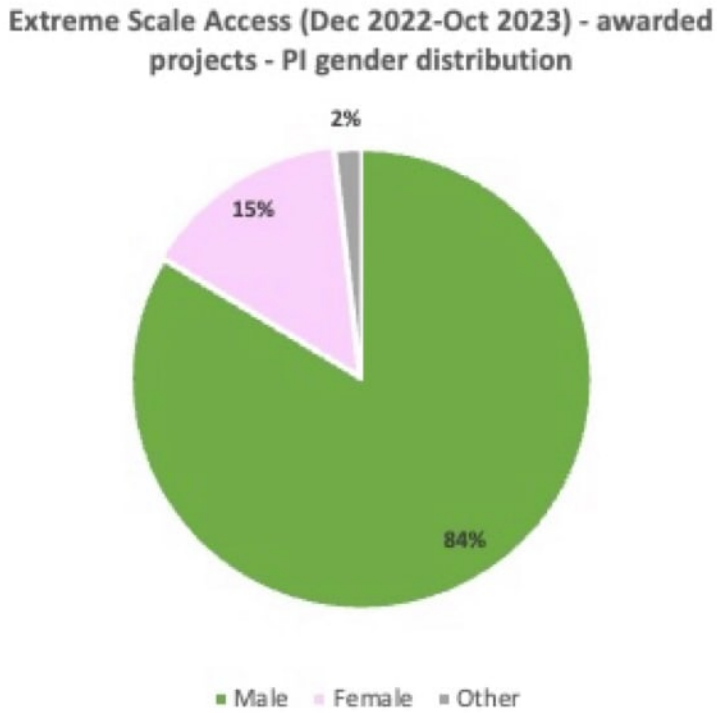


Extreme Scale Access statistics – proposal numbers





Extreme Scale Access statistics – gender distribution

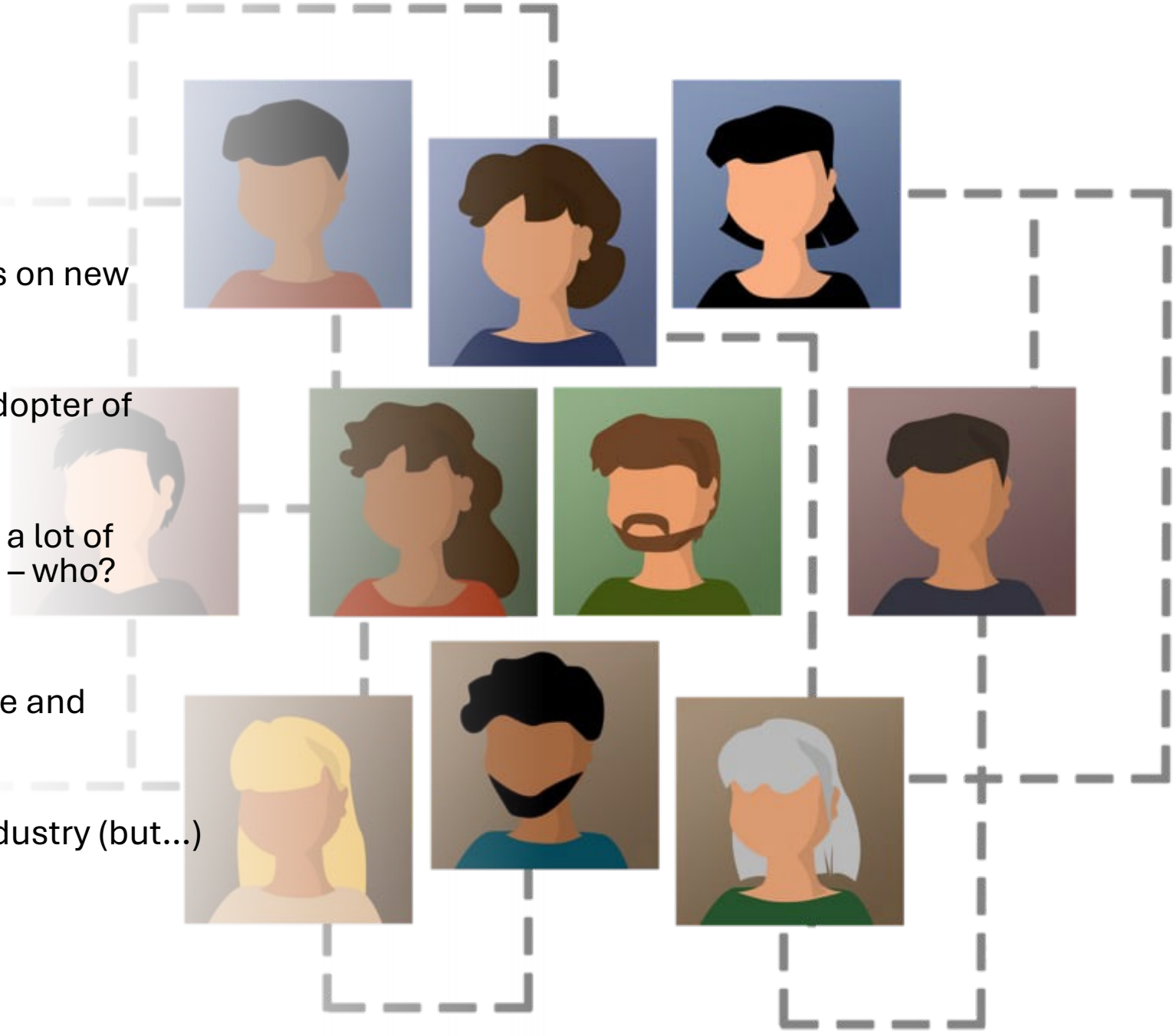


PI GENDER DISTRIBUTION – AWARDED PROJECTS		
GENDER	**Number of proposals**	**%**
Male	46	84%
Female	8	15%
Unspecified	1	2%
TOTAL	**55**	**100%**

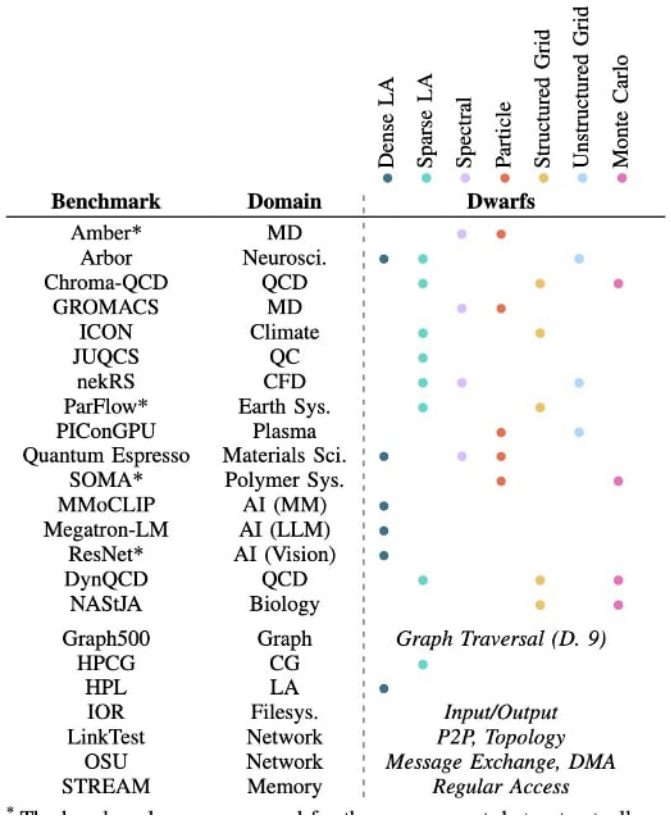
PI GENDER DISTRIBUTION – SUBMITTED PROPOSALS		
GENDER	**Number of proposals**	**%**
Male	64	80%
Female	15	19%
Unspecified	1	1%
TOTAL	**80**	**100%**

Human Resources

- Develop efficient algorithms on new architectures
- Lattice QCD was an early adopter of GPUs
- Requires (in my experience) a lot of domain-specific knowledge – who? where?
- Porting on multiple hardware and benchmarking, data (!)
- Engage in co-design with industry (but…)
- Career path for RSEs



procuring new systems



Benchmark	Domain	Dwarfs
Amber*	MD	Particle, Spectral
Arbor	Neurosci.	Dense LA, Sparse LA, Unstructured Grid
Chroma-QCD	QCD	Sparse LA, Structured Grid, Monte Carlo
GROMACS	MD	Particle, Spectral
ICON	Climate	Sparse LA, Structured Grid
JUQCS	QC	Sparse LA
nekRS	CFD	Sparse LA, Spectral, Unstructured Grid
ParFlow*	Earth Sys.	Sparse LA, Structured Grid
PIConGPU	Plasma	Particle, Unstructured Grid
Quantum Espresso	Materials Sci.	Dense LA, Spectral, Particle
SOMA*	Polymer Sys.	Particle, Monte Carlo
MMoCLIP	AI (MM)	Dense LA
Megatron-LM	AI (LLM)	Dense LA
ResNet*	AI (Vision)	Dense LA
DynQCD	QCD	Sparse LA, Structured Grid, Monte Carlo
NAStJA	Biology	Structured Grid, Monte Carlo
Graph500	Graph	*Graph Traversal (D. 9)*
HPCG	CG	Sparse LA
HPL	LA	Dense LA
IOR	Filesys.	*Input/Output*
LinkTest	Network	*P2P, Topology*
OSU	Network	*Message Exchange, DMA*
STREAM	Memory	*Regular Access*

* The

Aug 2024

Application-Driven Exascale: The JUPITER Benchmark Suite

Andreas Herten, Sebastian Achilles, Damian Alvarez, Jayesh Badwaik, Eric Behle, Mathis Bode, Thomas Breuer, Daniel Caviedes-Voullième, Mehdi Cherti, Adel Dabah, Salem El Sayed, Wolfgang Frings, Ana Gonzalez-Nicolas, Eric B. Gregory, Kaveh Haghighi Mood, Thorsten Hater, Jenia Jitsev, Chelsea Maria John, Jan H. Meinke, Catrin I. Meyer, Pavel Mezentsev, Jan-Oliver Mirus, Stepan Nassyr, Carolin Penke, Manoel Römmer, Ujjwal Sinha, Benedikt von St. Vieth, Olaf Stein, Estela Suarez, Dennis Willsch, Ilya Zhukov
Jülich Supercomputing Centre
Forschungszentrum Jülich
Jülich, Germany

***Abstract*—Benchmarks are essential in the design of modern** This paper introduces the *JUPITER Benchmark Suite*, a

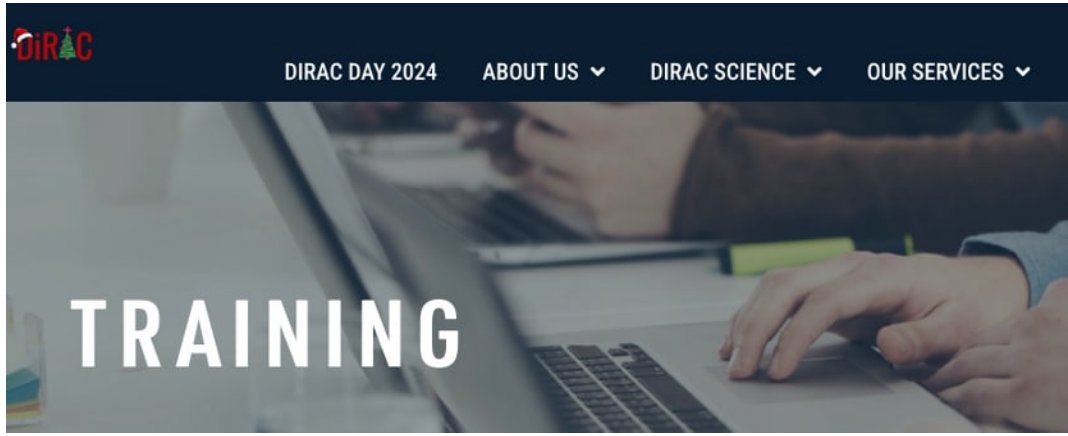
Governance of EuroHPC

- Centralised investment by the EU
- Matching funds from the participating countries
- Machines are split between EuroHPC and the national programmes
- National programmes are run independently

Resource Allocation

- A single panel across *all* domains – one-year focused applications
- Raises the bar for *all* domains
- Quality of applications has improved – frequency of calls
- Very expensive in terms of human resources (reviewers!)

training & education

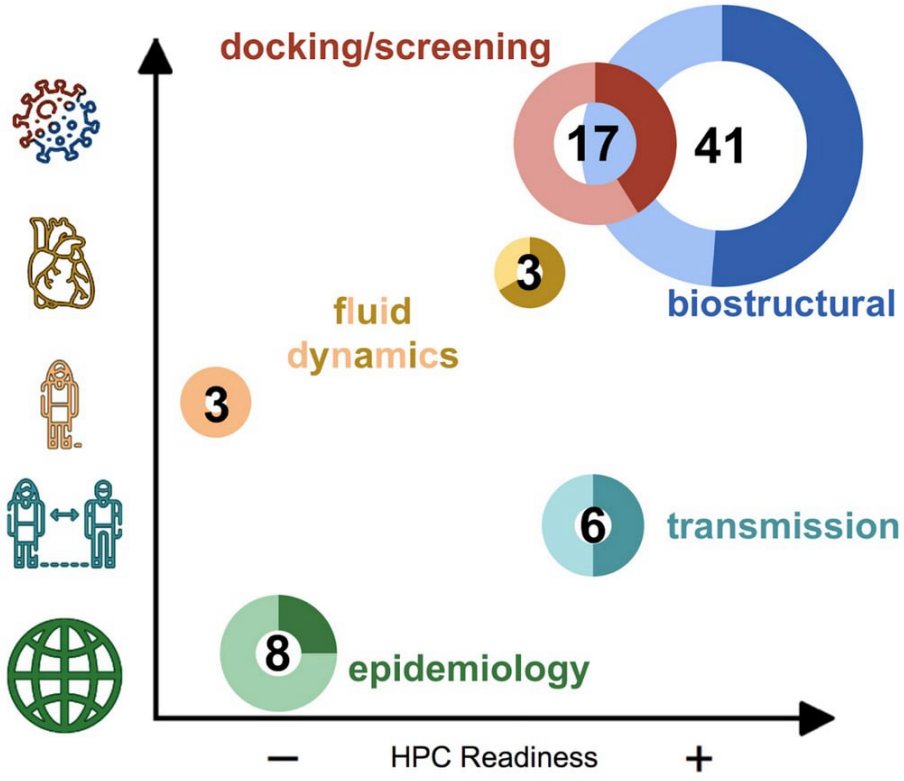


- training of current researchers & RSEs
- training of students – rooted in universities vs online
- industry placements
- hackathons with industry
- students are our biggest asset!

Atos THE UNIVERSITY of EDINBURGH

In collaboration with ATOS and the University of Edinburgh, DiRAC is pleased to invite applications for a 6-month Innovation Placement looking at the application of Quantum Computing to Quantum Field Theory.

Covid-19 and a PNAS paper



PNAS PERSPECTIVE OPEN ACCESS

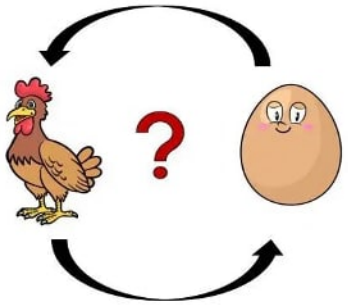
Lessons learned from urgent computing in Europe: Tackling the COVID-19 pandemic

Núria López [a,b], Luigi Del Debbio [a,c], Marc Baaden [a,d], Matej Praprotnik [a,e,f,1], Laura Grigori[a,g], Catarina Simões[a], Serge Bogaerts [a], Florian Berberich [a,h], Thomas Lippert[a,h], Janne Ignatius [a,i], Philippe Lavocat [a,j], Oriol Pineda [a,k], Maria Grazia Giuffreda[a,l], Sergi Girona [a,k], Dieter Kranzlmüller[a,m], Michael M. Resch [a,n], Gabriella Scipione[a,o], and Thomas Schulthess[a,l]

emphasizes the need for a thriving ecosystem

Back to Visions…

- HPC is an amazing catalyser of research
- AI/ML provide new opportunities/directions/synergies
- We urgently need a UK exascale **system**
- Onboarding/readiness of all communities
- People are at the core of Computing



CIUK 2024 Presentations

Mayank Kumar (Science and Technology Facilities Council)

Integration of SYCL with MPI in Multi-scale Universal Interface (MUI) library for Multi-Physics Coupling

Abstract: The Multiscale Universal Interface (MUI) is a versatile self-contained library designed to facilitate multi-physics and multi-scale simulations by providing an architecture that enables users to achieve seamless integration of solvers without requiring in-depth knowledge of their underlying topologies. This abstraction simplifies the coupling process, enabling researchers to focus on their scientific inquiries rather than the technical complexities of solver interaction. This work focuses on the improvements in portability and performance achieved by leveraging Message Passing Interface (MPI) and SYCL to handle computationally expensive parts of the coupling algorithm.

Bio: Mayank is a computational scientist working at STFC Daresbury Laboratory, with a background in high-performance computing, computational fluid dynamics and code coupling. Currently working on the heterogeneous parallelization and optimization of codes used within the computational engineering group.

Integration of SYCL with MPI in Multi-scale Universal Interface (MUI) library for Multi-Physics Coupling

Mayank Kumar, Wendi Liu, Omar Mahfoze, Stephen Longshaw

Scientific Computing Department, Daresbury Lab
UKRI-STFC

Motivation

Complex multiphysics/multiscale problems

Science and Technology Facilities Council

Scientific Computing

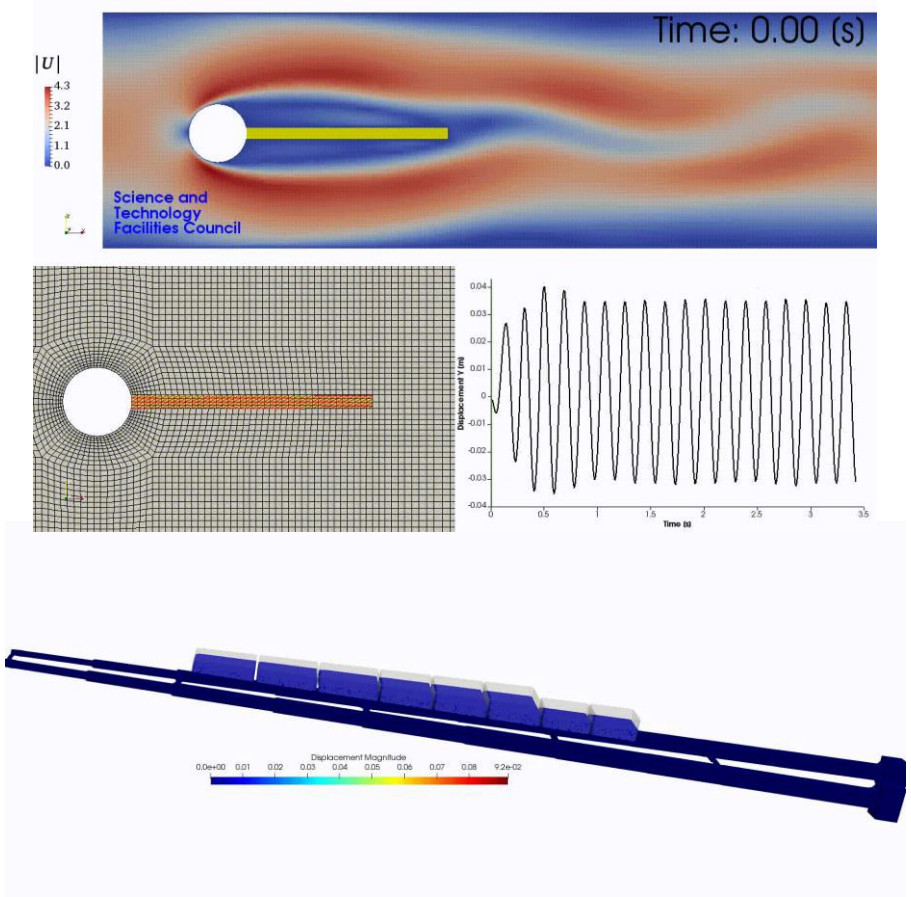
www.ccp-wsi.ac.uk/data_repository/test_cases/test_case_015

CCP-WSI a Collaborative Computational Project in Wave Structure Interaction
www.ccp-wsi.ac.uk

HEC-WSI a High End Computing Consortium for Wave Structure Interaction
www.hec-wsi.ac.uk

MUI

- A concurrent interface for coupling heterogeneous solvers[1,2]
- A header-only code-coupling library written in C++11 (wrappers for C, Fortran and Python)
- MUI transfers data through MPI MPMD as a cloud of points, submitted as frames of time and provides spatial and temporal data sampling
- Scales to over 100,000 MPI ranks with parallel efficiency in excess of 90% depending on the codes



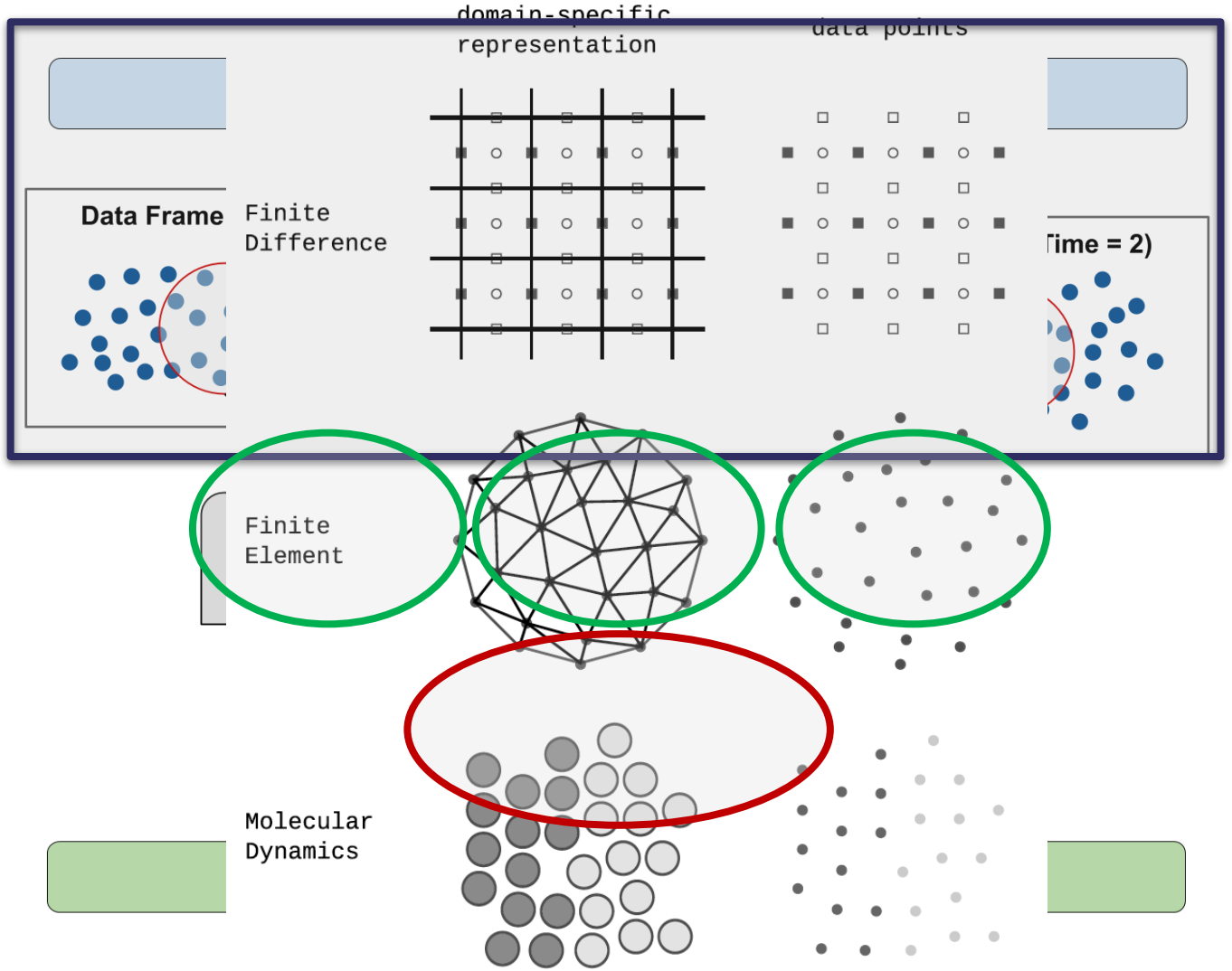
- Spatial Samplers
  - Exact
  - Linear
  - Quintic
  - RBF
- Temporal Samplers
  - Exact
  - Mean
  - Sum
- Algorithms
  - Fixed Relaxation
  - Aitken

https://github.com/MxUI/MUI.git

Science and Technology Facilities Council
Scientific Computing

MUI Workflow

- Couples using a set of discrete data samples and an **interface**:
  1. Convert domain-specific representations to a general form (a **cloud of points** with associated data)
  2. Solver **imparts** data (at a point) to interface with an **associated time-stamp** using **non-blocking** operations
  3. Other solver requests data at specific location and time from MUI interface using **spatial** and **temporal** sampler using **blocking** operations



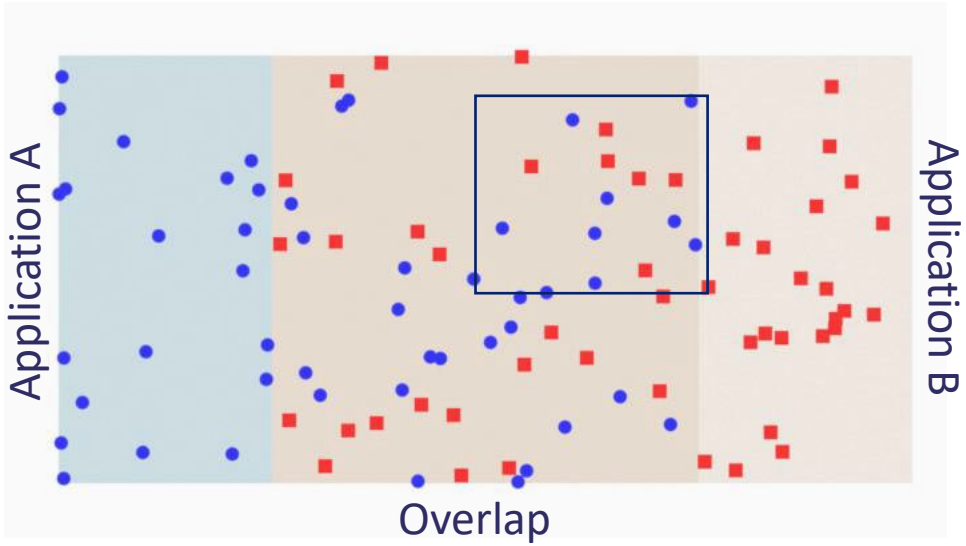
Science and Technology Facilities Council

Scientific Computing

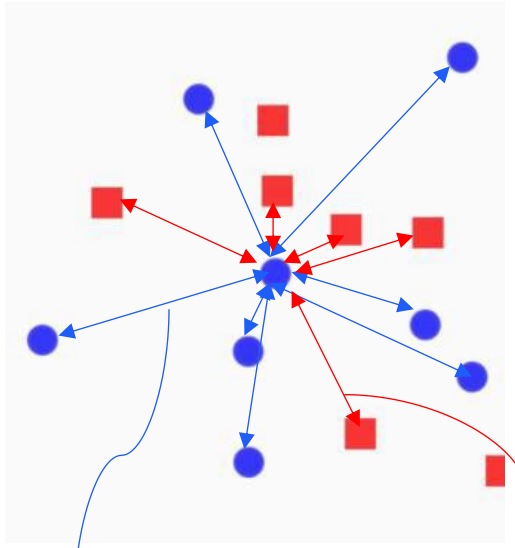
Radial Basis Function (RBF)

- General form of RBF on maintaining high-order consistent/conservative black-box coupling

$$s(\boldsymbol{r}) = \sum_{i=1}^{N} \alpha_i \, \varphi(\|\boldsymbol{r} - \boldsymbol{r_i}\|)$$



• Local Point ▪ Remote Point

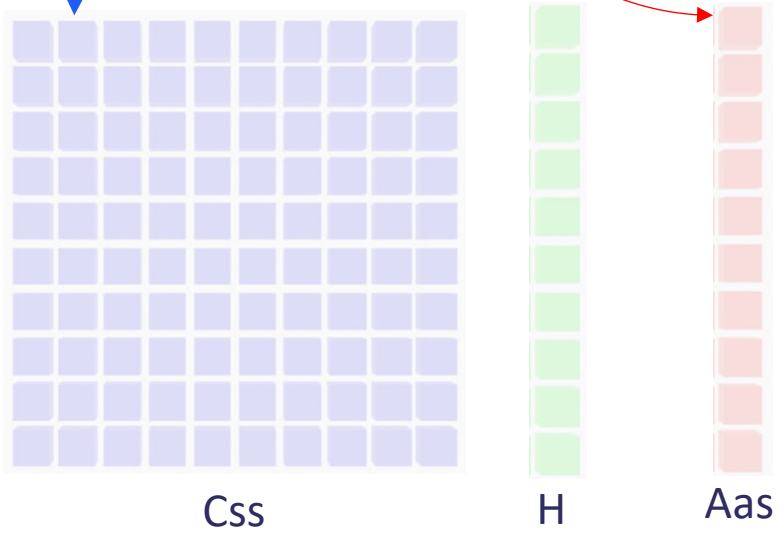


Steps:

Connectivity Matrix Generation

Matrix Solve

Smoothening operation



Science and Technology Facilities Council

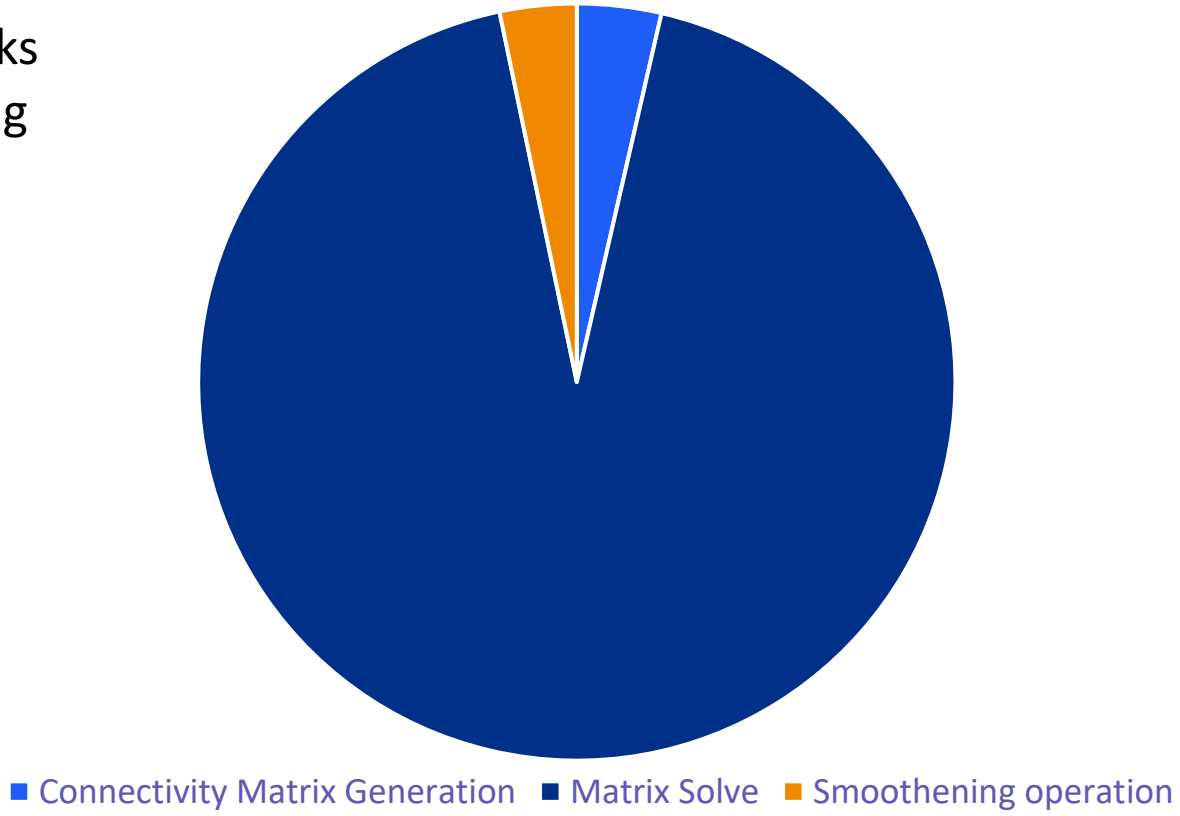
Scientific Computing

RBF filter profile

RBF Coupling workflow is profiled using MUI testing framework
https://github.com/Mayank-K17/MUI-Testing.git

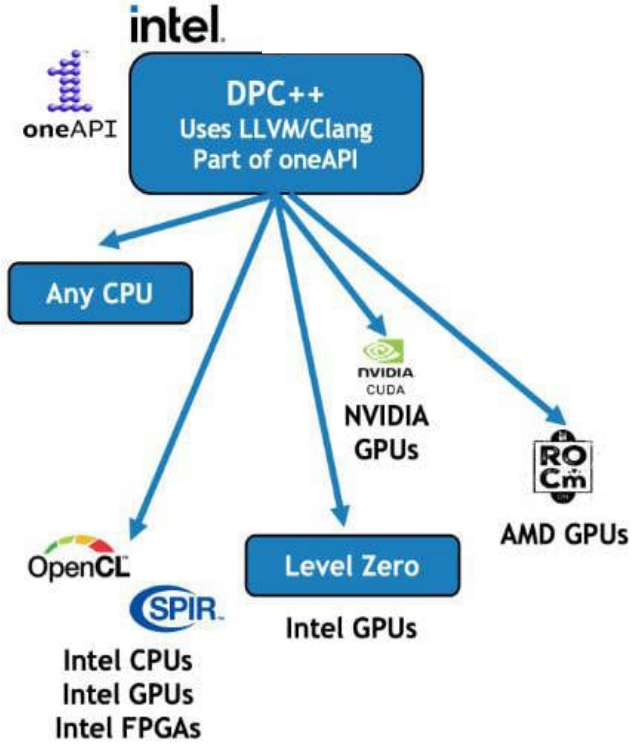
- Two 3D blocks with 5% overlap between the blocks
- RBF Conservative matrix approach with smoothing
- Small test case: 1m block divided into 50*50*50
- Matrix size (576*576)

Operation	Time (s)
Connectivity Matrix Generation	5.8 s
Matrix Solve	176.2 s
Smoothening operation	5.2 s



SYCL

- SYCL (pronounced ‘sickle’) is a royalty-free, cross-platform abstraction layer.
- Enables code for heterogeneous and offload processors to be written using modern ISO C++
- Provides APIs and abstractions to find devices (e.g. CPUs, GPUs, FPGAs) on which code can be executed, and to manage data resources and code execution on those devices



SYCL programming model

- Can use conventional pointers (USM) or SYCL buffers for data management
- Need to define SYCL queue for execution
- Parallel_for operation used to parallelize the for loop
- Intel oneAPI Software stack support is limited for Nvidia and AMD devices

C++

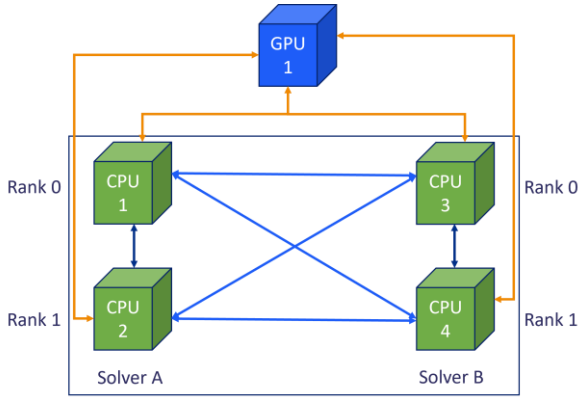
```
for (size_t i = 0; i < vec1.size(); ++i)
{
    res[i] = vec1[i] + vec2[i];
}
```

SYCL

```
sycl::queue q(sycl::default_selector{});
q.parallel_for(sycl::range<1>(vec1.size()), [=](sycl::id<1> i)
{
    res[i] = vec1[i] + vec2[i];
}).wait()
```

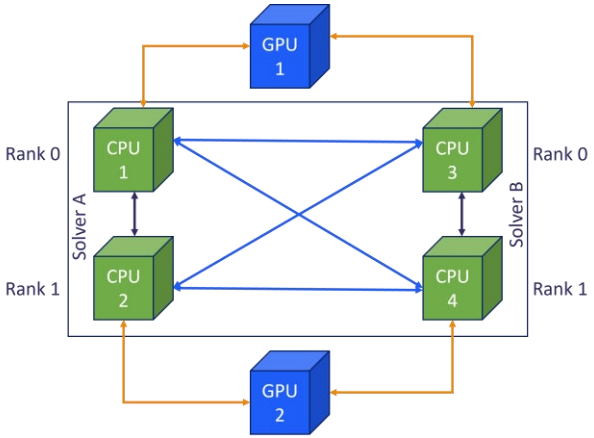
SYCL Integration with MPI

GPU shared across application



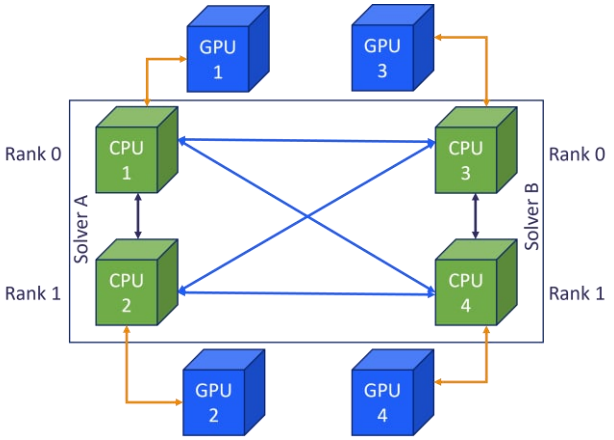

```
sycl::queue
q{sycl::default_selector_v};

sycl::queue
q{sycl::gpu_selector_v};
```

GPU shared across local rank




```
std::vector<sycl::device> Devs;
        for (const auto &plt :
sycl::platform::get_platforms())
{
   if (plt.get_backend() ==
sycl::backend::ext_oneapi_cuda  ||
plt.get_backend() == sycl::backend::hip ||
plt.get_backend() == ext_oneapi_level_zero
)
    {
        Devs.push_back(plt.get_devices()[0]);
    }
}
sycl::queue q{Devs[local_rank_]};
```

GPU exclusive to rank



Each rank on a separate node

```
sycl::queue
q{sycl::default_selector_v};
```

Ranks shared on nodes

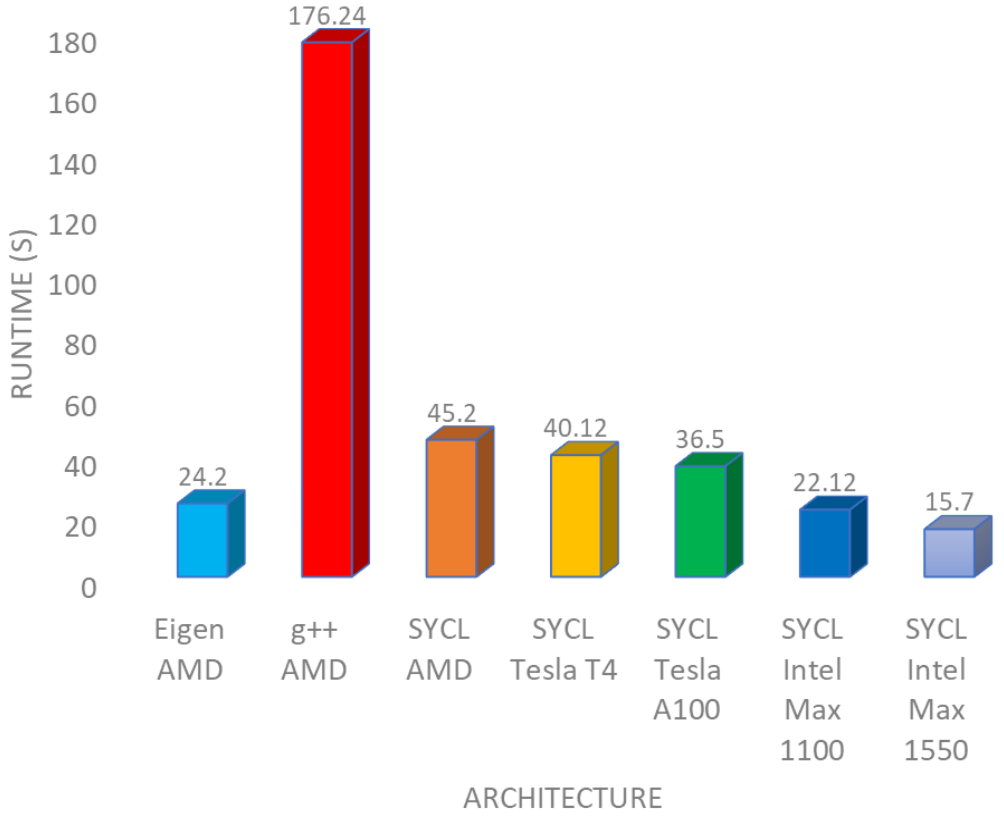
```
sycl::queue
q{Devs[global_rank_]};
```

Science and Technology Facilities Council

Scientific Computing

SYCL implementation

- Conjugate gradient method for solving linear algebra
- Diagonal preconditioner
- Uses SYCL USM (Shared)
- Tested for a range of CPUs and GPUs

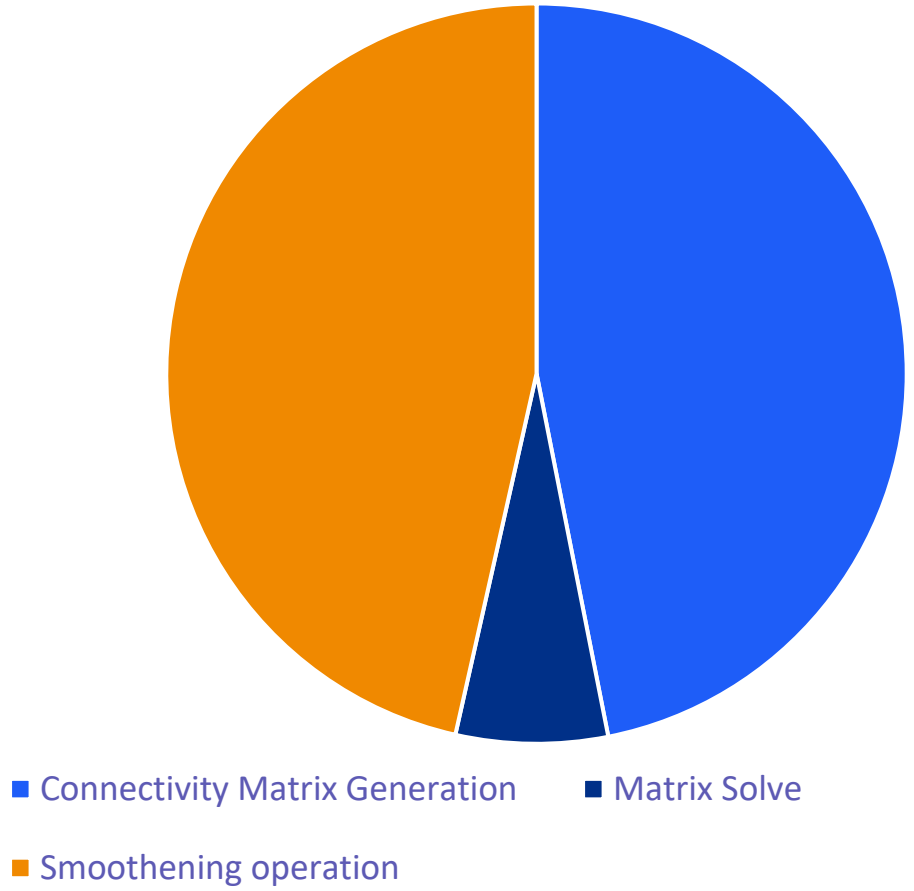


Profiling RBF filter

RBF Coupling workflow is profiled using MUI testing framework
https://github.com/Mayank-K17/MUI-Testing.git

- Two 3D blocks with 100*100*100 grid points
- 5% overlap between the blocks
- RBF Conservative matrix approach with smoothing

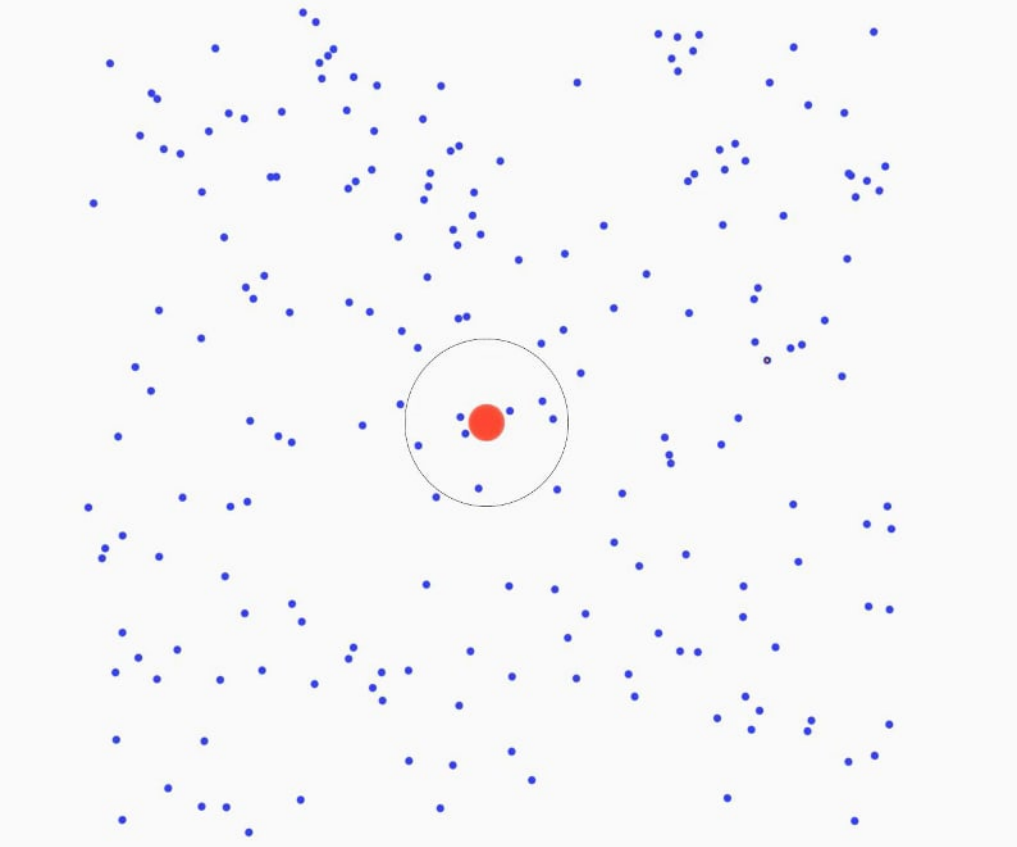
Operation	Time (s)
Connectivity Matrix Generation	341 s
Matrix Solve	48.2 s
Smoothening operation	338 s



Science and Technology Facilities Council
Scientific Computing

Neighbour search

- Each data point(N) checks every remote point(M) to check if it lies within the RBF cutoff radius (N*M complexity)
- Distance calculation is an expensive operation
- Most distance calculation are unnecessary for RBF but unavoidable for unordered set of points
- Other particle-based solvers (DEM, SPH) also have neighbour searches for unordered set of points



Science and Technology Facilities Council

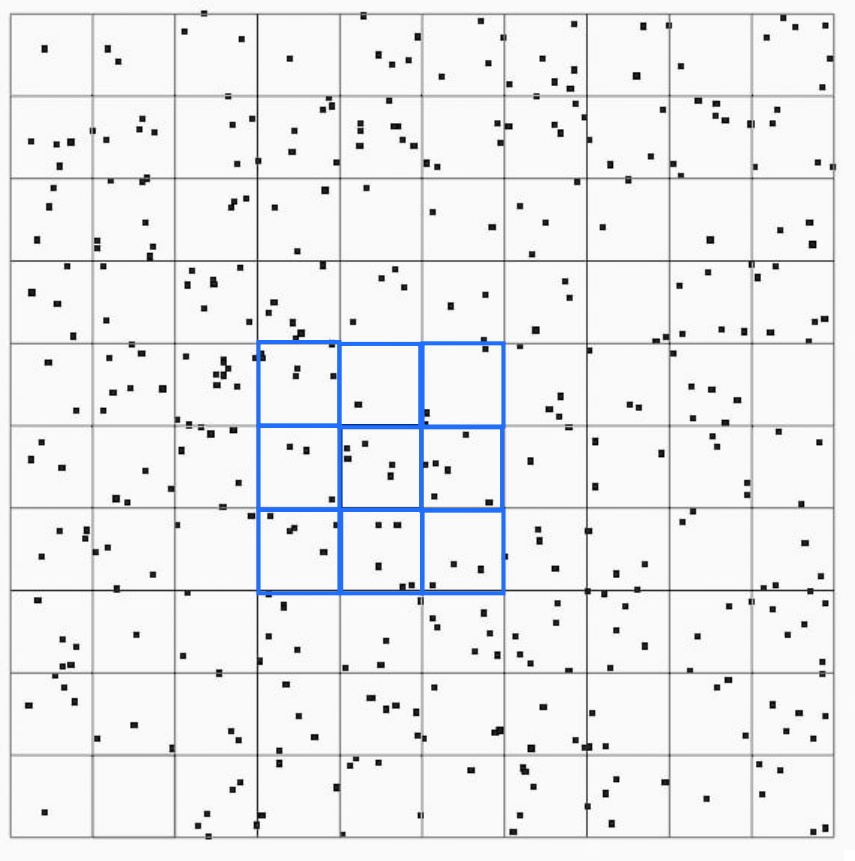
Scientific Computing

Boxing scheme

- Create an ordered grid of boxes of size of RBF cutoff radius across the extend of the points
- Each box is ordered sequentially specifying the bounds for each box ID
- Each point is binned in the specific box based on location and assigned box ID tag

$$\mathrm{P}_{\mathrm{boxId}} = \left(x_p - x_{min}\right) * \delta + Nx * \left(\left(y_p - y_{min}\right) * \delta\right)$$

- Box Structure has two objects first a vector of points and vector of Neighbour box IDs
- For a point within any box search for possible neighbour only within the same block and neighbour blocks



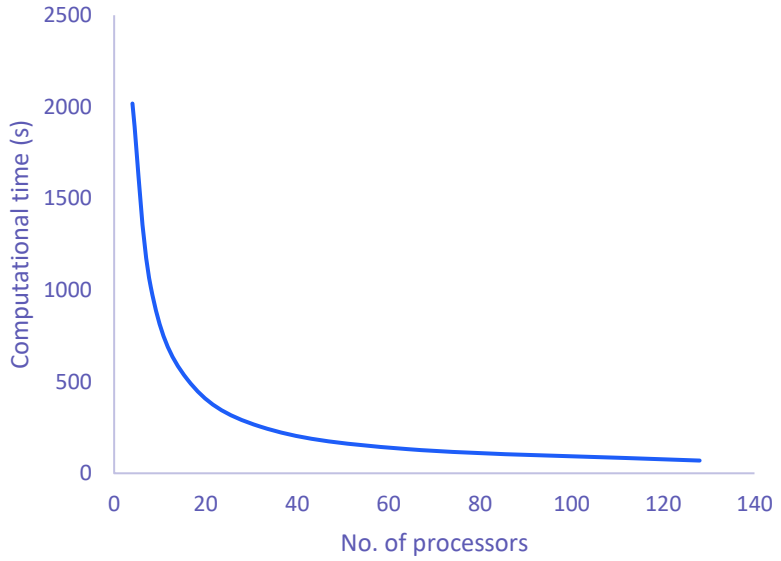
Performance and Scaling

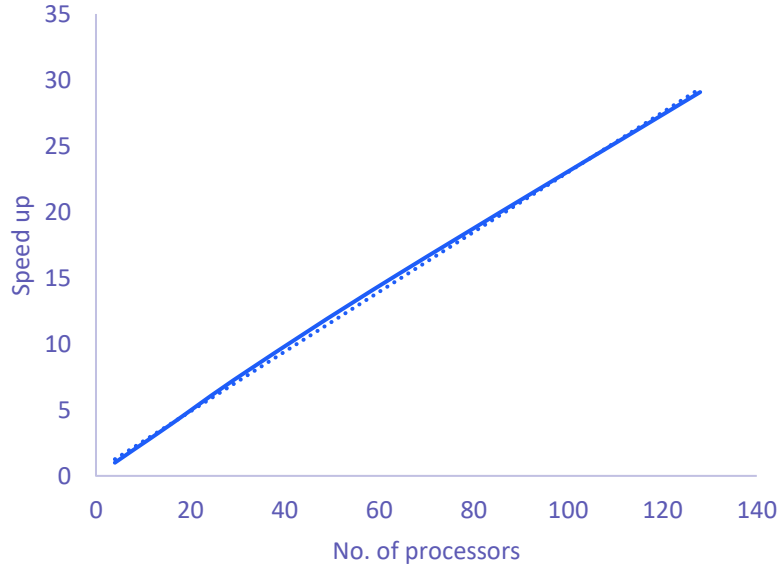
Performance for 1M MUI-Testing case

Operation	Time (s)
Connectivity Matrix Generation	3.6 s (341 s)
Matrix Solve	48.2 s (178 s)
Smoothening operation	3.4s (338 s)

Scaling for 64M MUI-Testing case up to 128 cores

- Two 3D blocks with 400*400*400 grid points
- 20% overlap between the blocks
- Scales well upto 128 cores (28X for 32X processors)



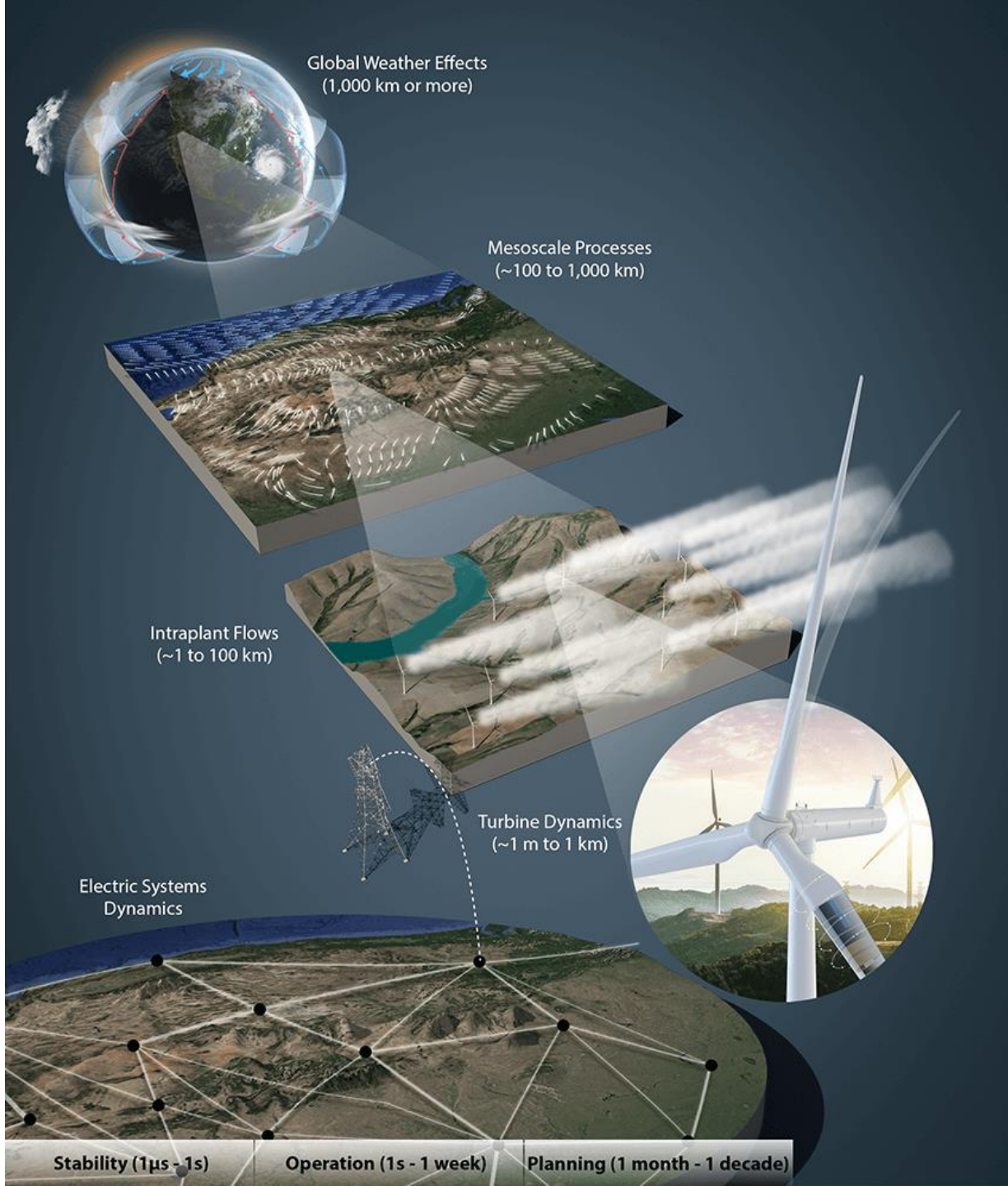


Science and Technology Facilities Council

Scientific Computing

Mesoscale and Microscale Coupling

- Part of the ExCALIBUR *Turbulence at Exascale* project
- **Aim:** perform near real time simulation of urban flows and wind-farms
- These simulations require real-time accurate weather forecasting model and high-fidelity computational fluid mechanics tools
- Efficient coupling between the WRF model and XCompact3D using MUI



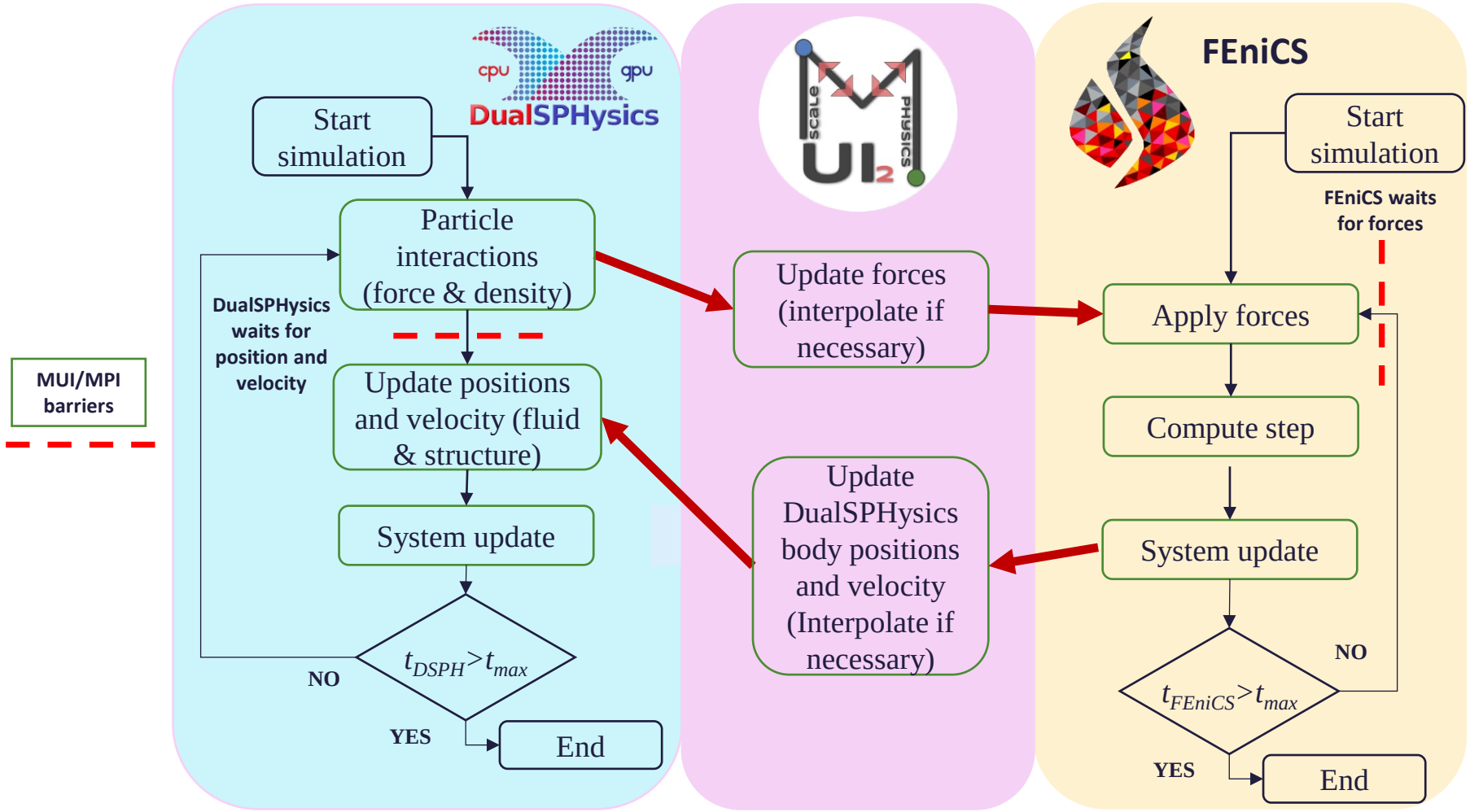
Science and Technology Facilities Council

Scientific Computing

Illustration by Josh Bauer and Besiki Kazaishvili, NREL
https://www.nrel.gov/news/program/2019/nrel-publishes-science-journal-article-posing-three-challenges-to-wind-energy-potential.html

Coupling Strategy DualSPHysics & FEniCS

- Simulations launched concurrently with MPI
- Forces acting on fluid and wall (structure) particles calculated in DualSPHysics
- MUI calls MPI barrier ➔ Forces sent to FEniCS
- Project CHRONO computes structural time step
- Velocities and centre of mass sent back to DualSPHysics
- The position and velocity of structure and fluid particles is then updated in DualSPHysics



Similar to strategy of Martínez-Estévezetal (2023) for coupling DualSPHysics with Project CHRONO

Science and Technology Facilities Council

Scientific Computing

Summary

- Profiled of a moderate test case (Matrix solve hotspot)
- Conjugate gradient solver and Diagonal preconditioner implemented using SYCL USM and tested for a range of processors.
- Profile of a large test case (Neighbour connectivity hotspot)
- Boxing scheme implemented to reduce the computational time of the Neighbour search algorithm by a factor of almost 100.
- Needs to be tested for real world test cases.

CIUK 2024 Presentations

Gregory Tourte (Research Systems Administrator, Advanced Research Computing, University of Oxford)

Use of WEKA FS for fast ephemeral storage powering academic HPC

Abstract: As part of this year's University of Oxford Advanced Research Computing (ARC) storage infrastructure upgrade, we implemented a solution using a 275 TiB Weka filesystem for our fast scratch area in our HPC and HTC infrastructure, replacing the existing GPFS solution. This is an ephemeral data storage area with the primary requirement of extremely high-speed data read and write operations, supporting our ARC, HTC, and Jade clusters. In this talk we present an overview of our requirements, hardware and setup design. We present benchmark results from our raw tests as well as the results from evaluating real world applications alongside our users in several disciples, each of whom are working with applications where I/O is critical.

Bio: Gregory J. L. Tourte is a Research Systems Administrator in the Advanced Research Computing Team at the University of Oxford which provides and supports high performance computing facilities for researchers throughout the institution. Prior to joining Oxford, he worked as a Senior Research Associate at the University of Bristol in the Digital Health Department and School of Geographical Sciences where he started as a research software engineer and system administrator for the research group's supercomputer and used this opportunity to develop an understanding of the research data management required for the large quantity of data being generated. He holds an undergraduate Master's in Mechanical Engineering and a Master's in Computer Science from the University of Bath.

Use of WEKA FS for fast ephemeral storage powering academic HPC

CIUK 2024

Gregory J. L. Tourte & The ARC Team

Contents

About Us

- Advanced Research Computing team at the University of Oxford;
- Part of central IT services;
- Provide generic, non-subject specific, high performance computing facilities for the whole institution;
- Provide help with software installation and mininal support for use on the cluster;
- Other facilities for more subject-specific uses at departmental/division levels are also available;
- About 1300 active users, across 600 active projects.

Oxford ARC estate

- ARC capability cluster (14 640 CPU cores):
  - 305× 48 core worker nodes;
  - 2× Intel Platinum 8628 24 core 2.90 GHz Cascade Lake CPUs;
  - 384 GB memory;
  - HDR 100 infiniband interconnect (The fabric has a 3:1 blocking factor with non-blocking islands of 44 nodes);

Oxford ARC estate

- ARC capability cluster (14 640 CPU cores):
  - 305× 48 core worker nodes;
  - 2× Intel Platinum 8628 24 core 2.90 GHz Cascade Lake CPUs;
  - 384 GB memory;
  - HDR 100 infiniband interconnect (The fabric has a 3:1 blocking factor with non-blocking islands of 44 nodes);
- HTC high throughput cluster:
  - 95× worker nodes;
  - Including 49× GPGPU nodes;
  - 2× high memory (3 TB) nodes;
  - About 20 nodes with HDR 100 interconnect;
  - 10 Gbit Ethernet;

Oxford ARC estate

- JADE2.5 technology pilot cluster:
  - 3× 128 core worker nodes;
  - 2× AMD EPYC 9534 64-Core CPUs and;
  - 8× AMD MI300X GPUs;
  - 2× NDR 200 networking;
  - 32 TB NVME SSD local scratch.

Storage Infrastructure (prior to update)

- 19 TB networked home on nfs via netapp;
- 18 TB shared architecture specific software install on nfs via netapp;
- 1.7 PB shared project bulk storage on GPFS;
- 51 TB ×3 shared scratch storage on GPFS (different ones per interconnect);
- 100 GB local scratch storage available on nodes for non MPI jobs.

Partial Upgrading of the storage

- GPFS storage was put in place in 2019;
- Running out of support;
- Running out of space on the bulk storage;

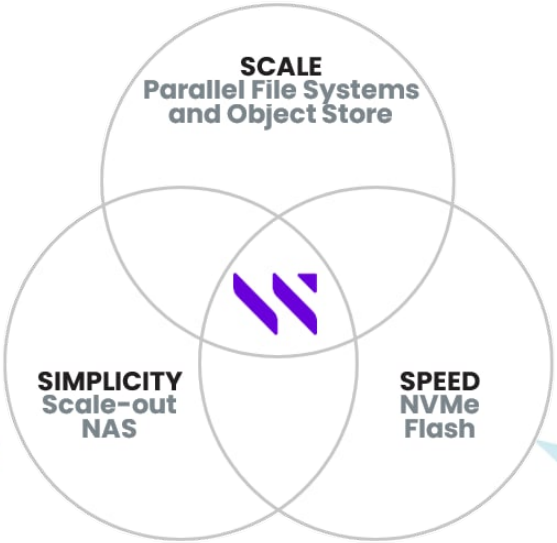
Partial Upgrading of the storage

- GPFS storage was put in place in 2019;
- Running out of support;
- Running out of space on the bulk storage;
- Need to upgrade both bulk and scratch storage;
- Went to tender little over a year ago;
- Chose a mixed solution of Lenovo ontap for bulk and WEKA for scratch.

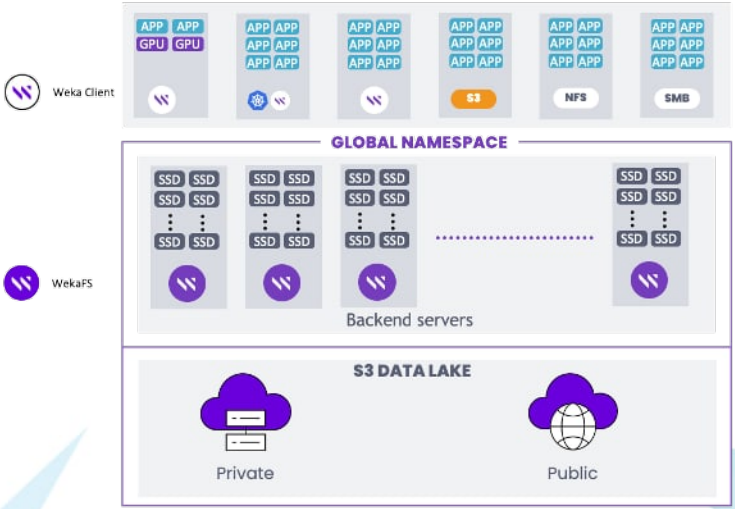
Bulk vs Scratch vs local

- Shared bulk storage is large but slow;
- Scratch only available during processing time, data need to be moved to bulk at the end of the run;
- Local scratch is very fast but cannot be use in MPI case unless single I/O process;
- Shared scratch should be comparable in speed to local scratch for use across cluster.

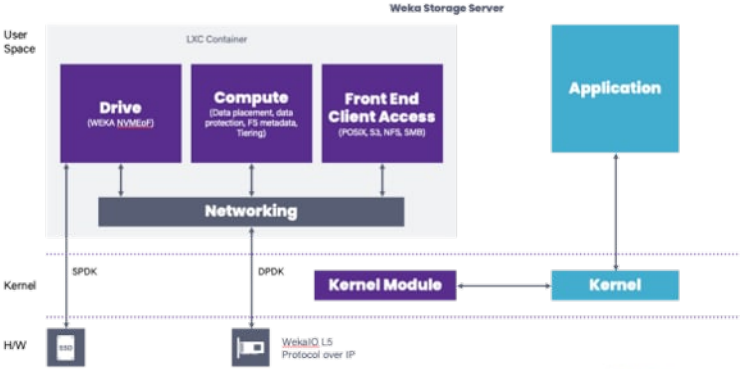
Weka architecture



Weka architecture



Weka architecture



WEKA implementation

- 10× Lenovo ThinkSystem SR635 v3, with:
  - AMD EPYC 9354P 32-Core Processor;
  - 770 GiB RAM;
  - 9× storage NVME SSD drives;
  - Mellanox/NVidia ConnectX-7 NDR 200 Infiniband dual port (using both ports);
  - Mellanox/NVidia ConnectX-6 HDR 100/100 GbE dual port in Ethernet mode (channel bonded);
- Total of ~700 TB of usable NVME SSD;
- Filesystem exported natively over both Ethernet and infiniband, and over NFS.

Caveats

- Standard FIO benchmarks;
- This is artificial!
- Performance is very workload specific;
- No benchmarks against GPFS as in full use so not comparable to a clean empty WEKA setup.

FIO Configuration

Example FIO configuration file

```
[global]
group_reporting
ioengine=posixaio
direct=1
directory=/scratch/
filename_format=$filenum/$jobnum
clocksource=gettimeofday
runtime=3600
iodepth=1
create_serialize=0
stonewall
disable_clat=1
disable_slat=1
clat_percentiles=0
size=4G
write_bw_log
write_iops_log
log_avg_msec=10000
write_lat_log
```

```
[R4096]
new_group
name=R4096
description=4096k random reads
loops=2
numjobs=16
readwrite=randrw
rwmixread=100
bs=4096k
```

Table: Aggregate FIO results for WEKA (Infiniband clients), key results

Test	Cores for FS	Hosts	Threads	Bandwidth (GiB/s)	Feature
4096k sequential reads and writes (mixed)	3	24	48	166.395	Highest mixed operation bandwidth
4096k sequential reads	2	44	32	190.397	Highest total bandwidth (also highest read)
4096k sequential writes	3	24	48	85.841	Highest write bandwidth
4096k random reads and writes (mixed)	2	40	48	157.461	Highest random I/O bandwidth
4096k random reads	2	44	32	157.026	Highest random read bandwidth
4096k random writes	3	44	48	81.162	Highest random write bandwidth

Table: Single HDR 100 client FIO results for WEKA, key results

Test	Cores for FS	Threads	Bandwidth (GiB/s)	Feature
2048k sequential reads and writes (mixed)	2	24	14.396	Highest mixed operation bandwidth
4096k random reads	3	24	14.611	Highest total bandwidth (also highest read)
4096k sequential writes	2	48	9.900	Highest write bandwidth
4096k random reads and writes (mixed)	3	40	13.312	Highest random I/O mixed operation bandwidth
4096k random reads	3	24	14.611	Highest random I/O bandwidth (also highest random read)
2048k random writes	2	40	9.555	Highest random write bandwidth

Table: Single 10 Gb Ethernet (UDP) client FIO results for WEKA, key results

Test	Cores for FS	Threads	Bandwidth (GiB/s)	Feature
4096k sequential reads and writes (mixed)	2	16	1.153	Highest total bandwidth (also highest mixed)
4096k sequential reads	2	4	1.025	Highest read bandwidth
4096k sequential writes	2	48	0.979	Highest write bandwidth
4096k random reads and writes (mixed)	2	40	1.082	Highest random I/O bandwidth (also highest mixed)
4096k random reads	2	40	1.000	Highest random read bandwidth
4096k random writes	2	48	0.937	Highest random write bandwidth

Metadata Tests

```
file system is wekafs
-- started at 06/23/2024 19:37:45 --

mdtest-3.3.0 was launched with 1408 total task(s) on 44 node(s)
Command line used: mdtest '-I' '10' '-z' '5' '-b' '5' '-i' '5' '-u' '-d' '/scratch/mdtest/'
Path: /scratch/mdtest
FS: 250.0 TiB   Used FS: 0.0%   Inodes: 64000.0 Mi   Used Inodes: 0.0%

Nodemap: 11111111111111111111111111111111000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000
1408 tasks, 54996480 files/directories

SUMMARY rate: (of 5 iterations)
Operation                      Max            Min           Mean        Std Dev
---------                      ---            ---           ----        -------
Directory creation        :     529114.286     465800.954     500533.339      24273.005
Directory stat            :    4285066.900    2275094.492    3839420.155     783616.985
Directory removal         :     469512.321     450655.868     463392.363       6758.779
File creation             :     468173.983     458545.503     463251.190       3410.682
File stat                 :    4942181.094    4729493.737    4806719.987      75601.964
File read                 :    2669518.810    2329518.244    2438056.149     121344.593
File removal              :     584956.751     576643.502     580552.044       3059.389
Tree creation             :        371.608        341.918        363.238         10.785
Tree removal              :        307.261        302.602        304.576          1.863
-- finished at 06/23/2024 20:20:41 --
```

Metadata Tests – Part 2

```
file system is wekafs
-- started at 06/21/2024 18:53:15 --

mdtest-3.3.0 was launched with 256 total task(s) on 8 node(s)
Command line used: mdtest '-I' '10' '-z' '5' '-b' '5' '-i' '5' '-u' '-d' '/scratch/mdtest/'
Path: /scratch/mdtest
FS: 250.0 TiB   Used FS: 0.0%   Inodes: 64000.0 Mi   Used Inodes: 0.0%

Nodemap: 11111111111111111111111111111111000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
256 tasks, 9999360 files/directories

SUMMARY rate: (of 5 iterations)
Operation                      Max            Min           Mean        Std Dev
---------                      ---            ---           ----        -------
Directory creation        :     375252.157     227191.091     340357.476      56741.951
Directory stat            :    1468507.584     577985.375    1271466.429     347717.278
Directory removal         :     314605.749     169318.364     280074.914      55501.138
File creation             :     325648.628     181526.007     293545.395      56019.209
File stat                 :    1600473.140     892610.218    1440337.883     274115.285
File read                 :     910013.495     420682.543     808968.167     194162.914
File removal              :     393163.245     188612.541     347378.663      79504.996
Tree creation             :       1315.135        708.321       1178.537        235.553
Tree removal              :       1080.176        984.062       1029.032         32.466
-- finished at 06/21/2024 19:07:39 --
```

Metadata Tests – Part 3

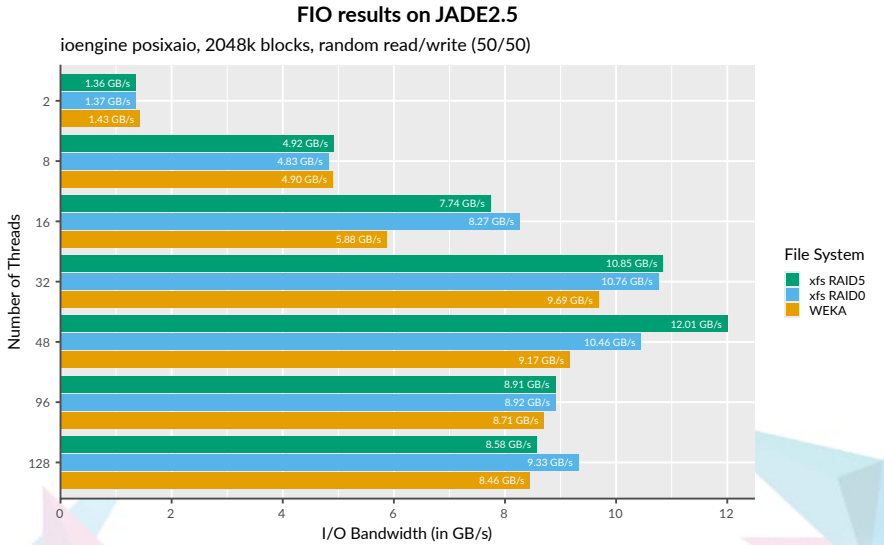
```
file system is wekafs
-- started at 06/21/2024 11:58:53 --

mdtest-3.3.0 was launched with 1 total task(s) on 1 node(s)
Command line used: mdtest '-I' '10' '-z' '5' '-b' '5' '-i' '5' '-u' '-d' '/scratch/mdtest/'
Path: /scratch/mdtest
FS: 250.0 TiB   Used FS: 0.8%   Inodes: 64000.0 Mi   Used Inodes: 0.0%

Nodemap: 1
1 tasks, 39060 files/directories

SUMMARY rate: (of 5 iterations)
Operation                     Max            Min           Mean        Std Dev
---------                     ---            ---           ----        -------
Directory creation        :        771.018        769.265        770.230          0.587
Directory stat            :      10377.942       9499.599       9922.347        287.072
Directory removal         :        775.049        773.970        774.653          0.370
File creation             :        777.549        764.939        774.045          4.620
File stat                 :      22620.889      21627.808      21930.556        355.085
File read                 :      14181.165      13310.001      13746.948        378.857
File removal              :        721.783        704.137        716.403          6.332
Tree creation             :        683.177        672.598        676.964          4.219
Tree removal              :        778.107        775.197        776.358          0.972
-- finished at 06/21/2024 12:17:41 --
```

Benchmarks against local NVME

- Carried out on JADE 2.5;
- Compared to local NVME in RAID0 and RAID5 configuration;
- XFS filesystem;
- Network is NDR 200 (single link only).

Benchmarks against local NVME

random read/write



Caveats

- Still in ‘early stages’ of implementations;
- Not fully rolled out as it is done in parallel with other migration.

User experience

- How was it to install? ‘Super easy, barely an inconvenience’;
- Initial hardware compatibility issue with the wrong network cards provided but was exchanged by Lenovo immediately;
- Support provided for setup;
- Support provided for fine tuning;
- Support provided to updates;
- Very responsive interaction via dedicated slack channel.

User experience – Part 2

- Started with nfs before getting native;
- Quite a different way to setup from other system;
- Native support can be tricky to obtain in mixed hardware environment;
- Can be hardware and software (and version) dependent;
- Still having issues with OFED stack, MPI implementation, and software installed.

Thank You!

Any Questions?

CIUK 2024 Presentations

Oliver Brown (EPCC, University of Edinburgh)

The Scalability of Quantum Air Traffic Control

Abstract: We describe a “top-down” approach for applying quantum computing to aviation; specifically, the problem of routing airplanes through an idealised airspace such that the carbon emissions associated with air travel are minimised. We follow a graph-based model, whereby airspace sectors are represented by vertices connected by weighted edges. The edge weights determine the cost of travelling through the airspace. The problem is naturally scalable in that one can adjust the number of vertices and edges, from a simple 3-by-3 lattice digraph with 24 edges to an irregular digraph representing the UK airspace containing 80 edges. Finding the minimum cost path can of course be solved using classical techniques (e.g., Dijkstra, Bellman-Ford), but our concern here is to measure the scalability of a quantum-based approach, one that uses a Quadratic Unconstrained Binary Optimisation (QUBO) formalism suitable for a D-Wave quantum annealer. The QUBO formalism makes it possible to represent within a single matrix, the airspace network, the entry and exit sectors for an airplane requesting to enter the airspace, as well as the level and location of air traffic congestion at the time of the request.

Bio: Dr Oliver Thomson Brown is a Chancellor’s Fellow at EPCC and leads their Quantum Group. His research interests are all things quantum and HPC, including applications of quantum computing, classical emulation of quantum computers, and programming models for hybrid quantum-HPC. He is also a member of Edinburgh’s Quantum Software Lab.

epcc

The Scalability of Quantum Air Traffic Control

CIUK 2024

Dr Oliver Thomson Brown
o.brown@epcc.ed.ac.uk

THE UNIVERSITY of EDINBURGH

Industry and Academia

NATS

Quantum Base Alpha

UKRI National Quantum Computing Centre

D-Wave The Quantum Computing Company™

Innovate Project no. 10072541.

THE UNIVERSITY of EDINBURGH

Personnel

Dr Michael Bareford
EPCC
University of Edinburgh

Dr Ross Grassie
Quantum Software Lab
University of Edinburgh

epcc

Simulating Air Traffic Control

Python 3.9.12

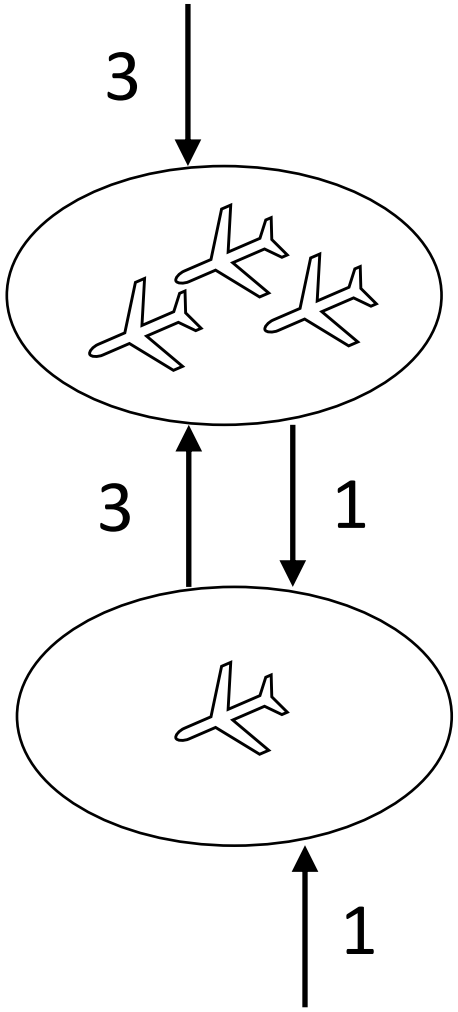
```
networkx 3.2.1
pygraphviz 1.11
graphviz 0.20.3
```

`Graphviz 10.0.1`

THE UNIVERSITY of EDINBURGH

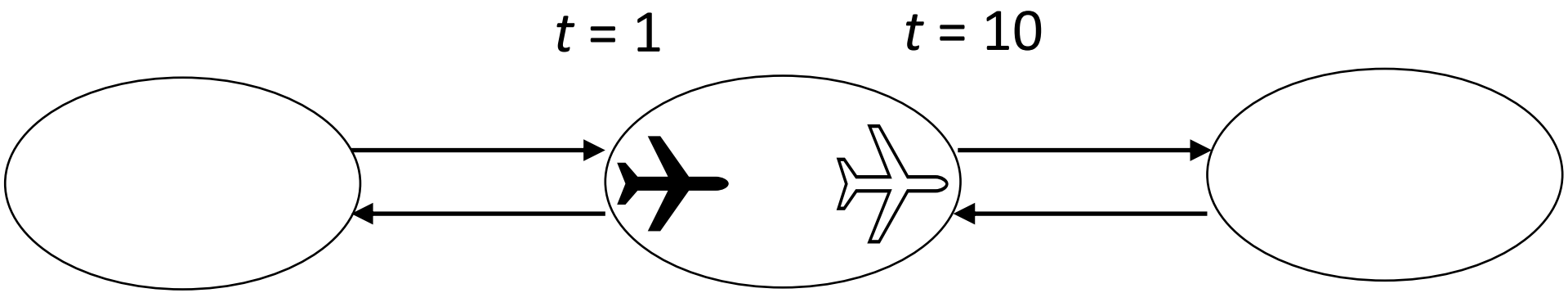
Airspace Graph

- 2D Directed Lattice
- Vertices are Airspace Sectors
- Vertex weight is the number of airplanes in the sector
- Edge weight is equal to the weight of the vertex pointed to by the edge
- All weights initialized to zero before any airplanes have entered the airspace



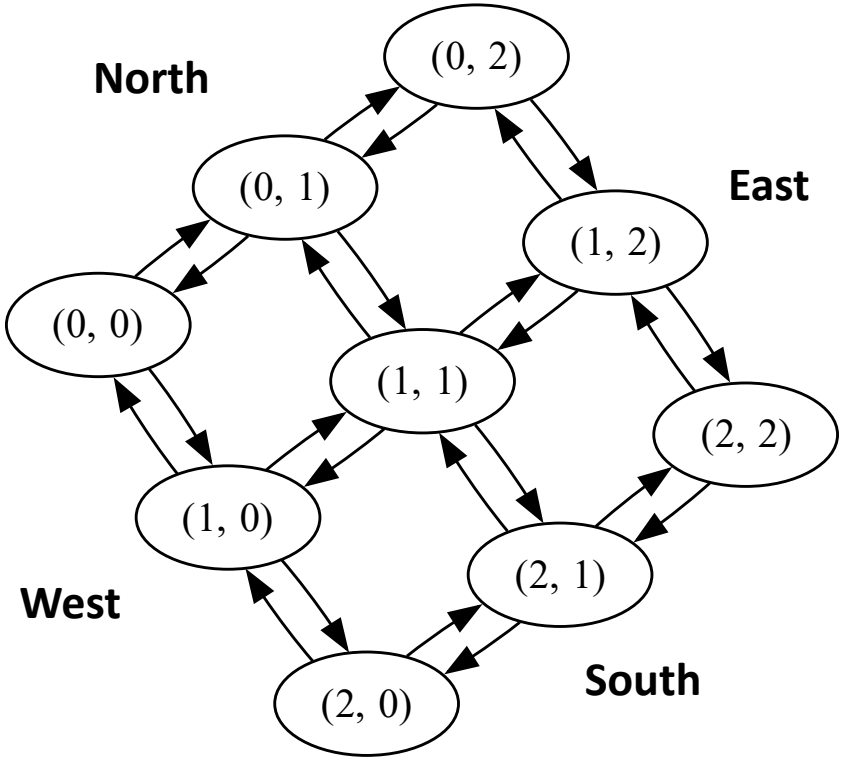
Airspace Travel Costs

- All vertices have an associated cost: the number of steps it takes for an airplane to traverse the sector.
- The cost is the same regardless of the airplane's origin/destination.
- All vertex/sector costs initialized to 10 steps.



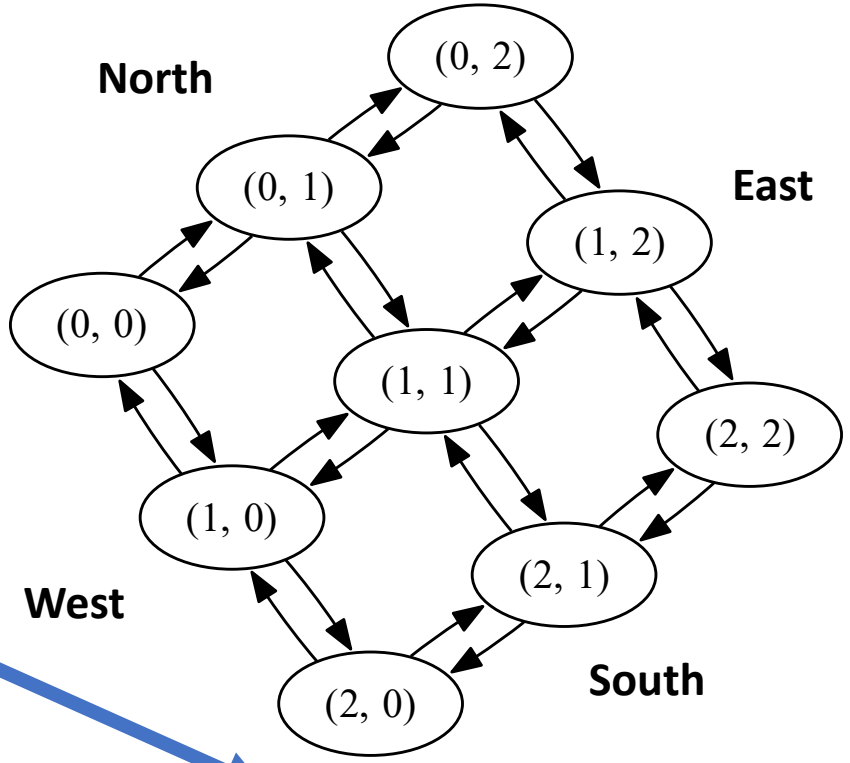
Air Traffic Simulator (part 1)

```
For 1 to MaxSimulationSteps

    Randomly select the number of planes to
    enter airspace

    For each new plane
        Randomly choose…
            1. entry and exit boundaries
            2. entry sector from entry boundary
            3. exit sector from exit boundary

        Call shortest path algorithm to
        obtain route between entry and exit
        sectors

        Update weight of entry sector

        Add airplane to airtraffic list

    …
```



Air Traffic Simulator (part 2)

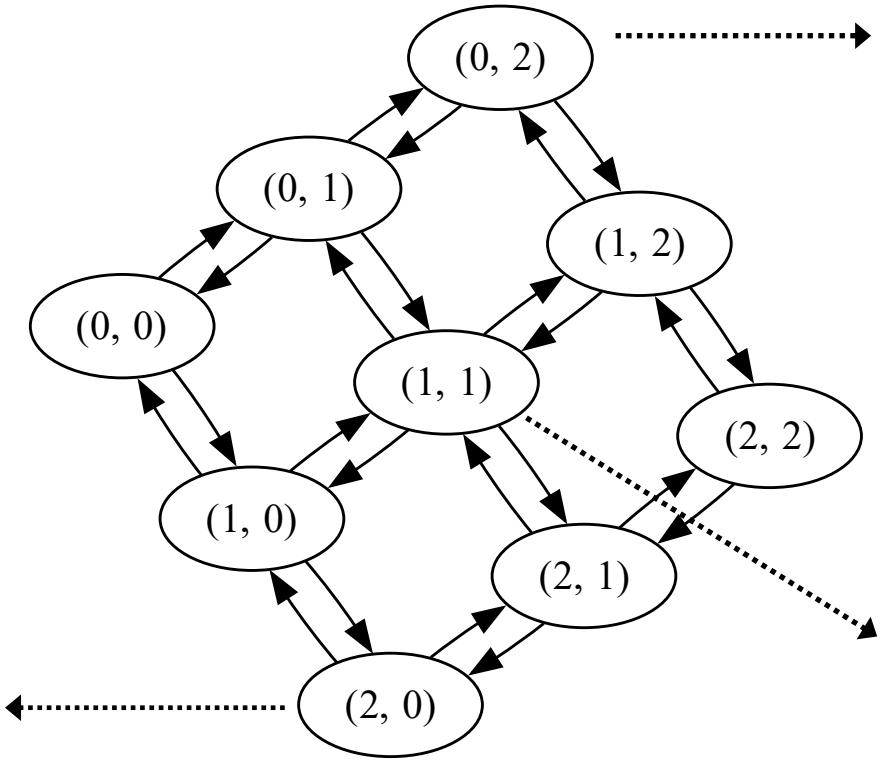
```
For 1 to MaxSimulationSteps

    ...

    For each airplane in airtraffic
        Update position
            1. plane continues to traverse sector

            2. plane has traversed sector and
               is moving on to next sector

            3. plane has traversed exit sector
               and is now leaving airspace


    For each edge in airspace
        set weight to the weight of the
        destination sector
```

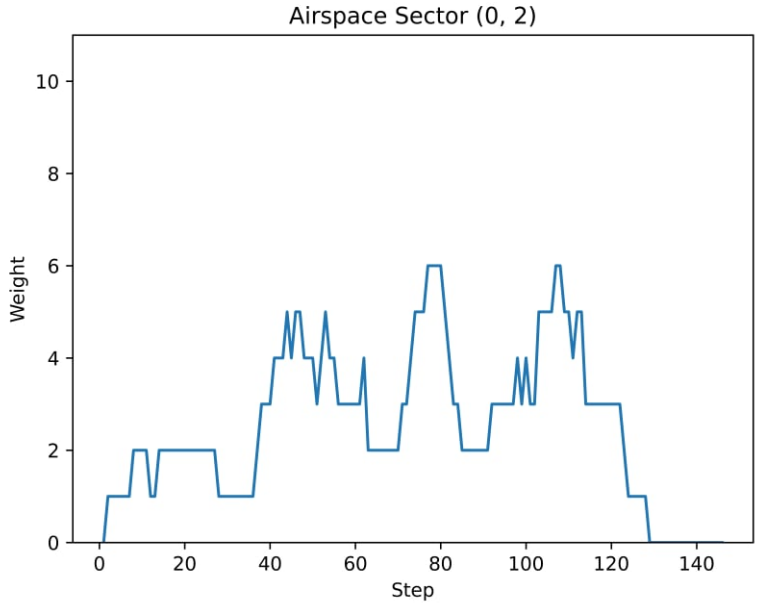


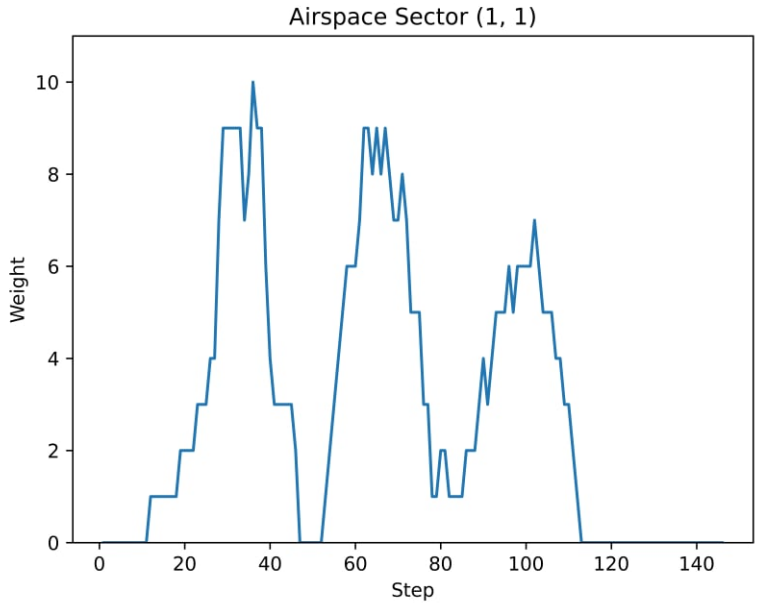
call shortest path routine to obtain route from next sector to exit sector

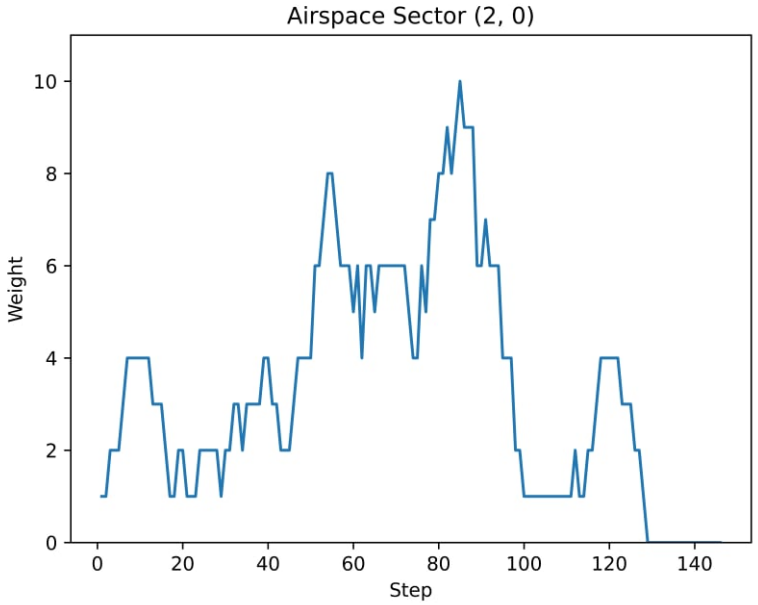
Air Traffic Simulator Plots

Plots show how the weight changes over time for particular sectors.

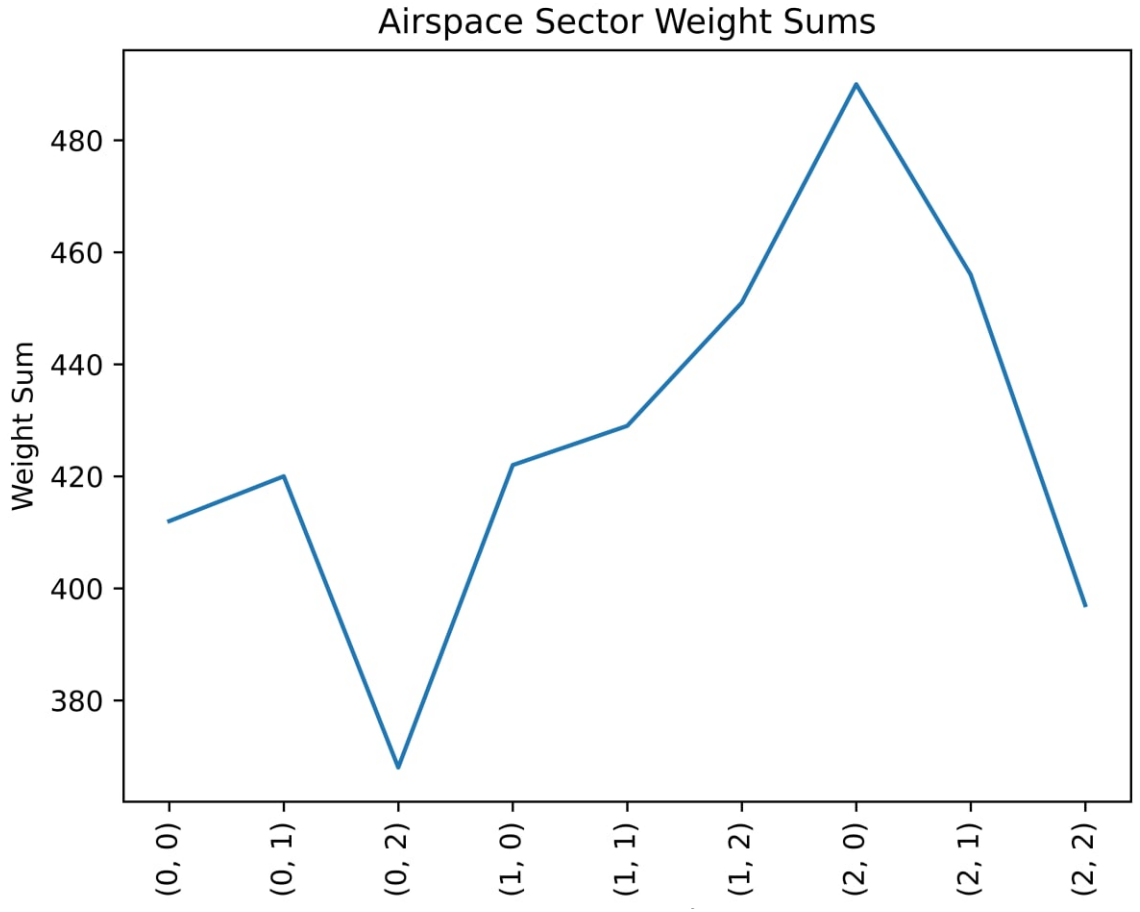


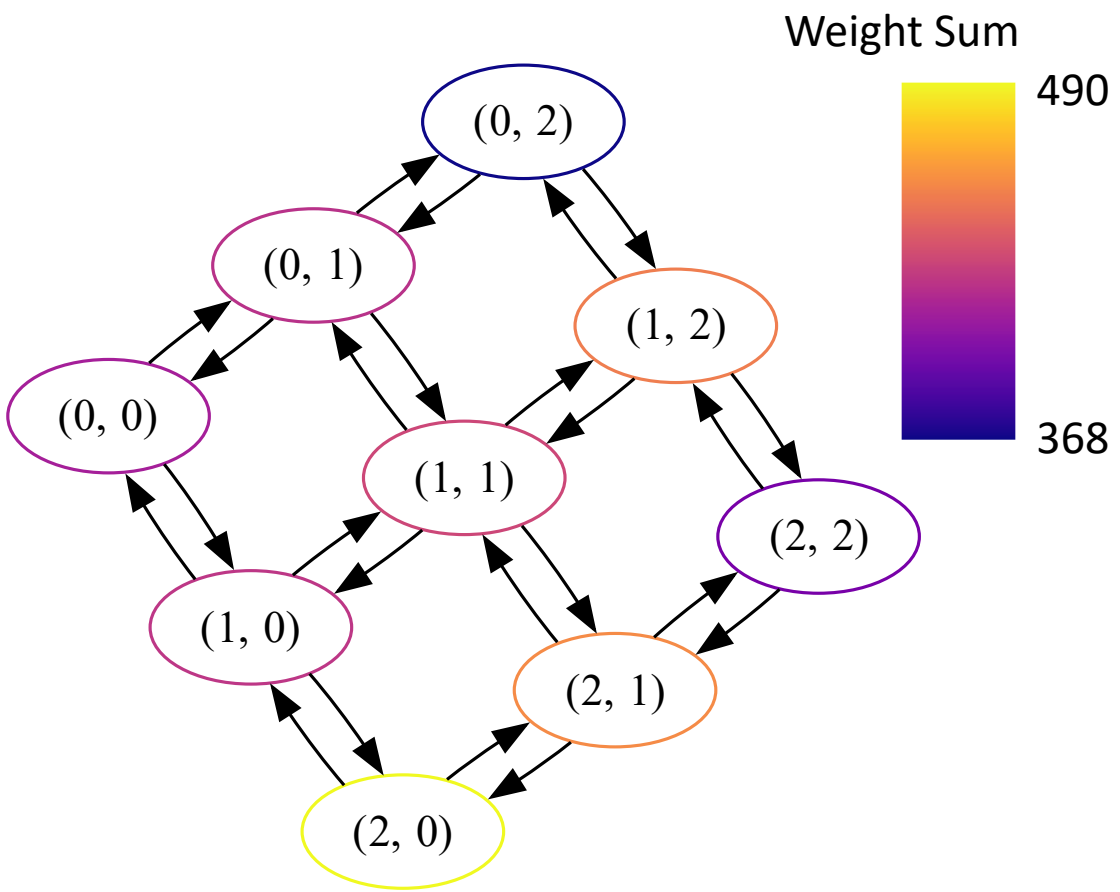






Air Traffic Simulator Plots





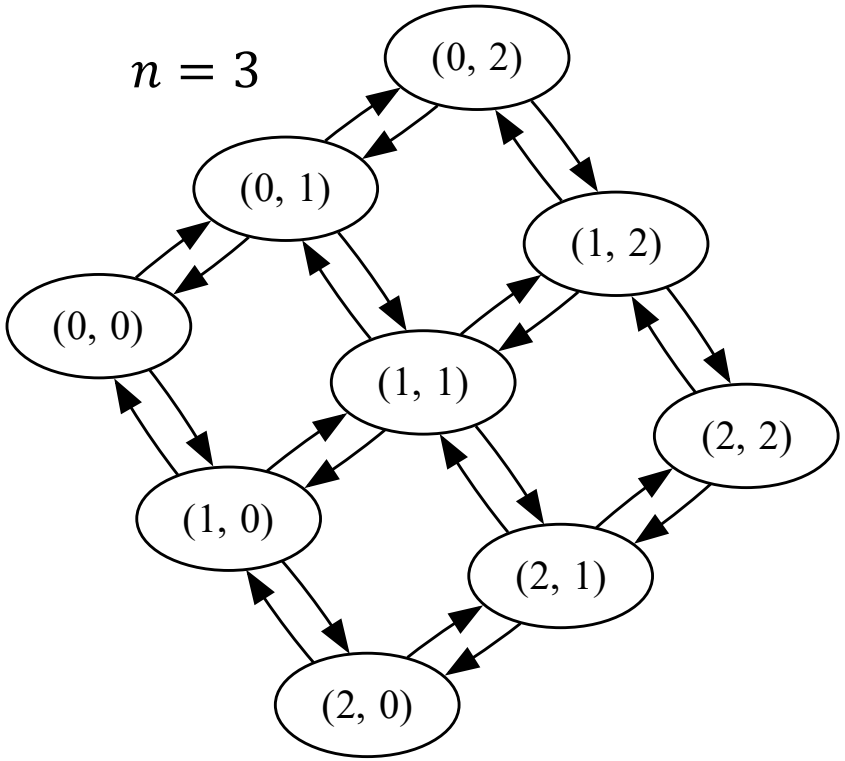
Quantum Implementation

Solving the Network Shortest Path Problem on a Quantum Annealer
Krauss & McCollum, 2020
https://ieeexplore.ieee.org/document/9186612

Directed edge-based approach

$|E|$ qubits required

$|E| \times |E|$ Q Matrix



$|V| = n^2$

$|E| = 4(n-1)n$

$|V| = 9$

$|E| = 24$

QUBO Formulation

QUBO = Quadratic Unconstrained Binary Optimisation

Constraint Functions

$$H_s = \left[\sum_j x_{s,j} - \sum_k x_{k,s} - 1\right]^2$$

the **source** vertex (*s*) has one more edge leaving it than entering it

$$H_t = \left[\sum_j x_{t,j} - \sum_k x_{k,t} + 1\right]^2$$

the **terminal** vertex (*t*) has one more edge entering it than leaving it

$$H_i = \left[\sum_j x_{i,j} - \sum_k x_{k,i}\right]^2$$

for all **other** vertices, the number of edges entering equals the number of edges leaving

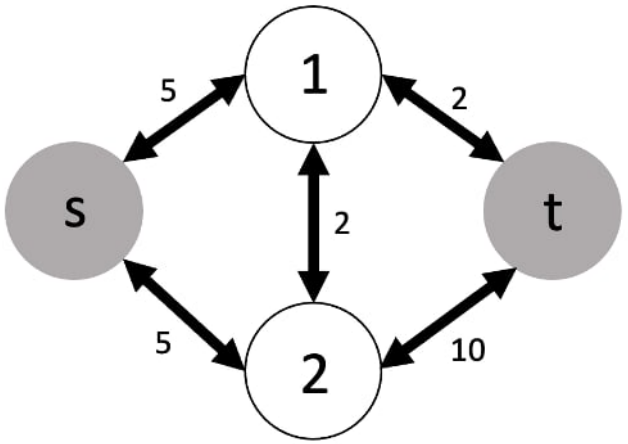
Objective Function

$$H_C = \sum_{(i,j)\in E} c_{i,j} \times x_{i,j}^2$$

the edges in the shortest path

DOI: 10.1109/TQE.2020.3021921

Populating the Q Matrix



$$H = \alpha\left(H_s + H_t + \sum_{i \notin \{s,t\}} H_i\right) + H_C$$

	$x_{s,1}$	$x_{s,2}$	$x_{1,s}$	$x_{1,2}$	$x_{1,t}$	$x_{2,s}$	$x_{2,1}$	$x_{2,t}$	$x_{t,1}$	$x_{t,2}$
$x_{s,1}$	5	2α	-4α	-2α	-2α	-2α	2α	0	2α	0
$x_{s,2}$	0	5	-2α	2α	0	-4α	-2α	-2α	0	2α
$x_{1,s}$	0	0	$5+4\alpha$	2α	2α	2α	-2α	0	-2α	0
$x_{1,2}$	0	0	0	$2+2\alpha$	2α	-2α	-4α	-2α	-2α	2α
$x_{1,t}$	0	0	0	0	2	0	-2α	2α	-4α	-2α
$x_{2,s}$	0	0	0	0	0	$5+4\alpha$	2α	2α	0	-2α
$x_{2,1}$	0	0	0	0	0	0	$2+2\alpha$	2α	2α	-2α
$x_{2,t}$	0	0	0	0	0	0	0	10	-2α	-4α
$x_{t,1}$	0	0	0	0	0	0	0	0	$2+4\alpha$	2α
$x_{t,2}$	0	0	0	0	0	0	0	0	0	$10+4\alpha$

Quantum Time Complexity

Time required to prepare QUBO formulation is $\mathcal{O}(|E|)$.

Preparation Time for directed edge-based approach is $\mathcal{O}(\Delta|E|)$,
where Δ is the maximum degree.

Assume annealing time is of $\mathcal{O}(1)$.

DOI: 10.1109/TQE.2020.3021921

Time Complexity Comparison between Dijkstra and quantum annealing

epcc

Dijkstra Time: $\mathcal{O}(|V|^2)$

Quantum Annealing Time: $\mathcal{O}(\Delta|E|)$

$|E|$ is the number of edges

$|V|$ is the number of vertices

Δ is the maximum degree

Therefore, for quantum annealing to beat Dijkstra,

$$\Delta|E| < |V|^2$$

$$32(n-1)n < n^4$$

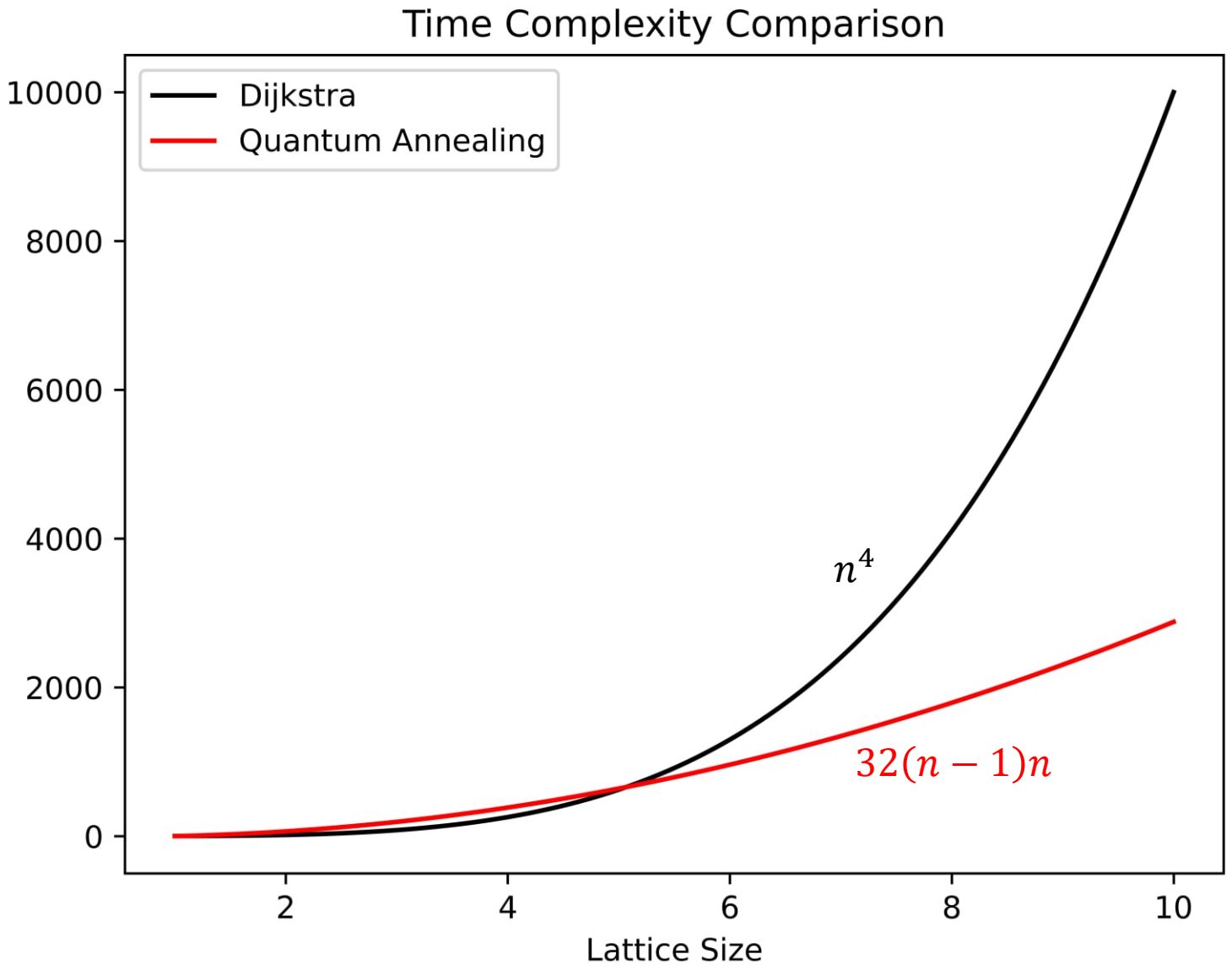
For a 2D lattice,

$$\Delta = 8$$

$$|V| = n^2$$

$$|E| = 4(n-1)n$$

Time Complexity Comparison between Dijkstra and quantum annealing



At crossover point lattice size $(n \times n)$ is approximately,

$n \approx 5$

or

80 edges/qubits!

$$|E| = 4(n-1)n$$

Time Complexity Comparison between Dijkstra and quantum annealing for non-negatively weighted graphs

epcc

Dijkstra Time: $\mathcal{O}(|E| + |V|log|V|)$

Quantum Annealing Time: $\mathcal{O}(\Delta|E|)$

$|E|$ is the number of edges

$|V|$ is the number of vertices

$$\mathcal{O}(\Delta|E|) = \mathcal{O}\big((1 + \delta)|E|\big)$$

$\Delta = 1 + \delta$ is the maximum degree

Therefore, for quantum annealing to beat Dijkstra,

$$\delta|E| < |V|log|V|$$

$$28(n-1)n < n^2 logn^2$$

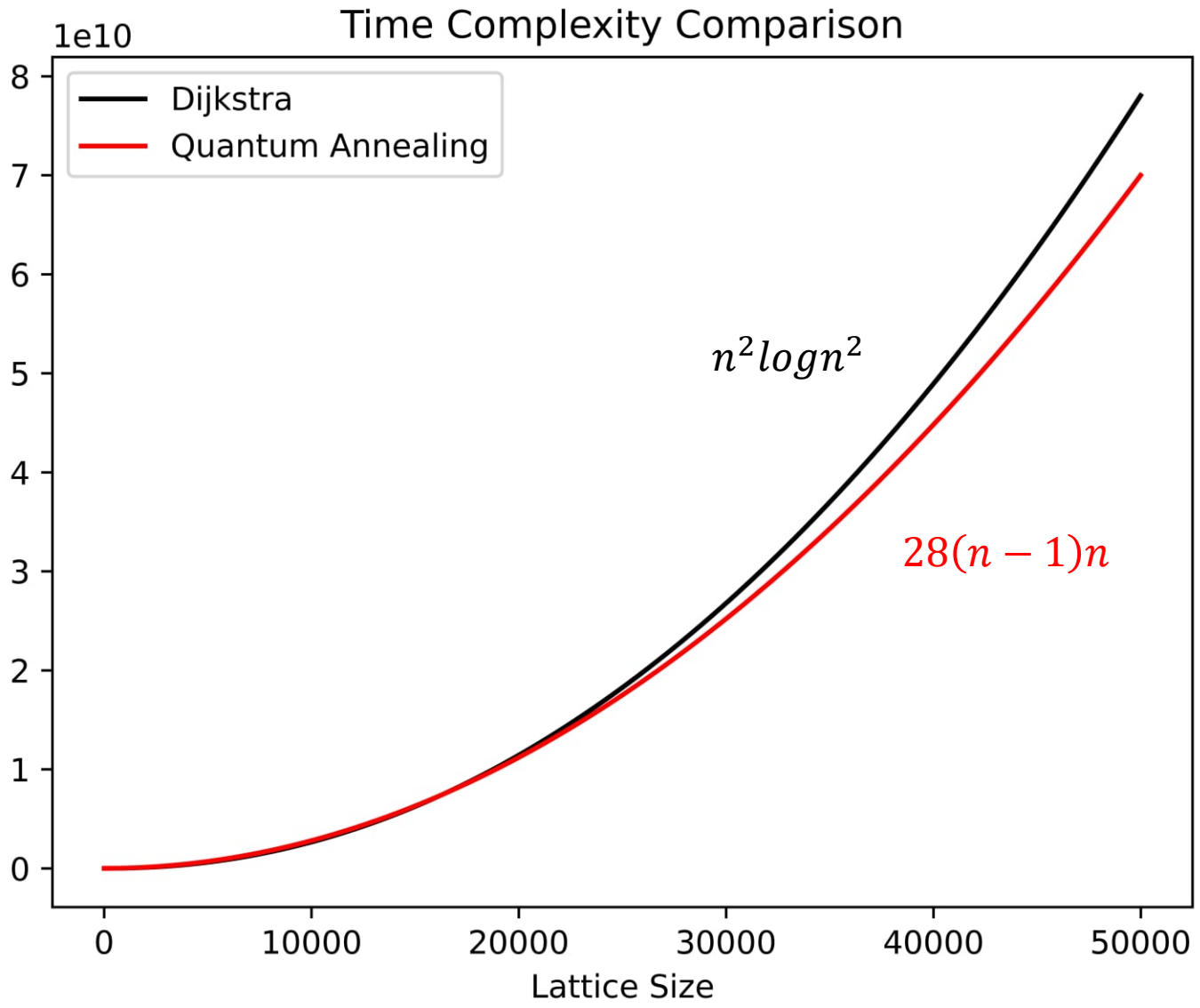
For a 2D lattice,

$$\Delta = 8 \quad \therefore \delta = 7$$

$$|E| = 4(n-1)n$$

$$|V| = n^2$$

Time Complexity Comparison between Dijkstra and quantum annealing for non-negatively weighted graphs



At crossover point lattice size $(n \times n)$ is approximately,

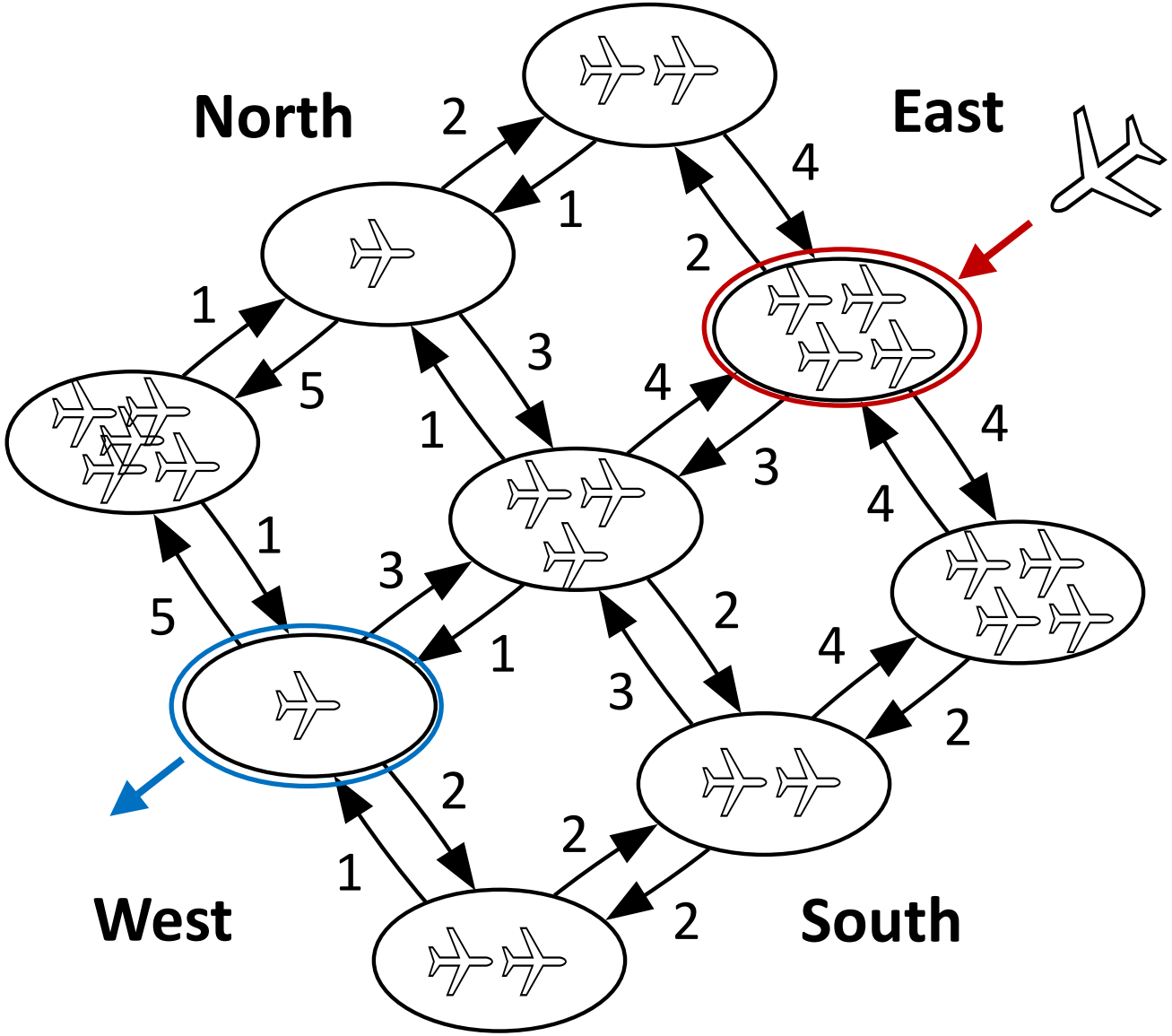
$n \approx 16375$

or

1.1 billion edges/qubits!

$$|E| = 4(n-1)n$$

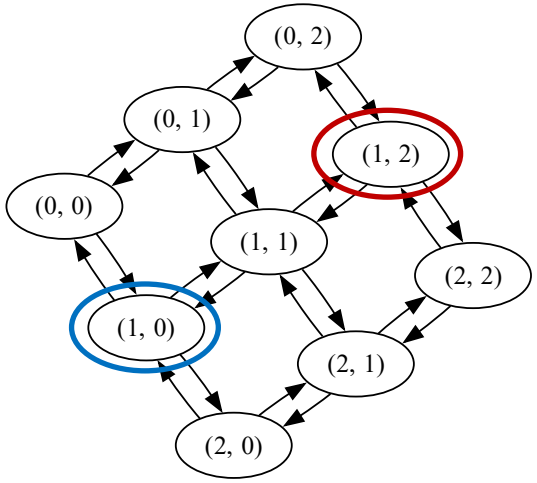
ATC Simulation Snapshot



Performance Comparison for ATC Snapshot

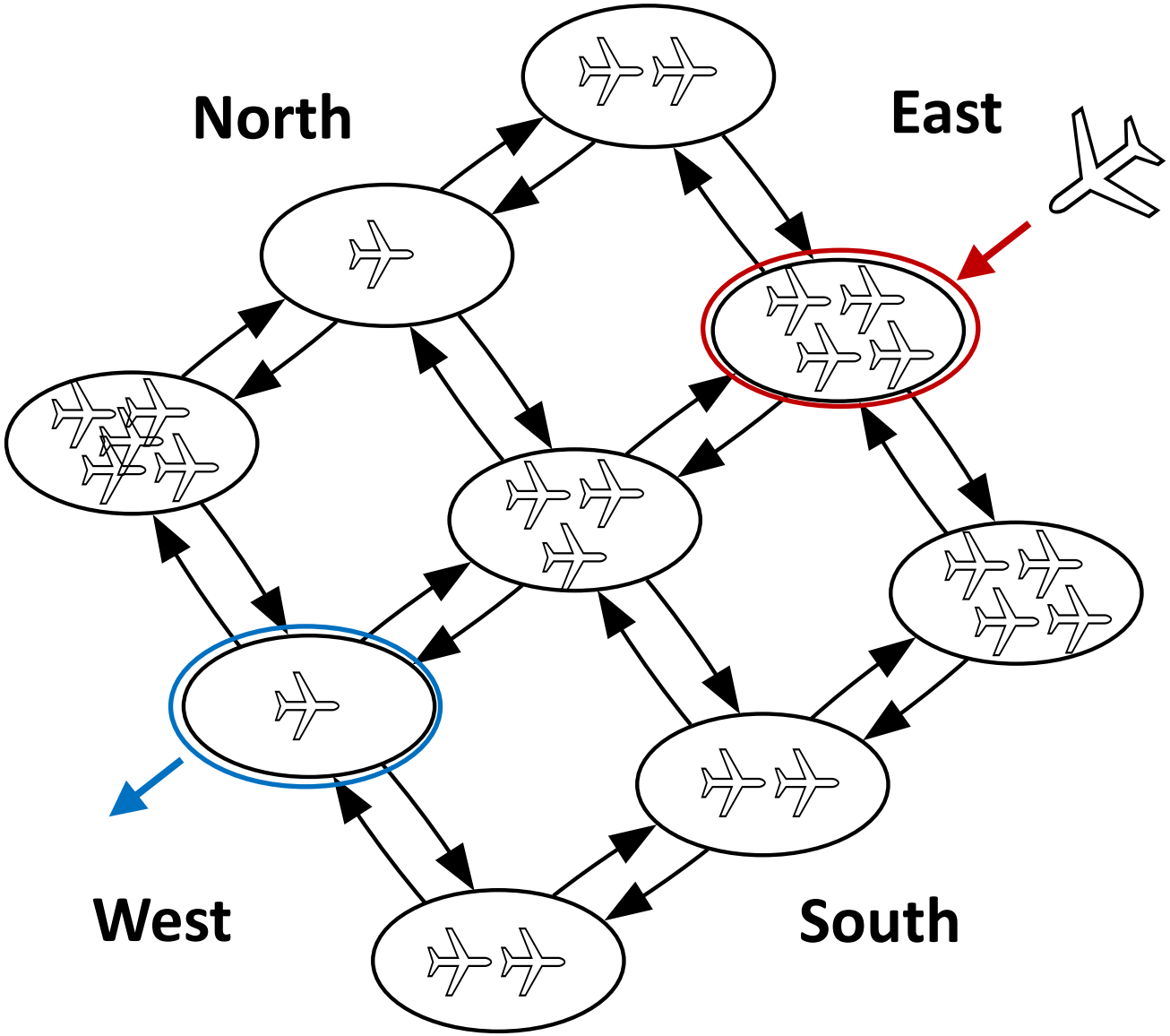
Time taken to calculate shortest path from sector (1,2) to (1,0)

Shortest Path Method	Runtime [µs]	Slowdown
Classical		
Dijkstra	9 ± 2	n/a
Bellman-Ford	15 ± 3	1.55
D-Wave Quantum Annealing		
`SimulatedAnnealingSampler`	$700{,}312 \pm 3{,}474$	~ 78,000
`Advantage_systems4.1`	$506{,}887 \pm 100{,}091$	~ 56,000



- Runtimes are average of ten runs following a single warm-up run.
- Each quantum annealing run performed 4000 samples (or reads).
  - 20 µs per sample
- The Advantage_systems4.1 QPU result is based on QPU access time.
  - QPU access time is used to calculate QPU usage charge
  - Excludes latencies due to internet comms and QPU queue wait time

ATC Airspace to QUBO Q-matrix to Problem Graph

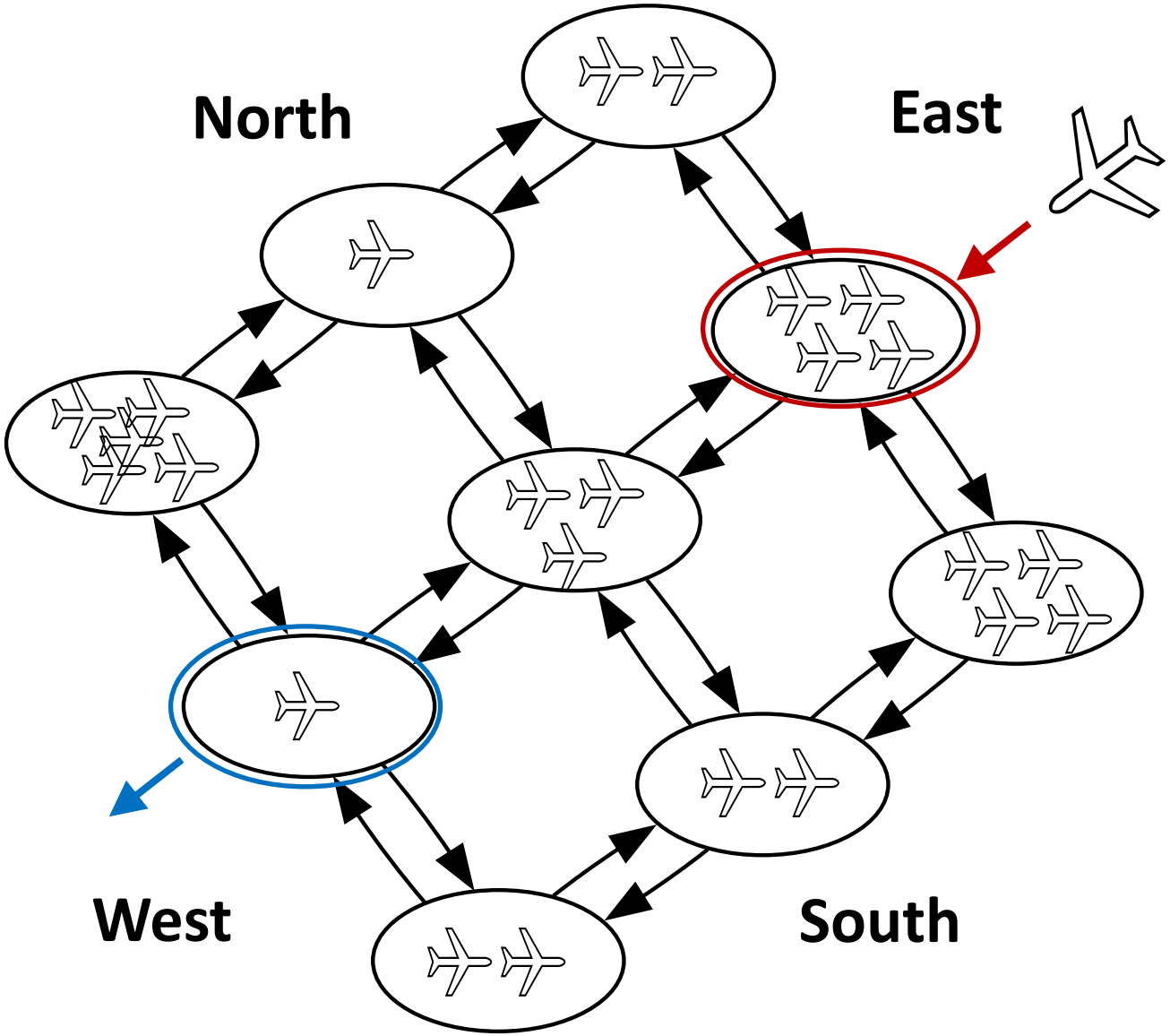


3 x 3 Airspace Lattice
- 9 vertices
- 24 edges

24 x 24 QUBO Q-matrix
- 24 diagonal terms
- 100 non-zero off-diagonal terms

Problem Graph
- 24 vertices
  - qubit biases: 1 to 253
- 100 edges
  - coupling strengths: -248 to 125
- Graph is sparsely connected
  - K_{24} complete graph has 276 edges

ATC Airspace to QUBO Q-matrix to Problem Graph



The number of edges in n x n lattice grows as n^2, but the path across the lattice grows as n.

Hence, the larger the (lattice) airspace, the sparser the shortest-path problem graph.

For a 5 x 5 lattice, a typical problem graph will have about 15% of the edges in a K_{80} complete graph.

D-Wave QPU Embedding

- For Advantage System QPUs, the qubits are linked according to the Pegasus topology.
  - D-Wave Pegasus Graph, https://docs.dwavesys.com/docs/latest/c_gs_4.html#pegasus-graph
- The Pegasus topology specifies which qubits are linked via couplers (i.e., entangled).
- There will not necessarily be a one-to-one mapping between the edges in the problem graph and qubit couplers.
- To achieve the right mapping, D-Wave uses chains, whereby a single node in the problem graph can be represented by a chain of physical qubits in the QPU.
  - A node in the problem graph is a logical qubit (representing an edge from the airspace).

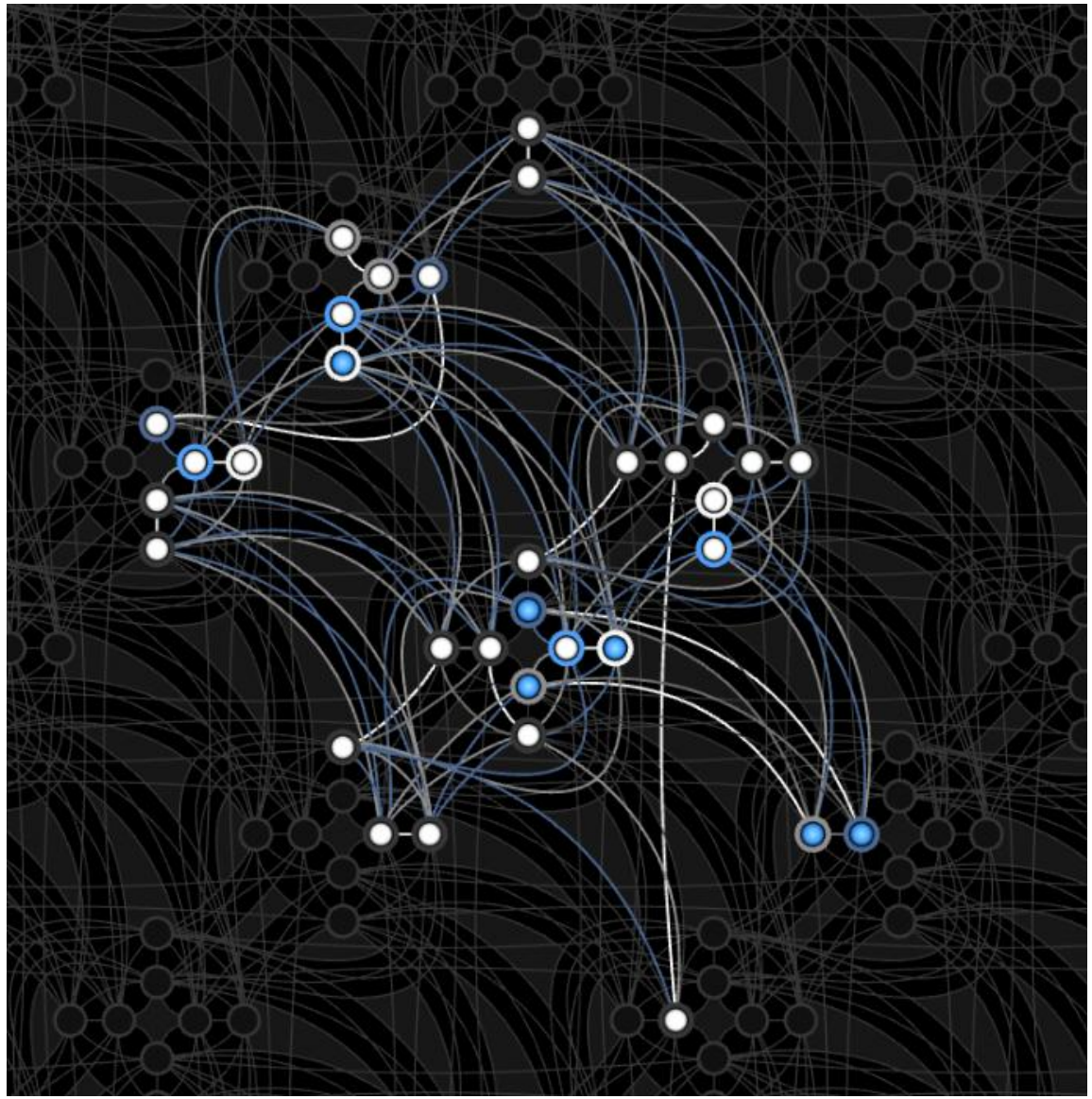
D-Wave QPU Embedding (Ocean SDK)

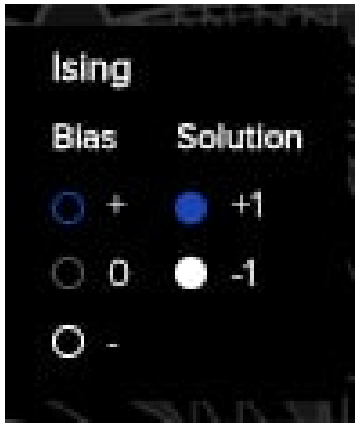
epcc

`Qdict[(i,i)]` = bias for qubit i
`Qdict[(i,j)]` = coupling strength between qubits i and j

```
import dimod
from dwave.system import DWaveSampler
import minorminer


h, J, offset = dimod.qubo_to_ising(Qdict, offset=alpha)

qpu = DWaveSampler()

emb = minorminer.find_embedding(prob_graph.edges, qpu.edgelist)
```

D-Wave Inspector: Target Embedding





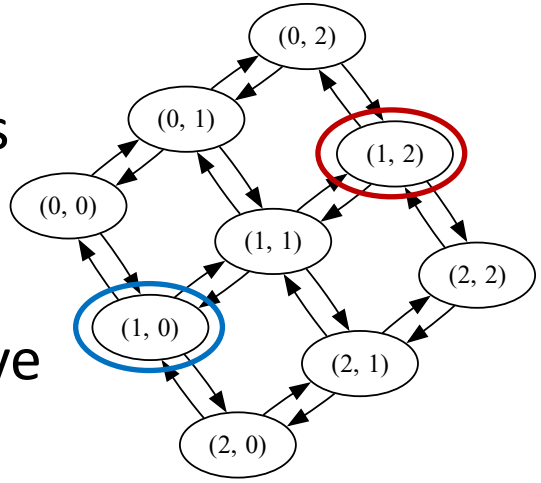
All physical qubits shown.

- 33 in this case.
- 7 two-qubit chains
- 1 three-qubit chain

- Coupling stengths for qubits within chains must be stronger than coupling strengths between logical qubits, otherwise chains may break.

- However, too high a chain strength prevents chained qubits from flipping together which may result in a non-optimal final state.

Looking for Performance Improvements



1. Use D-Wave's minorminer tool to find a minimal embedding that targets the Advantage System 4.1 QPU and requires 34 physical qubits.

2. Introduce 10 µs pause just after midway (0.55) through the anneal to give time for thermalization.

3. Reduce the number of reads from 4000 to 2000.

4. Extend the anneal time from 20 to 100 µs + 10 µs pause.
   - Total anneal time increases from 80,000 to 220,000 µs

Average time to solution (over 3 runs) is now 501,883 $\pm$ 142 µs compared to 506,887 $\pm$ 100,091 µs from before – deviation is much reduced because the same 34-qubit embedding is used for each read/sample,
but not much benefit otherwise.

Does this scale to larger airspaces?

Unfortunately, greater than 60 s runtimes are required to to find optimal shortest paths for 4x4 (48 qubits) and 5x5 (80 qubits) lattices.

For Developer access, D-Wave anneal times are limited to 1 s per job and 60 s per month.

If we run the larger simulations with the same parameters as for the 3 x 3 lattice, except for the embedding of course...

Lattice Size	Number of qubits		Average runtime [µs]
	Logical	Physical	
3x3	24	34	501,833 ± 142
4x4	48	83	487,373 ± 30,075
5x5	80	158	525,156 ± 572

Does this scale to larger airspaces?

Unfortunately, greater than 60 s runtimes are required to find optimal shortest paths for 4x4 (48 qubits) and 5x5 (80 qubits) lattices.

Advantage System 6.4

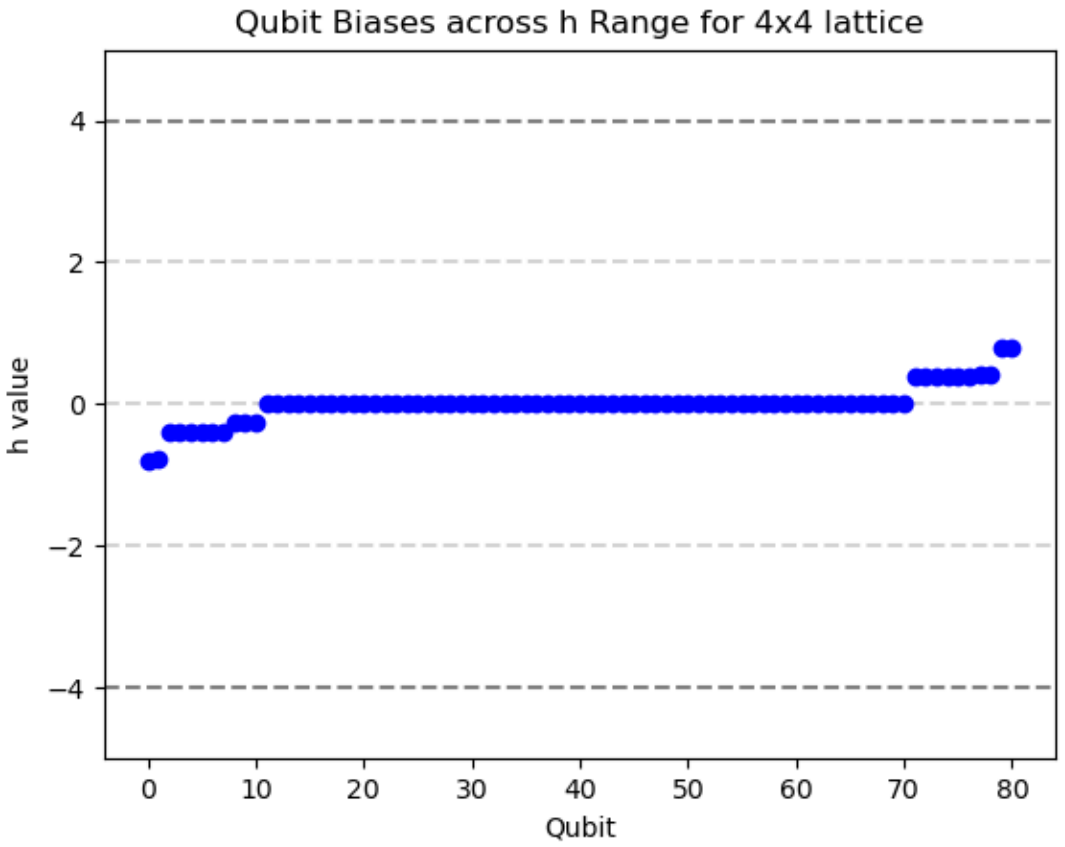
For Developer access, D-Wave anneal times are limited to 1 s per job and 60 s per month.

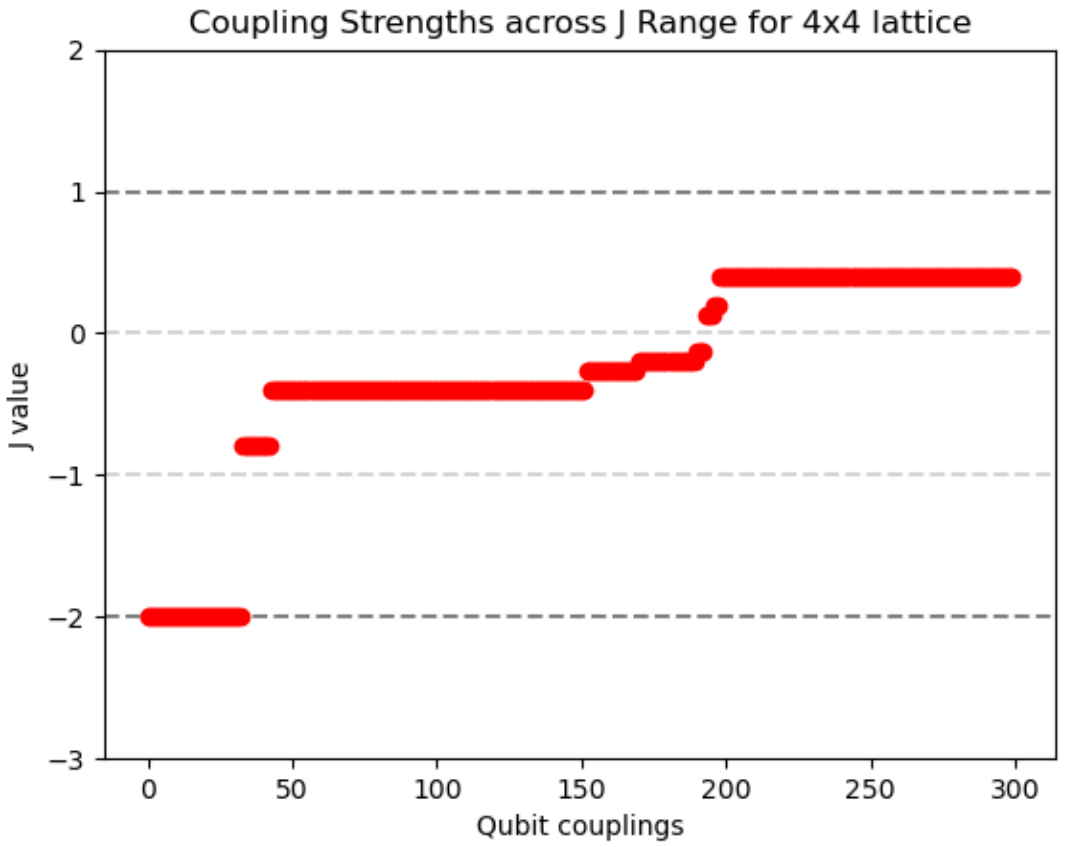
If we run the larger simulations with the same parameters as for the 3 x 3 lattice, except for the embedding of course...

Lattice Size	Number of qubits		Average runtime [µs]
	Logical	Physical	
3x3	24	34	382,766 ± 115
4x4	48	83	412,021 ± 31,309
5x5	80	158	443,135 ± 119

What's next to try?

Qubit biases and coupling strengths do not use all of available range.





What's next to try?

- Look more closely at the 48 qubit problem.
- Check sparsity for full simulation.
- Benchmark tensor network simulation on GPU.
  - 80-qubit solution takes around 45 minutes on my laptop.

Longer term:

- A shortest path problem is not a *natural* candidate for quantum acceleration…
- More complex model could include aircraft choices such as altitude and airspeed, but we of course then have a problem which scales much worse!
  - (Which is probably what you want for quantum).

Thank you!

Any questions?

CIUK 2024 Presentations

Duncan McBain (Codeplay Software)

Portable SYCL code using oneMKL on AMD, Intel and Nvidia GPUs

Abstract: In this session we will show you how to make your software run faster using oneMKL, an accelerated math library for AMD, Intel and Nvidia GPUs. oneMKL is based on the oneAPI specification and can be used to target multi-vendor and multi-architecture accelerators from a single code base. It is also now governed by the Unified Acceleration Foundation (UXL), an open governance body that is part of the Linux Foundation.

Bio: Duncan has worked with SYCL and GPUs at Codeplay Software after completing an MSc in high performance computing at the University of Edinburgh. Currently he is working with Intel's SYCL implementation, helping to deploy it on HPC systems across the world and providing support to users of Codeplay's plugins for oneAPI. He has a continued interest in how we can further leverage GPU hardware to accelerate the software that is trying to solve the biggest problems tackled by HPC systems today.

Portable SYCL code using oneMKL on AMD, Intel and NVIDIA GPUs

Duncan McBain, Codeplay

CIUK – 6th December2024

codeplay

Enabling AI & HPC To Be Open, Safe & Accessible To All

Established 2002 in **Edinburgh, Scotland**.

Grown successfully to around 100 employees.

In 2022, we became a **wholly owned subsidiary** of Intel.

Committed to expanding the **open ecosystem** for heterogeneous computing.

Through our involvement in oneAPI and SYCL governance, we help to **maintain and develop** open standards.

Developing at the forefront of **cutting-edge research**.

Currently involved in two research projects - **SYCLOPS** and **AERO**, both funded by the Horizon Europe Project.

Slide summary

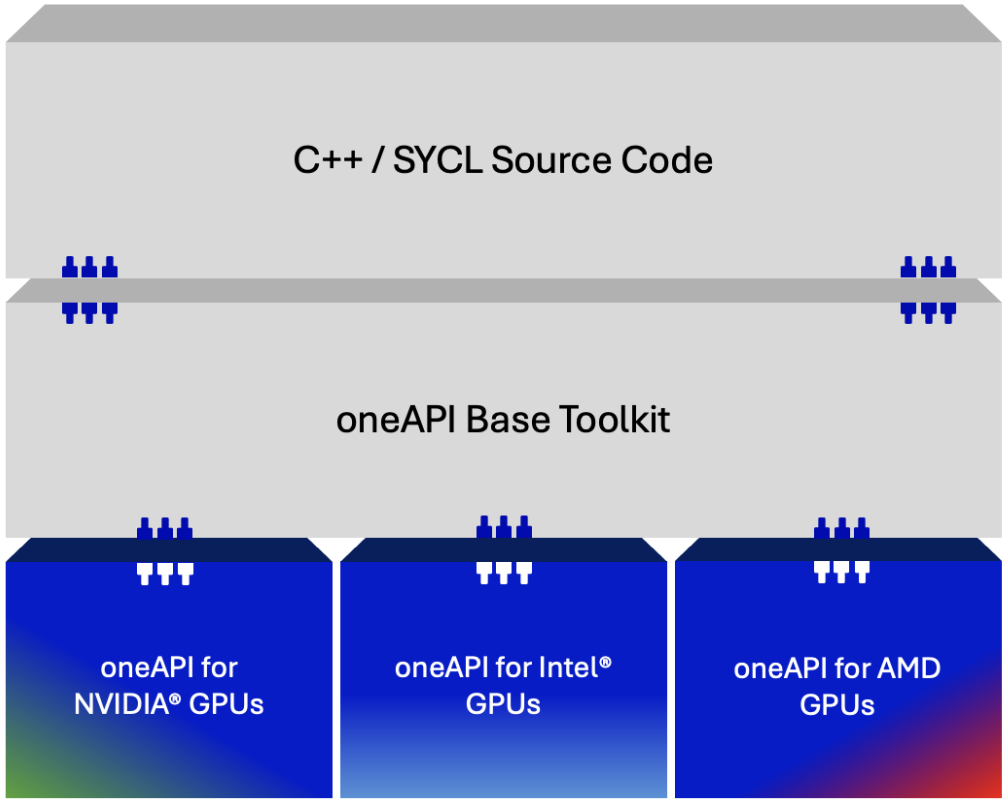
Show how to achieve **portability** of mathematical computations **across GPU vendors** using **oneMKL**

- Run SYCL on AMD and NVIDIA GPUs with oneAPI plugins
- What is oneMKL - a brief recap
- oneMKL Interface Library
  - What can it do?
  - How does it work?
  - How do you use it?
  - Building it
  - Gotchas
  - Performance

Run SYCL on AMD and NVIDIA GPUs with oneAPI plugins

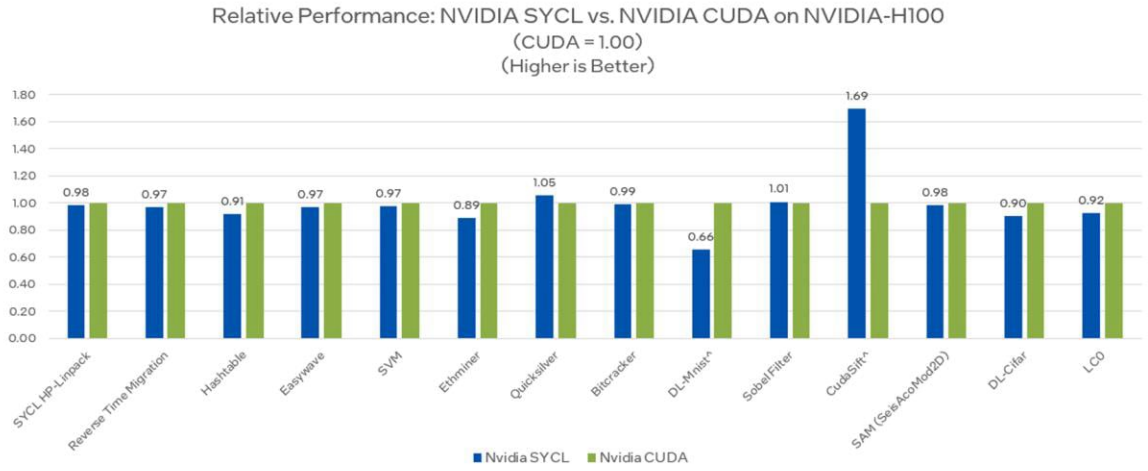
The DPC++ compiler and SYCL runtime

- Developed by Intel and Codeplay
- Extends **llvm/clang** with SYCL implementation
- Open-source version with support for **Intel/NVIDIA/AMD GPUs** + other backends available at github.com/intel/llvm
- Free proprietary version (icpx) shipped by Intel as part of the oneAPI Toolkit
  - NVIDIA/AMD GPU support available as plugins from developer.codeplay.com
  - Includes Intel's proprietary optimisation passes



SYCL performance is comparable to native CUDA/HIP

On NVIDIA GPU – SYCL Provides Comparable Performance to CUDA



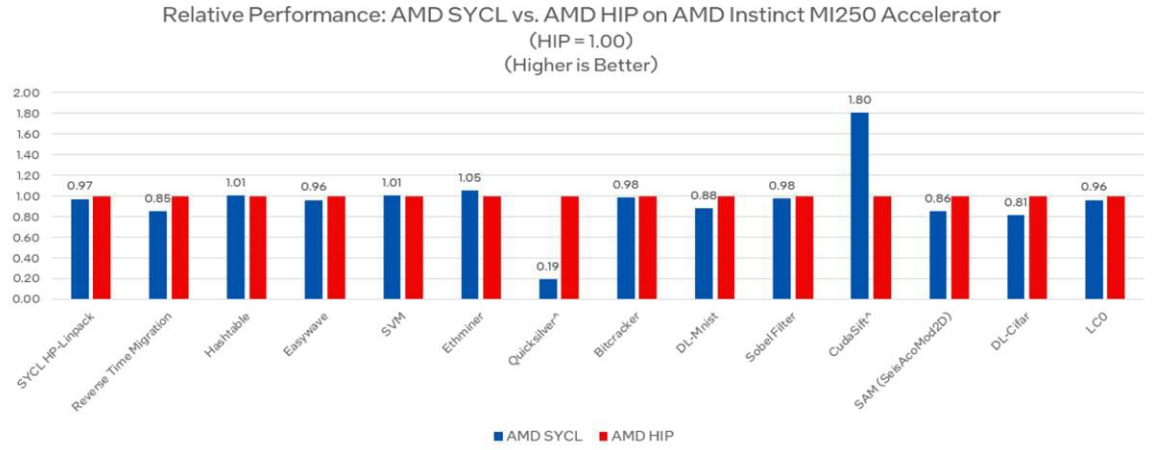
Testing Date: Performance results are based on testing by Intel as of **August 1, 2023** and may not reflect all publicly available updates.

Configuration Details and Workload Setup: Intel® Xeon® Platinum 8360Y CPU @ 2.4GHz, 2 socket, Hyper Thread On, Turbo On, 256GB Hynix DDR4-3200, ucode 0xd000389. GPU: Nvidia H100 PCIe 80GB GPU memory. Software: Velocity Bench benchmark suite branch from 8/1/23. SYCL open source/CLANG 17.0.0, CUDA SDK 12.0 with NVIDIA-NVCC 12.0.76, cuMath 12.0, cuDNN 12.0, Ubuntu 22.04.1. SYCL open source/CLANG compiler switches: -fsycl-targets=nvptx64-nvidia-cuda -Xsycl-target-backend --cuda-gpu-arch=sm_90, NVIDIA NVCC compiler switches: -O3 –gencode arch=compute_90, code=sm_90. Represented workloads with Intel optimizations.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Performance varies by use, configuration, and other factors. Learn more at www.Intel.com/PerformanceIndex. Your costs and results may vary

On AMD GPU – SYCL Provides Comparable Performance to HIP



Testing Date: Performance results are based on testing by Intel as of **August 1, 2023** and may not reflect all publicly available updates.

Configuration Details and Workload Setup: AMD EPYC 7313 CPU @ 3.0GHz, 2 socket, AMD Simultaneous Multi-Threading Off, AMD Precision Boost Enabled, 512GB DDR4, ucode 0xa001144. GPU: AMD Instinct MI250 OAM, 128GB GPU memory. Software: Velocity Bench benchmark suite branch from 8/1/23, SYCL open source/CLANG 17.0.0, AMD RoCm 5.6.0 with roc-5.6.31061, hipSolver 5.6.0, rocBLAS 5.6.0, Ubuntu 20.04.4. SYCL open source/CLANG compiler switches: -O3 -fsycl -fsycl-targets=amdgcn-amd-amdhsa -Xsycl-target-backend --offload-arch=gfx90a, AMD-ROCm compiler switches: -O3. Represented workloads with Intel optimizations.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Performance varies by use, configuration, and other factors. Learn more at www.Intel.com/PerformanceIndex. Your costs and results may vary

See our blog post for more details on these benchmark results

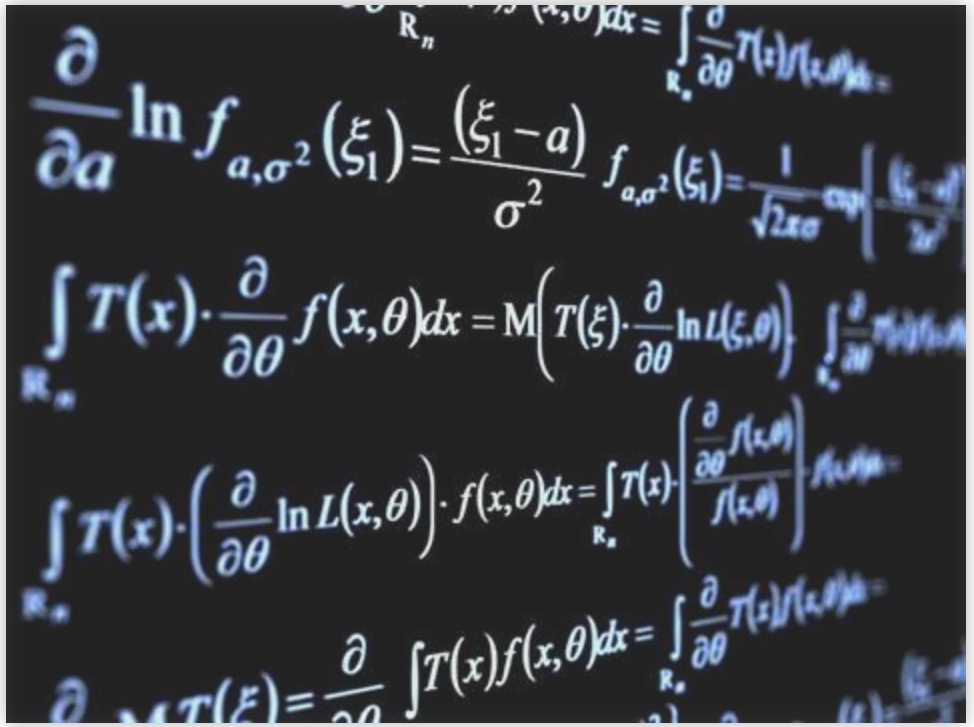
But what about numerical libraries?

The open-standard **oneAPI ecosystem** centred around SYCL delivers a solution!

You might be familiar with some of the vendor-specific GPU numerical libraries

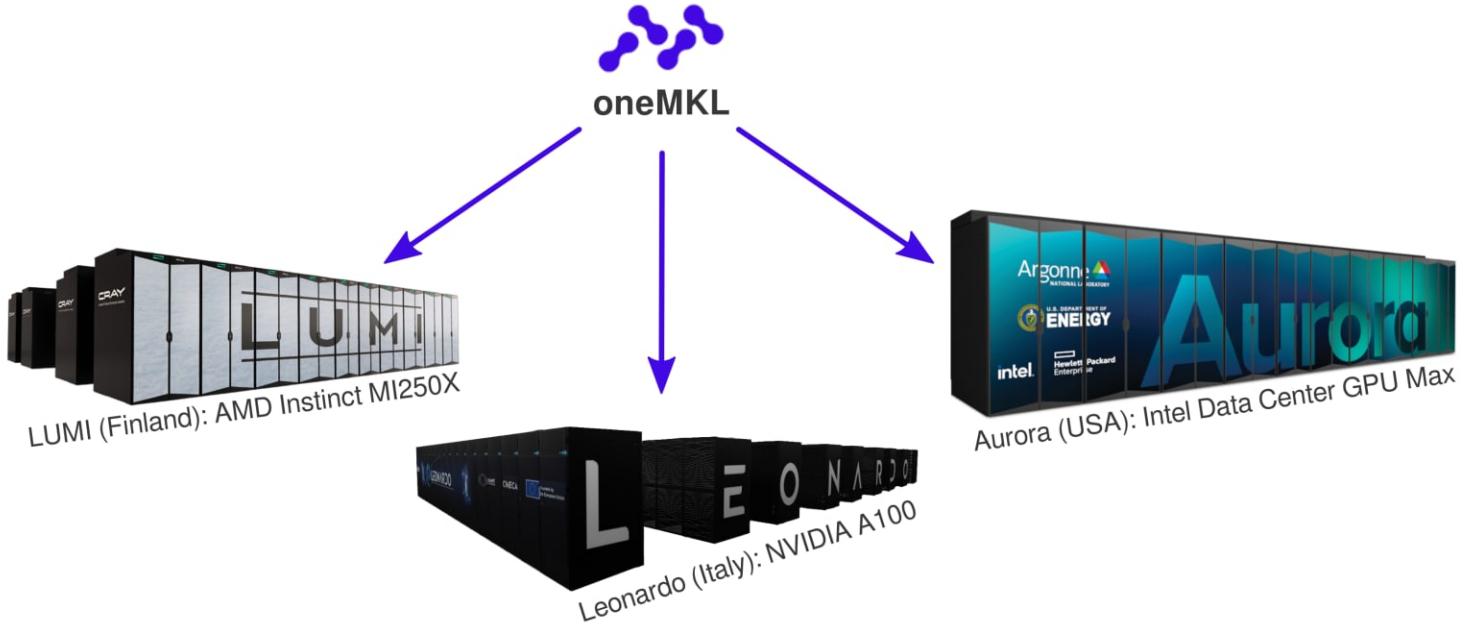
- Intel: *Math Kernels Library*
- NVIDIA: *cuBLAS*, *cuSOLVER*, *cuRAND*, *cuFFT*
- AMD: *rocBLAS*, *rocSOLVER*, *rocRAND*, *rocFFT*

Imagine being able to use all of them with single source code → **oneMKL**



oneMKL provides performance and portability

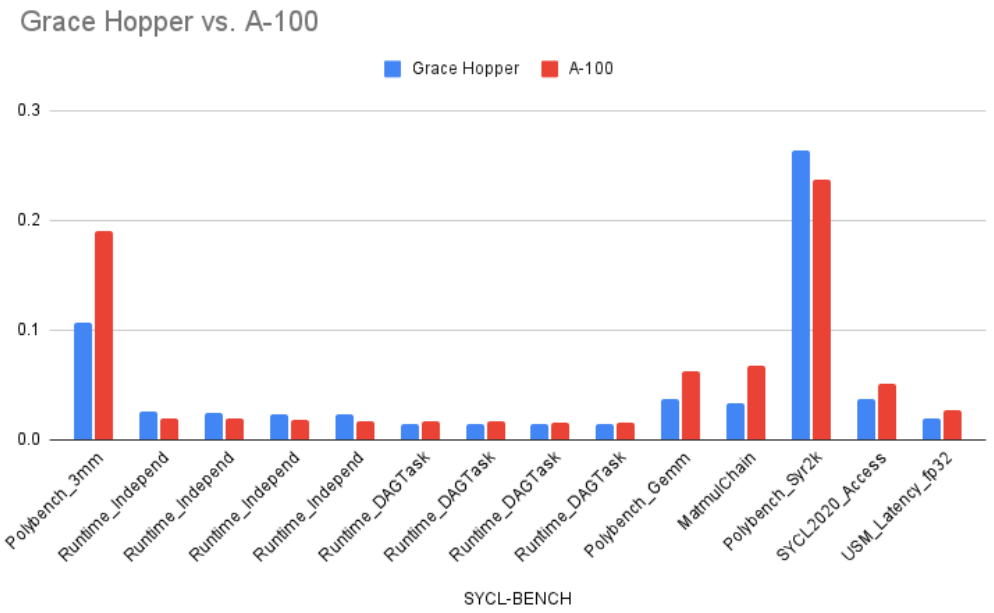
write single source code



run everywhere

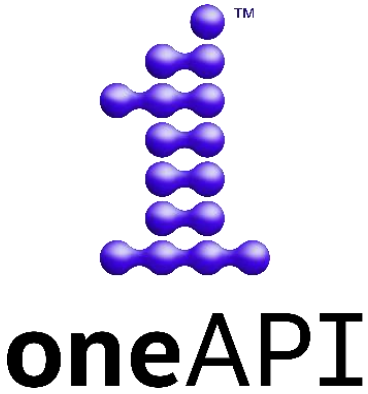
Even more cross-platform

- This year we tried oneMKL on Durham University's Bede pilot system, with Grace Hopper nodes
- Many thanks to them for this opportunity
- Build process required no changes
- All samples ran successfully



The oneAPIs and oneMKLs

What is oneAPI?



oneAPI specification

is a collection of **open-standard APIs**

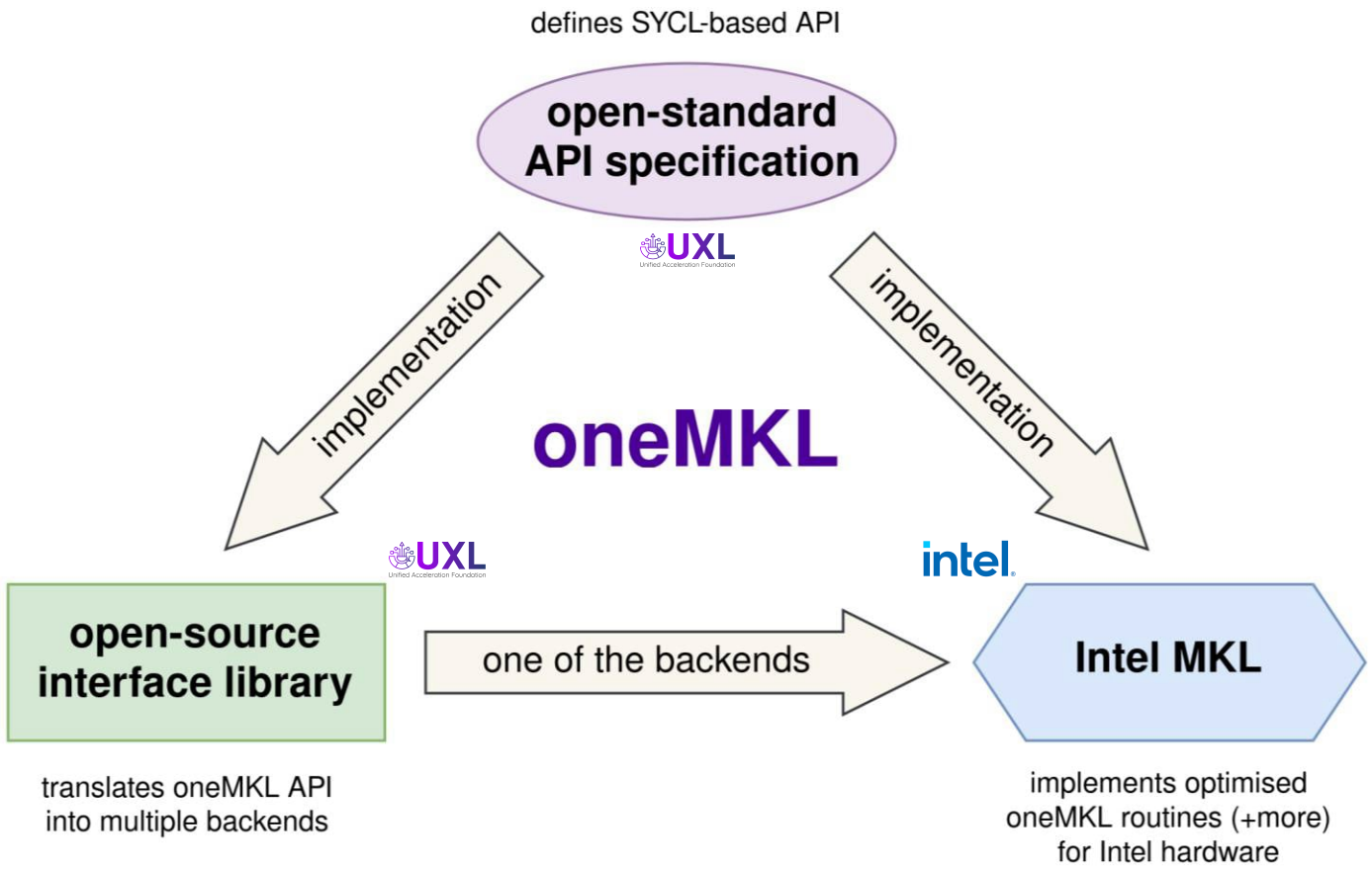
governed by

Intel oneAPI Toolkit

is a collection of **tools and libraries**

offered by

Three projects use the name oneMKL



Resources

oneAPI specification

oneapi.io

Intel oneAPI toolkits

intel.com/content/www/us/en/developer/tools/oneapi/overview.html

oneMKL specification *(part of oneAPI specification)*

oneapi-spec.uxlfoundation.org/specifications/oneapi/latest/elements/onemkl/source/

oneMKL interface library

source: github.com/oneapi-src/oneMKL

docs: oneapi-src.github.io/oneMKL

Intel oneMKL implementation *(part of Intel oneAPI toolkits)*

intel.com/content/www/us/en/developer/tools/oneapi/onemkl.html

oneMKL: capabilities

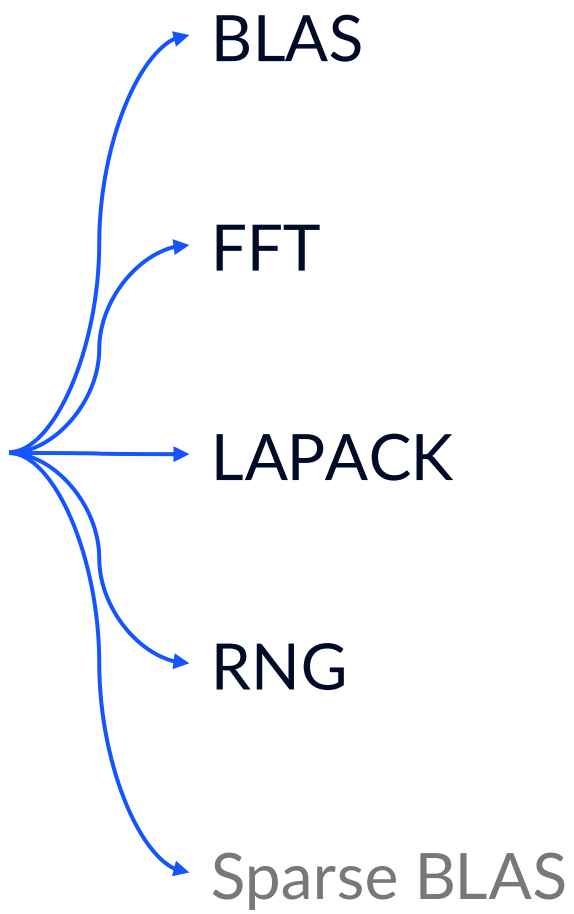
Domains

- BLAS
- LAPACK
- DFT
- RNG
- Sparse BLAS (WIP)

Not yet implemented:

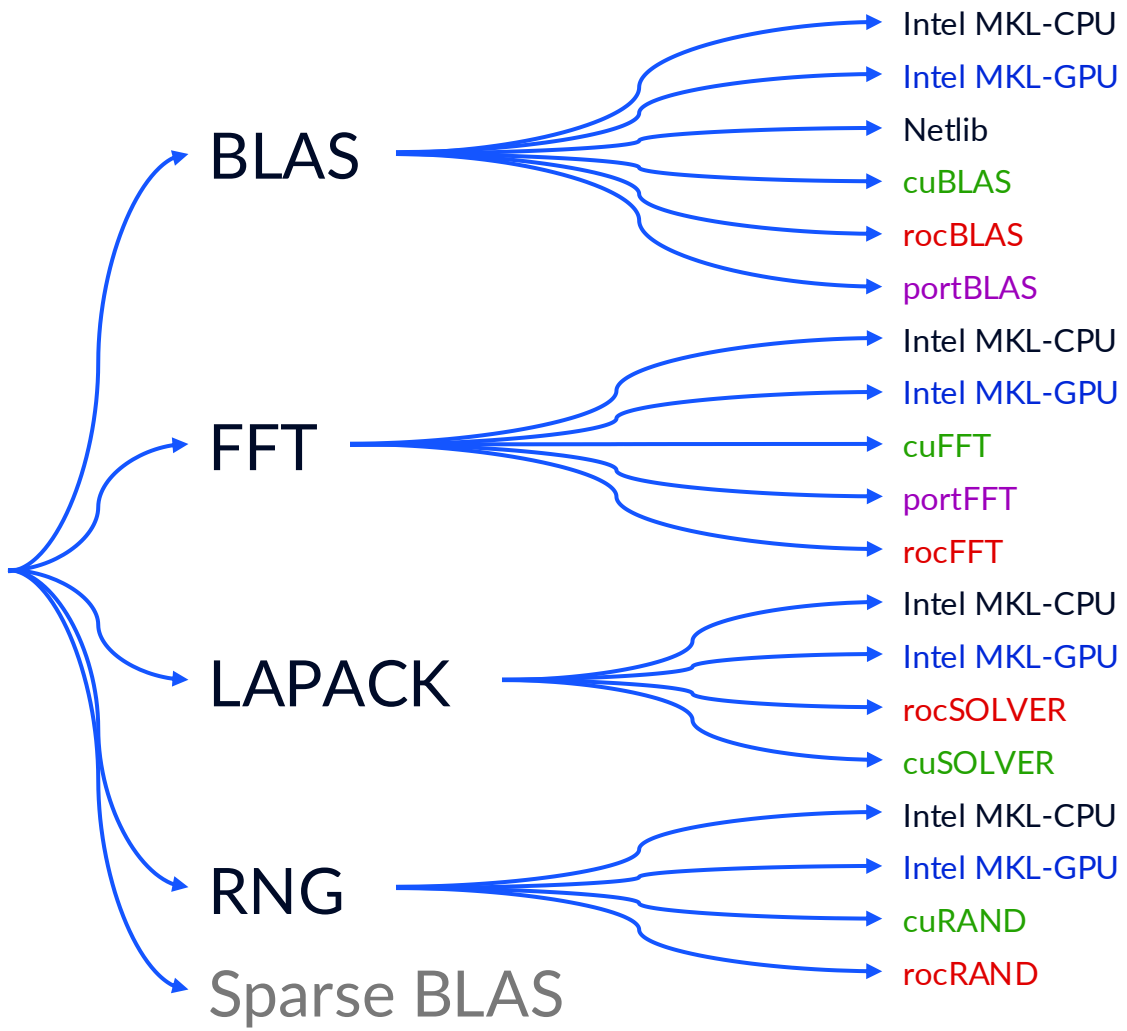
- Summary statistics
- Vector math

oneMKL Interface Library



Backends

oneMKL Interface Library



from now on, referred to as just "oneMKL"

Runtime dispatch

```
// Get a sycl::queue on any vendor’s device.
sycl::queue myQueue;

// oneMKL handles the dispatch to the
// correct backend library.
oneapi::mkl::<fn>(myQueue, ...);
```

- oneMKL can build with support for multiple vendors at once
- oneMKL can automatically dispatch to the correct backend library
- Backends are lazily `dlopen`'ed

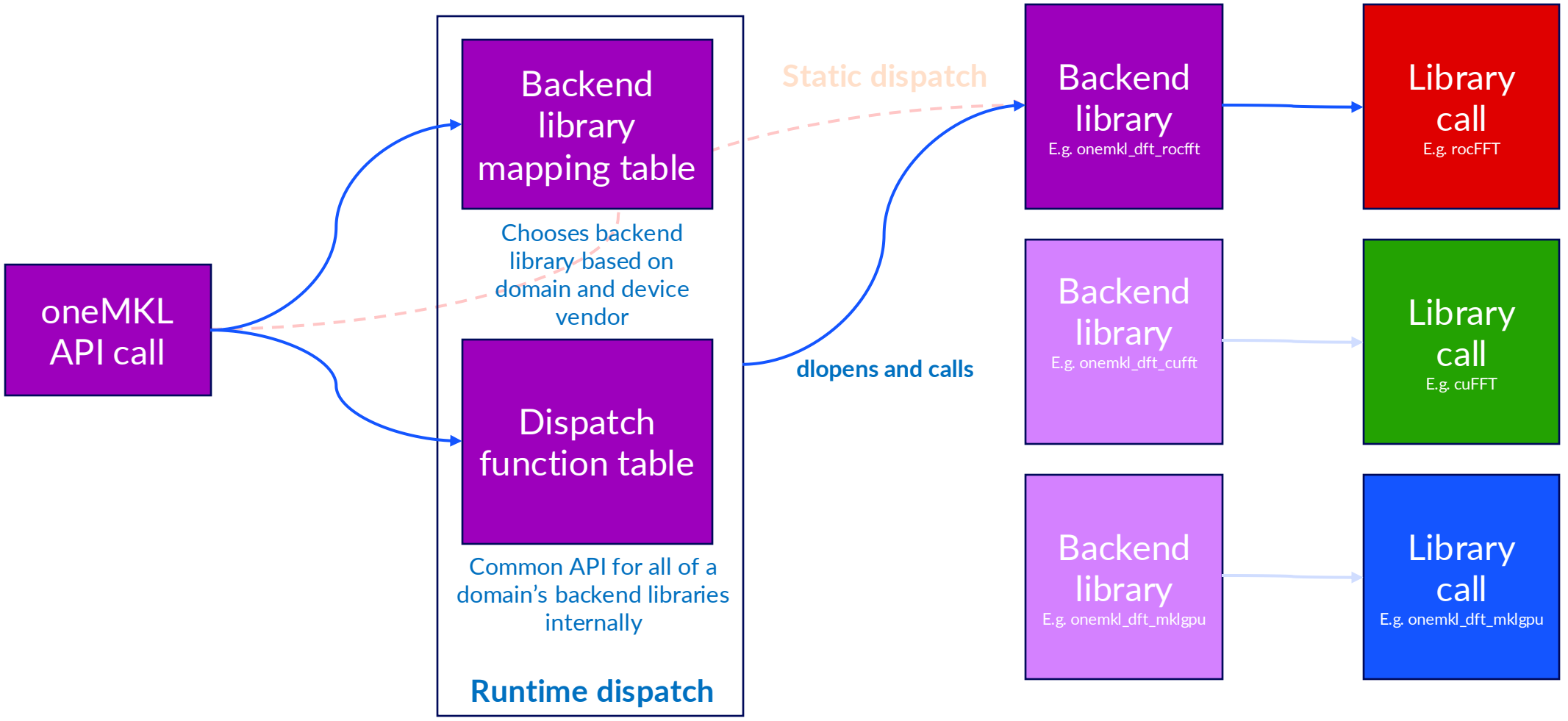
Static dispatch

```
using namespace oneapi::mkl;


// Choose a particular device
sycl::queue intelGpuQ{myIntelGPUSelector};


// Use a selector that uses a particular
// oneMKL backend
backend_selector<backend::mklgpu>
  mklgpuSelector{intelGpuQ};


// Call a backend function with the selector
oneapi::mkl::<fn>(mklgpuSelector, ...);
```

You can trade portability for avoiding the (small) overhead of dispatch tables

- Pick a specific backend in code
- Link the application directly against the used backend libraries

The runtime dispatch mechanism



Example (runtime dispatch)

Using Intel oneAPI 2024.2 and the latest oneMKL Interfaces Library, we compile:
https://github.com/oneapi-src/oneMKL/blob/develop/examples/blas/run_time_dispatching/level3/gemm_usm.cpp
and run on Intel CPU, Intel iGPU, NVIDIA GPU from **a single binary**

```
$ icpx -fsycl -o gemm gemm_usm.cpp -I../../../include -I$ONEMKL_DIR/include -L$ONEMKL_DIR/lib -lonemkl
```

```
$ ONEAPI_DEVICE_SELECTOR=opencl:cpu ./gemm

Running BLAS GEMM USM example on CPU device.
Device name is: 12th Gen Intel(R) Core(TM) i9-12900K
Running with single precision real data type:

        GEMM parameters:
            transA = trans, transB = nontrans
            m = 45, n = 98, k = 67
            lda = 103, ldB = 105, ldC = 106
            alpha = 2, beta = 3

        Outputting 2x2 block of A,B,C matrices:

            A = [ 0.340188, 0.260249, ...
                [ -0.105617, 0.0125354, ...
                [ ...


            B = [ -0.326421, -0.192968, ...
                [ 0.363891, 0.251295, ...
                [ ...


            C = [ 0.00698781, 0.525862, ...
                [ 0.585167, 1.59017, ...
                [ ...

BLAS GEMM USM example ran OK.
```

```
$ ONEAPI_DEVICE_SELECTOR=level_zero:gpu ./gemm

Running BLAS GEMM USM example on GPU device.
Device name is: Intel(R) UHD Graphics 770
Running with single precision real data type:

        GEMM parameters:
            transA = trans, transB = nontrans
            m = 45, n = 98, k = 67
            lda = 103, ldB = 105, ldC = 106
            alpha = 2, beta = 3

        Outputting 2x2 block of A,B,C matrices:

            A = [ 0.340188, 0.260249, ...
                [ -0.105617, 0.0125354, ...
                [ ...


            B = [ -0.326421, -0.192968, ...
                [ 0.363891, 0.251295, ...
                [ ...


            C = [ 0.00698781, 0.525862, ...
                [ 0.585167, 1.59017, ...
                [ ...

BLAS GEMM USM example ran OK.
```

```
$ ONEAPI_DEVICE_SELECTOR=cuda:gpu ./gemm

BLAS GEMM USM example on GPU device.
Device name is: NVIDIA GeForce RTX 3060
Running with single precision real data type:

        GEMM parameters:
            transA = trans, transB = nontrans
            m = 45, n = 98, k = 67
            lda = 103, ldB = 105, ldC = 106
            alpha = 2, beta = 3

        Outputting 2x2 block of A,B,C matrices:

            A = [ 0.340188, 0.260249, ...
                [ -0.105617, 0.0125354, ...
                [ ...


            B = [ -0.326421, -0.192968, ...
                [ 0.363891, 0.251295, ...
                [ ...


            C = [ 0.00698793, 0.525862, ...
                [ 0.585168, 1.59017, ...
                [ ...

BLAS GEMM USM example ran OK.
```

oneMKL: using it

BLAS

```
sycl::queue syclQueue;

// Your data needs to be accessible on the GPU.
auto dev_A = sycl::malloc_device<float>(sizeA, syclQueue);
// … allocate memory, give it relevant values.

// It's like the BLAS API, but taking a queue argument. The USM API returns an event.
gemm_done = oneapi::mkl::blas::column_major::gemm(syclQueue, transA, transB, m, n, k, alpha,
                                                  dev_A, ldA, dev_B, ldB, beta, dev_C, ldC);

// Wait for the work to finish.
gemm_done.wait_and_throw();
```

https://github.com/oneapi-src/oneMKL/blob/develop/examples/blas/run_time_dispatching/level3/gemm_usm.cpp

Random number generation

```
using namespace oneapi::mkl;

// A random number generator is linked to a sycl::queue
rng::default_engine engine(syclQueue, seed);
rng::uniform<float> distribution(low, high);

// Use the state we generated earlier.
auto eventOut = rng::generate(distribution, engine, n, deviceMem);

// Wait for the work to finish.
eventOut.wait_and_throw()
```

https://github.com/oneapi-src/oneMKL/blob/develop/examples/rng/run_time_dispatching/uniform_usm.cpp

DFT

```
using namespace oneapi::mkl;

// A descriptor describes the DFT you want…
dft::descriptor<dft::precision::SINGLE, dft::domain::REAL> desc(N);
desc.set_value(dft::config_param::PLACEMENT, dft::config_value::INPLACE);

// Once set, it is committed on for the chosen queue.
desc.commit(syclQueue);

// Compute the DFTs...
auto computeEvent = dft::compute_forward(desc, x_usm);

// Wait for the result
computeEvent.wait_and_throw();
```

https://github.com/oneapi-src/oneMKL/blob/develop/examples/dft/run_time_dispatching/real_fwd_usm.cpp

LAPACK

```
using namespace oneapi::mkl;
// Some APIs need scratch memory to be pre-allocated.
std::int64_t getrf_scratchpad_size = lapack::getrf_scratchpad_size<float>(syclQueue, m, n, lda);
float* getrf_scratchpad = sycl::malloc_shared<float>(getrf_scratchpad_size, syclQueue);

// ... More allocs, etc.
// LU factorization on device
auto getrfDone = lapack::getrf(syclQueue, m, n, devA, lda, dev_ipiv, getrf_scratchpad, getrf_scratchpad_size);
// Use LU factorization to solve system on device. Needs LU factorization to be complete.
auto getrsDone = lapack::getrs(syclQueue, trans, n, nrhs, devA, lda, dev_ipiv,
                               devB, ldb, getrs_scratchpad, getrs_scratchpad_size, {getrfDone});

// Wait until calculations are done
syclQueue.wait_and_throw();
```

https://github.com/oneapi-src/oneMKL/blob/develop/examples/lapack/run_time_dispatching/getrs_usm.cpp

CMake

oneMKL is installed

```
find_package(oneMKL REQUIRED)


// Link everything, runtime dispatch
target_link_library(mytarget PRIVATE MKL::onemkl)


// Link against specific backend
target_link_library(mytarget PRIVATE
          MKL::onemkl_<domain>_<backend>)
```

Using FetchContent

```
include(FetchContent)
set(BUILD_FUNCTIONAL_TESTS OFF)
set(BUILD_EXAMPLES OFF)
set(ENABLE_<BACKEND_NAME>_BACKEND ON)
FetchContent_Declare(
        onemkl_interface_library
        GIT_REPOSITORY https://github.com/oneapi-src/oneMKL.git
        GIT_TAG develop
)
FetchContent_MakeAvailable(onemkl_interface_library)

target_link_libraries(myTarget PRIVATE onemkl)
// or for a specific backend
target_link_libraries(myTarget PRIVATE onemkl_<domain>_<backend>)
```

… And add <install_dir>/lib to your **LD_LIBRARY_PATH** if it's installed in a non-standard location, otherwise dlopen doesn't work.

oneMKL: building it

The documentation makes it look harder than it is

With DPC++

```
cmake     $ONEMKL_DIR \
          -GNinja \
          -DCMAKE_CXX_COMPILER=icpx \
          -DCMAKE_C_COMPILER=icx \
          -DENABLE_MKLGPU_BACKEND=ON \
          -DENABLE_MKLCPU_BACKEND=ON \
          -DENABLE_CUFFT_BACKEND=ON \
          -DENABLE_CUBLAS_BACKEND=ON \
          -DENABLE_ROCRAND_BACKEND=ON \
          -DENABLE_FUNCTIONAL_TESTS=OFF \
          -DHIP_TARGETS=gfx90a
```

- Building isn’t that complicated
  - Enable the backends you want
  - Set HIP_TARGETS on AMD
  - Disable functional tests in most cases
- README lists some "supported" compiler/backend combinations
  - "Supported" is what we actually test with
  - Using icpx + Codeplay plugins for HIP/CUDA backends also works well and it's probably your best option
- On AMD, you can only have a single arch right now

oneMKL: gotchas

oneMKL: gotchas

- Backend libraries don't all support every feature
  - e.g. The cuFFT backend doesn't support scaling.

- Backend libraries make different guarantees
  - e.g. The rocFFT backend can modify input.

oneMKL: gotchas

- Variadic functions like
  - `desc.set_value(dft::config_param::INPUT_STRIDES, myStrides);`
  - The spec uses `int64_t`
  - Variadic arguments means that the compiler won't tell you you're wrong.

- LD_LIBRARY_PATH
  - Backend libraries need to be findable for dynamic dispatch

Coming from Intel® MKL

Intel® MKL

```
#include <oneapi/mkl/dfti.hpp>

DFTI_INPLACE
```

oneMKL Interface Library

```
#include <oneapi/mkl/dft.hpp>

oneapi::mkl::dft::config_value::INPLACE
```

oneMath: the new oneMKL Interface

Becoming oneMath

- The MKL terminology has been used by Intel for 30 years
- The open source repository is migrating to oneMath
- This was agreed via public RFC
- Changes are being merged

The repository has also migrated to the UXL Foundation GitHub
https://github.com/uxlfoundation/oneMath

The Intel library remains called Intel oneMKL

oneMKL: performance

Performance

oneMKL is a thin wrapper calling native backend libraries

- very little overhead, negligible in typical HPC use cases
  - we are working on improving the overhead for small workloads where it may be more visible
- you get **comparable performance + portability**

Let's test this with a simple GEMM example!

$$C \leftarrow alpha * op(A) * op(B) + beta * C$$

op(X) is one of **op(X) = X** or **op(X) = X^T** or **op(X) = X^H**

alpha and **beta** are scalars

A, **B** and **C** are matrices

op(A) is an **m**-by-**k** matrix

op(B) is a **k**-by-**n** matrix

C is an **m**-by-**n** matrix

Performance

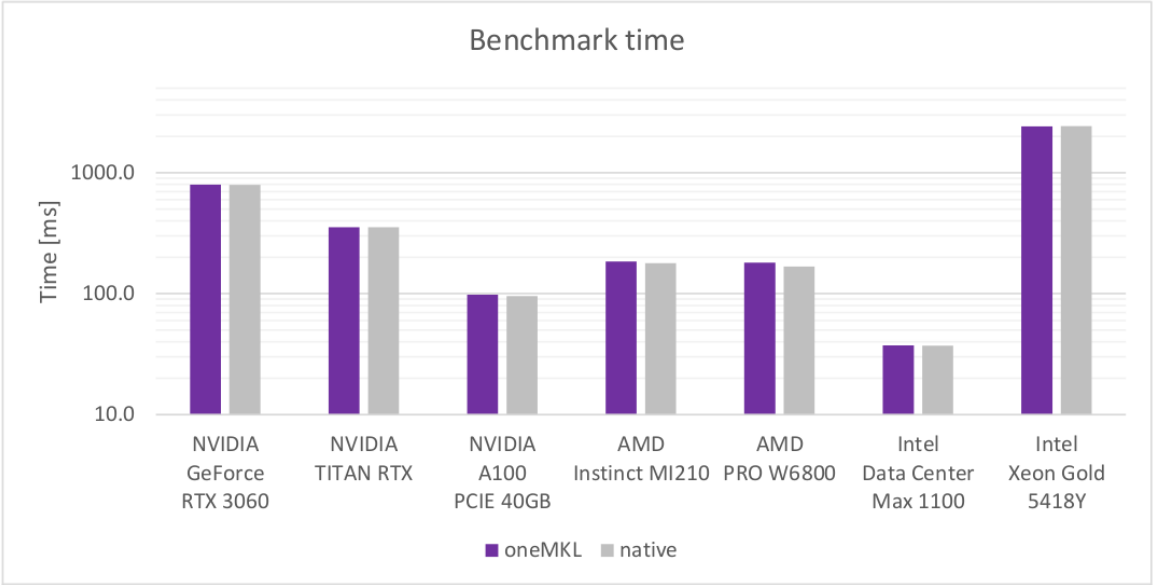
- Code from VelocityBench hplinpack DPC++ example
- We call it with double-precision matrices with {m,n,k} = {16384, 2048, 2048} filled with random values in the range 0.0–1.0
- Three code versions compiled into four executables:
  - CUDA: `cublasDgemm`
  - HIP: `hipblasDgemm`
  - IntelMKL / oneMKL (same API): `oneapi::mkl::blas::column_major::gemm`

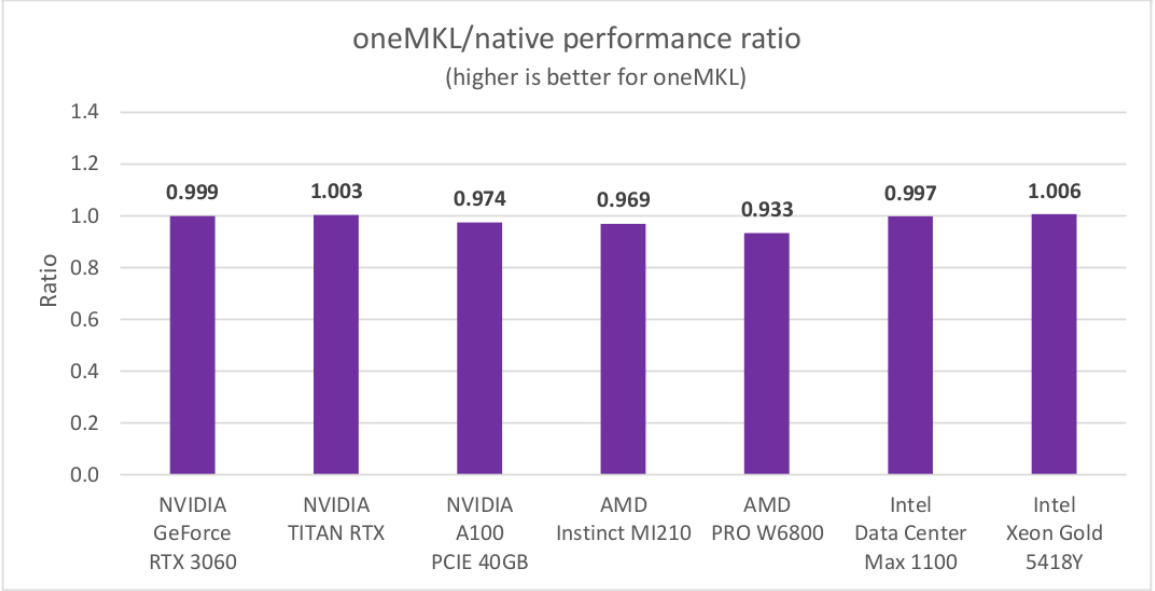
Performance

Same code runs on 7 different devices from 3 different vendors (6 GPUs and 1 CPU)

Comparable results to the native library in all cases

No need to maintain three versions of the code if just one does it!





"native" means cuBLAS on NVIDIA GPU, hipBLAS on AMD GPU and Intel MKL on Intel GPU/CPU

Summary

We learned about the **oneMKL interface library, now oneMath**:

- what it is and how it works
- its place in the oneAPI ecosystem
- how to use it
- that it provides portability and performance for your mathematical computations on GPUs from all major vendors

Your feedback

https://github.com/oneapi-src/oneMKL/issues

- Want something that oneMKL doesn't have / support?
  - Make an issue!
- Found a bug?
  - Make an issue!
- Finding something confusing?
  - Make an issue!

Issues let us justify spending time on improving oneMKL

Questions?

oneAPI Plugins for NVIDIA/AMD

Scan QR code or visit **developer.codeplay.com**

Disclaimers

A wee bit of legal

Performance varies by use, configuration and other factors.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details.

No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Performance benchmark details

Main function: https://gist.github.com/rafbiels/e93b70098d46e947ce825eb1cc95f6b5
VelocityBench dpcpp_dgemm.cpp: https://github.com/oneapi-src/Velocity-Bench/blob/50343b438e838ceae1eb11a10196d3ae90aebb67/hplinpack/dpcpp/hpl-2.3/src/dpcpp/dpcpp_dgemm.cpp

Base compilation command: `icpx -fsycl -fsycl-targets=${SYCL_TARGET} ${OFFLOAD_ARCH_FLAGS} -o onemkl main.cpp dpcpp_dgemm.cpp`

Extra flags:
oneMKL: `-lonemkl`
Intel MKL: `-DMKL_ILP64  -qmkl=parallel -qtbb`
cuBLAS: `-DUSE_CUBLAS -lcublas -lcuda -lcudart -L$(dirname $(which nvcc))/../lib64`
hipBLAS: `-DUSE_HIPBLAS -D__HIP_PLATFORM_AMD__=${HIP_TARGET} -L${ROCM_PATH}/hipblas/lib/ -L${ROCM_PATH}/hip/lib -lhipblas -lamdhip64`

Software stack: Ubuntu 22.04.4 LTS, oneAPI Base Toolkit 2024.1, CUDA 12.4, ROCm 5.4.3, oneMKL interfaces commit 6d6a7b711dbc55c49370b8ddbcc9db6e81a6ac27 + PR #490

Hardware (6 machines):
1. Intel i9-12900K CPU + NVIDIA GeForce RTX 3060 GPU
2. Intel Xeon Platinum 8268 CPU + NVIDIA TITAN RTX GPU
3. Intel Xeon Gold 6326 CPU + NVIDIA A100 PCIE 40GB GPU
4. 2x AMD EPYC 7402 CPU + AMD Instinct MI210 GPU
5. Intel i9-12900K CPU + AMD Radeon PRO W6800 GPU
6. 2x Intel Xeon Gold 5418Y CPU + Intel Data Center Max 1100 GPU

Tested on 31 May 2024

Performance varies by use, configuration and other factors.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details.

No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

CIUK 2024 Presentations

Dimitrios Bellos (The Rosalind Franklin Institute)

FlowCron: Improving access to HPC via a Function-as-a-Service which utilizes a Globus Flow and a cron service

Abstract: Over the last 10 years, data generation rates in structural biology and material science have exploded, increasing computing power requirements for techniques such as cryo-EM and X-ray CT. Many UK institutes, including the Rosalind Franklin Institute, rely on HPCs, like the Baskerville Tier 2 HPC. Our research produces multi-terabyte datasets that need to be analysed quickly, thus requiring high computing resources and fast data transfers. Furthermore, computationally demanding AI algorithms are becoming more popular to process such datasets. In the Franklin we encourage all our scientists to adopt the use of our HPC provider, Baskerville. However, making HPC use more convenient for them is a significant challenge. To tackle this challenge, we would like to showcase FlowCron, a Function-as-a-Service solution that utilises a Globus Flow and a cron service, to increase automation and minimise the number of steps necessary to transfer and process data on Baskerville.

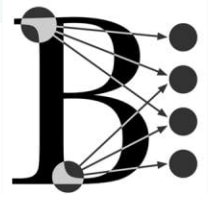
Bio: Dimitrios is a researcher and software developer, in the Artificial Intelligence and Informatics (AI & I) theme in Rosalind Franklin Institute. He is a member of the Franklin's Advanced Research Computing (ARC) team which specialises in developing data and compute infrastructure for biological scientists and offers a centralised Research Software Engineering (RSE) capability through collaboration with the Franklin and external scientists to provide excellent software for research. Furthermore, he researches into DL/ML approaches for the processing or enhancing of Electron Tomography and X-ray Computed Tomography data. He studied in the school of Electrical and Computer Engineering In the Aristotle University of Thessaloniki in Greece before completing his PhDs in Computer Science at the University of Nottingham. His main research interests are regarding the development of AI solutions with focus on denoising, classification, segmentation, deep/machine learning, computer vision and image processing.

The Rosalind Franklin Institute

FlowCron: Improving access to HPC via a Function-as-a-Service which utilizes a Globus Flow and a cron service

Dr. Dimitrios Bellos[1], Dr. James Allsopp[4], Dr. Elaine M. L. Ho[1],
Dr. Tibor Auer[5], Dr. Gavin Yearwood[2], Prof. Andrew J. Morris[3],
Dr. Mark Basham[1]

1. Advanced Research Computing, The Rosalind Franklin Institute
2. Advanced Research Computing, Baskerville HPC, University of Birmingham
3. School of Metallurgy and Materials, University of Birmingham
4. Platform Engineer, HPC and cloud infrastructure for Secure Data Environments, University of Sheffield
5. Advanced Research Computing, Brain Sciences, University College London



High-Performance Computing (HPC) clusters

Advantages & Disadvantages from non-expert HPC users

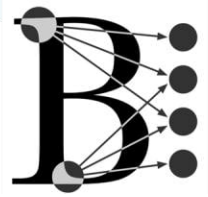
HPC Advantages

- Test multiple processing configurations in parallel
- Process multiple datasets in parallel
- Accelerate processing - Reduce time
- Use computationally expensive pipelines

HPC Disadvantages

- Transferring data to/from the HPC
- Learn terminal commands & text editors
- Monitor data and HPC jobs
- Periodically clean the HPC storage

Can something be done to remove or lessen their negative effect?



The Rosalind Franklin Institute

The importance of accelerating science in the Franklin

ARTICLES
https://doi.org/10.1038/s41594-020-0469-6
nature structural & molecular biology

Check for updates

Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2

Jiangdong Huo[1,2,3], Audrey Le Bas[2,3], Reinis R. Ruza[2], Helen M. E. Duyvesteyn[2], Halina Mikolajek[4], Tomas Malinauskas[2], Tiong Kit Tan[5], Pramila Rijal[5,6], Maud Dumoux[1], Philip N. Ward[2,3], Jingshan Ren[2], Daming Zhou[2], Peter J. Harrison[2,3], Miriam Weckener[1], Daniel K. Clare[4], Vinod K. Vogirala[4], Julika Radecke[4], Lucile Moynié[1], Yuguang Zhao[2], Javier Gilbert-Jaramillo[7], Michael L. Knight[7], Julia A. Tree[8], Karen R. Buttigieg[8], Naomi Coombes[8], Michael J. Elmore[8], Miles W. Carroll[8], Loic Carrique[2], Pranav N. M. Shah[2], William James[7], Alain R. Townsend[5,6], David I. Stuart[2,4], Raymond J. Owens[1,2,3] ✉ and James H. Naismith[1,2,3] ✉

The SARS-CoV-2 virus is more transmissible than previous coronaviruses and causes a more serious illness than influenza. The SARS-CoV-2 receptor binding domain (RBD) of the spike protein binds to the human angiotensin-converting enzyme 2 (ACE2) receptor as a prelude to viral entry into the cell. Using a naive llama single-domain antibody library and PCR-based maturation, we have produced two closely related nanobodies, H11-D4 and H11-H4, that bind RBD (K_D of 39 and 12 nM, respectively) and block its interaction with ACE2. Single-particle cryo-EM revealed that both nanobodies bind to all three RBDs in the spike trimer. Crystal structures of each nanobody–RBD complex revealed how both nanobodies recognize the same epitope, which partly overlaps with the ACE2 binding surface, explaining the blocking of the RBD–ACE2 interaction. Nanobody-Fc fusions showed neutralizing activity against SARS-CoV-2 (4–6 nM for H11-H4, 18 nM for H11-D4) and additive neutralization with the SARS-CoV-1/2 antibody CR3022.

- An example of highly impactfully publication enabled by RELION
- RELION can be accelerated by running distributed on a HPC
- Is available on Baskerville Tier 2 HPC
- Operations that took 2 weeks in a single machine can now be done in only a few days.

RELION

The Rosalind Franklin Institute

Cryogenic Electron Micrography (cryo-EM)

Biologists: Crystallography ⇨ cryo-EM

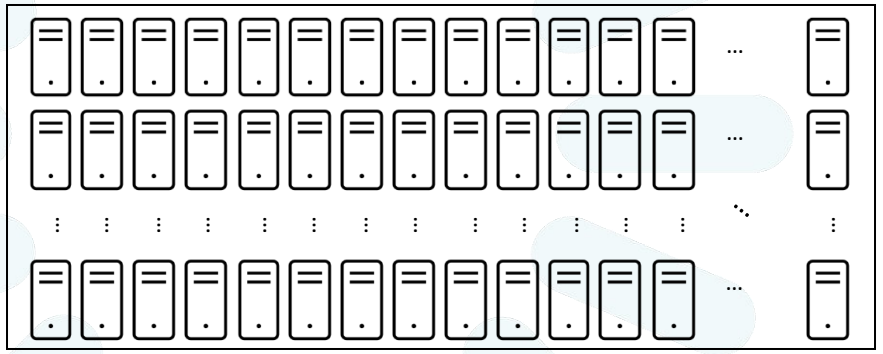
On single machines:

- ⇧ Size of Data - 200-600 TB per dataset
- ⇧ Compute resources to process
- ⇧ Time to process

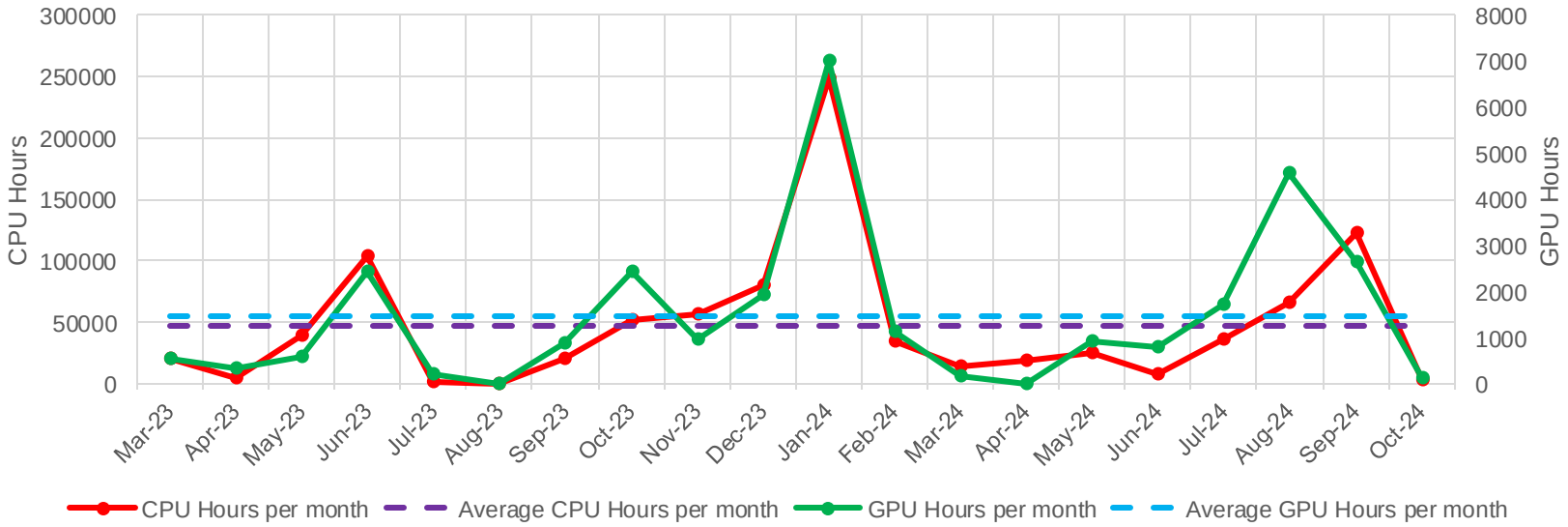
Migrate to HPC

High barrier to entry

HPC



CPU & GPU Hours per month only for RELION jobs on Baskerville HPC



RELION

Average CPU Hours per month	Average GPU Hours per month
47,588.64	1,465.767

The Rosalind Franklin Institute

Function-as-a-Service (FaaS)

for academic HPCs

User Input - Unit of Work

- Data
- Function - processing script

FaaS →

HPC

← FaaS

FaaS output

- Processed Data
- FaaS logs

FaaS bundles:

1. Data transfer to the HPC
2. Submit job
3. Data transfer back from the HPC
4. (Optional) Delete any remaining data in the HPC

The Rosalind Franklin Institute

Baskerville Tier 2 HPC cluster

228
GPUs

57
Compute Nodes

5400 TB
Storage

GPU Utilisation for 30 Nov 2024:

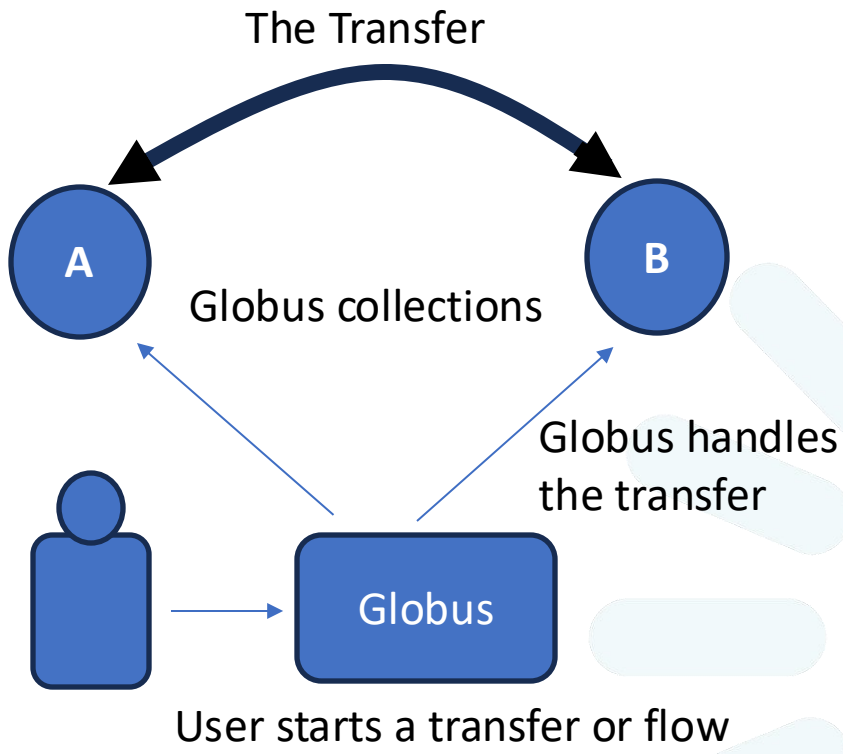
88.28%

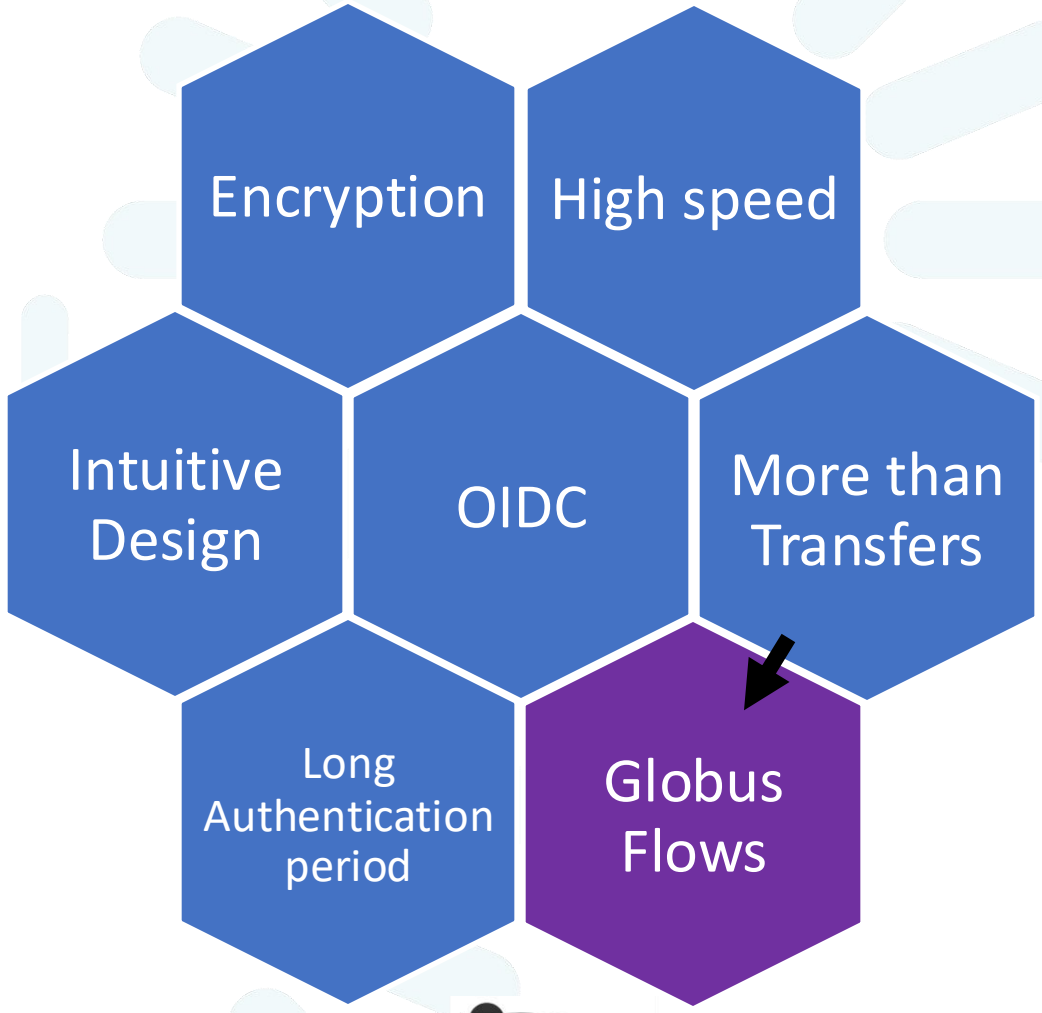
Baskerville HPC:

- 57 Nodes
- 228 Nvidia A100 GPUs
- Large Data Storage (5.4PB)
- Fast Internet connection via **a Globus collection**

The Rosalind Franklin Institute

Globus: Fast data transfers via GridFTP

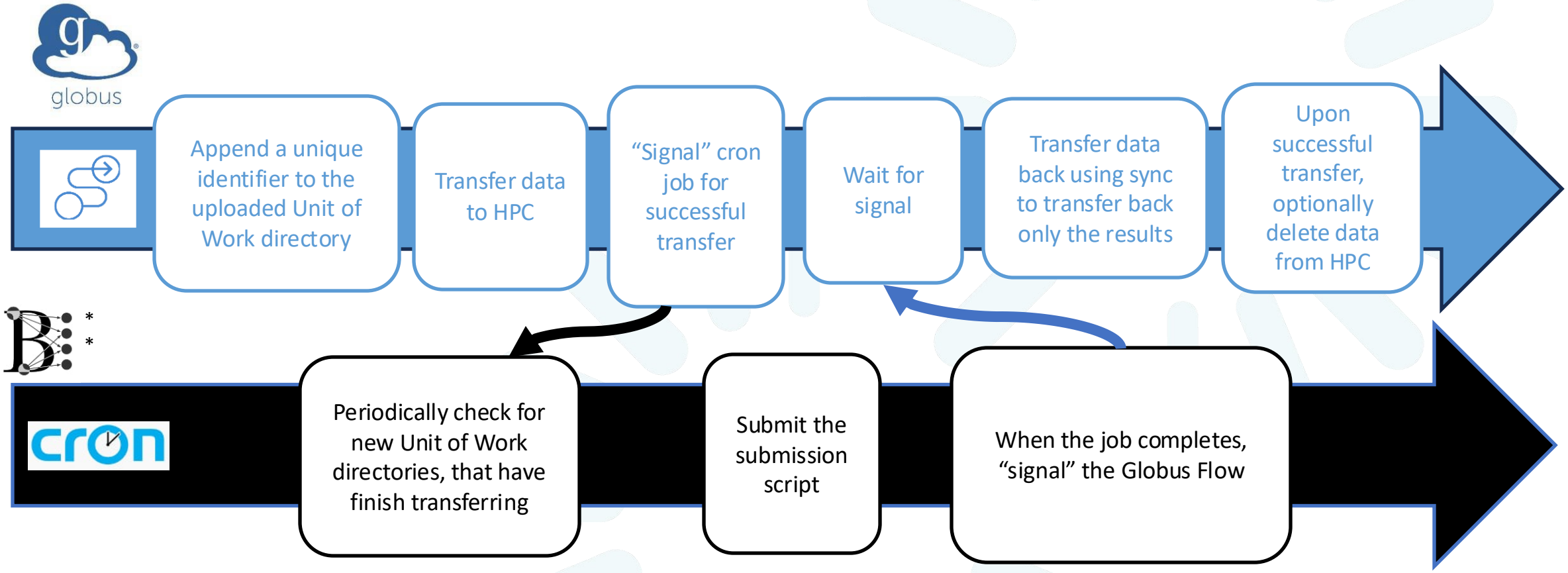




The Rosalind Franklin Institute

Globus Flow + Cron Job = FlowCron

An FaaS for fast data transfers and processing on an academic HPC



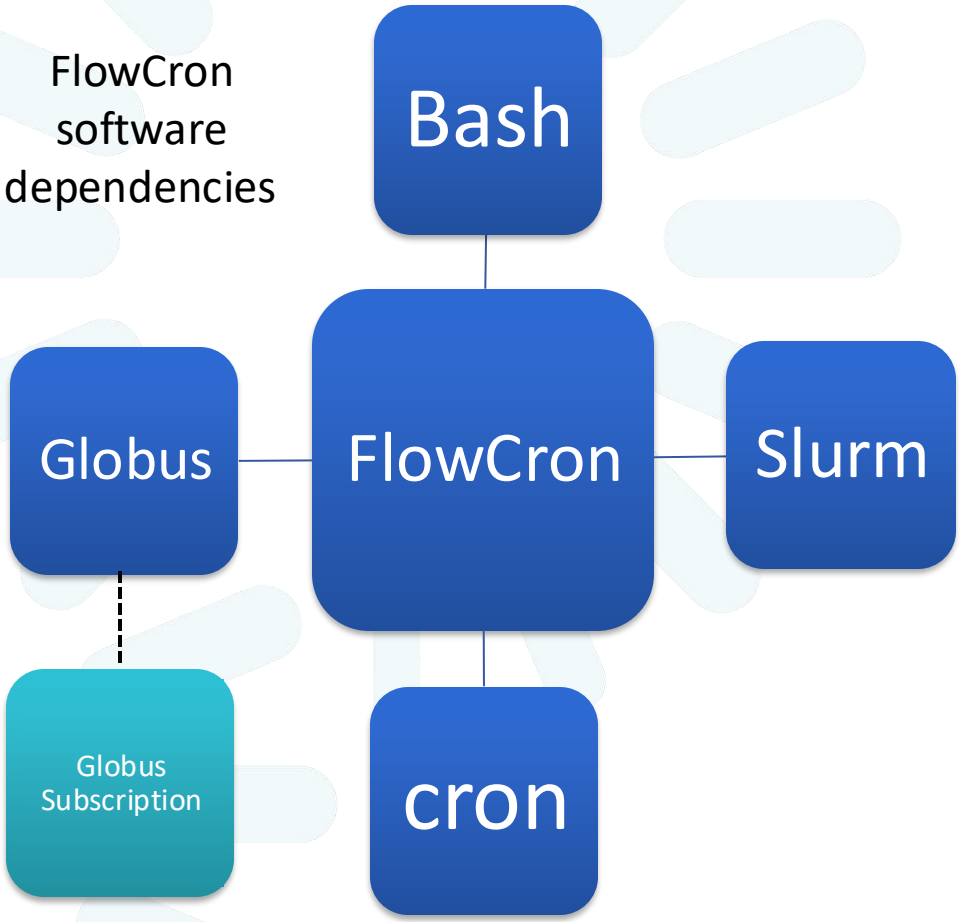
** The computations described in this research were performed using the Baskerville Tier 2 HPC service (https://www.baskerville.ac.uk/). Baskerville was funded by the EPSRC and UKRI through the World Class Labs scheme (EP/T022221/1) and the Digital Research Infrastructure programme (EP/W032244/1) and is operated by Advanced Research Computing at the University of Birmingham.

The Rosalind Franklin Institute

Why Globus Flow + Cron job

Other FaaS → Administrative costs ↑

Remote Code Execution (RCE) → Security issues ↑



The Rosalind Franklin Institute

Install/configure

One FlowCron per project per HPC

1. Install FlowCron (HPC-side) in a project directory in a HPC

2. Generate JSON files and Create Globus Flow

With Globus Subscription

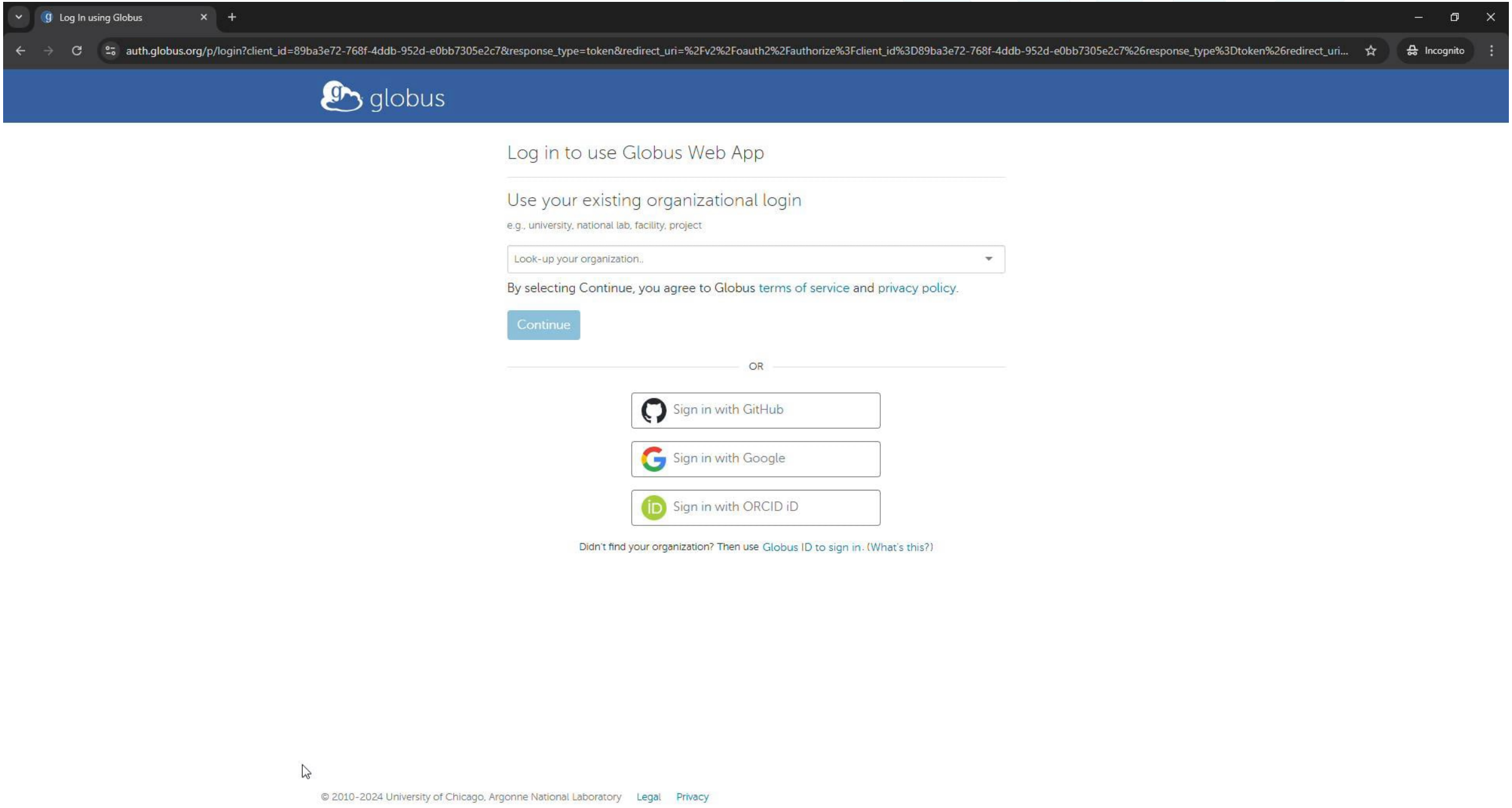
3. Create Globus group & Invite members
4. Allow Globus group to run the FlowCron's Globus Flow

Without Globus Subscription

3. Distribute Globus Flow JSON files to all users
4. Each user creates FlowCron's Globus Flow

FlowCron

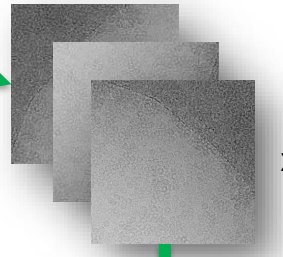
UoW submission



Processing of EM datasets

Reconstruction of EM datasets

Hours – days of data collection

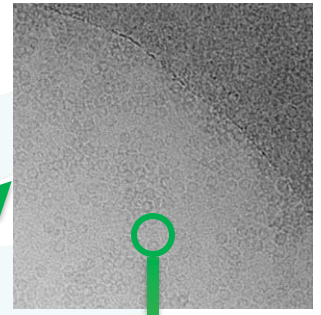


x 1,000 – 20,000

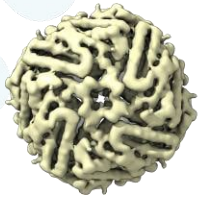
Micrographs, each with hundreds of particles

Motion Correction → CTF Estimation → Tilt-series Alignment → Inverse-Radon algorithm → Denoise Reconstructions

Use Case 1

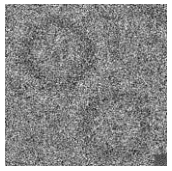


Particle analysis



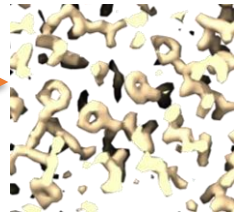
x 50,000 – 5,000,000

Particles



Use Case 2

High-resolution 3D refinement ← 3D classification ← Remove duplicate particles ← 3D Refinement ← Reconstruct Particle ← Particle 3D model generation ← Extract subtomograms ← Pick candidate particles



2.0Å Apoferritin protein structure to determine electron damage in vitreous ice

RELION

The Rosalind Franklin Institute

FlowCron

Use Case 1: Preprocessing of Cryo-EM

	Process	Size of data	Compute Resources	Time
1	Motion correction	600GB dataset	108 CPU cores	~3 hours
2	CTF estimation	600GB dataset	108 CPU cores	~10 minutes
3	Tilt-series alignment	600GB dataset	3 GPUs (A100)	~10 hours

- Including data transfers → 19 hours and 15 minutes
- No further human attention

RELION

The Rosalind Franklin Institute

FlowCron

Use Case 2: 3D Classification of Cryo-EM (benchmark)

	Process	Size of data	Compute Resources	Time
1	3D classification Only 5 out of 25 iterations	50GB dataset	2 GPUs	~1 hour

- Including data transfers → 56 minutes
- No further human attention

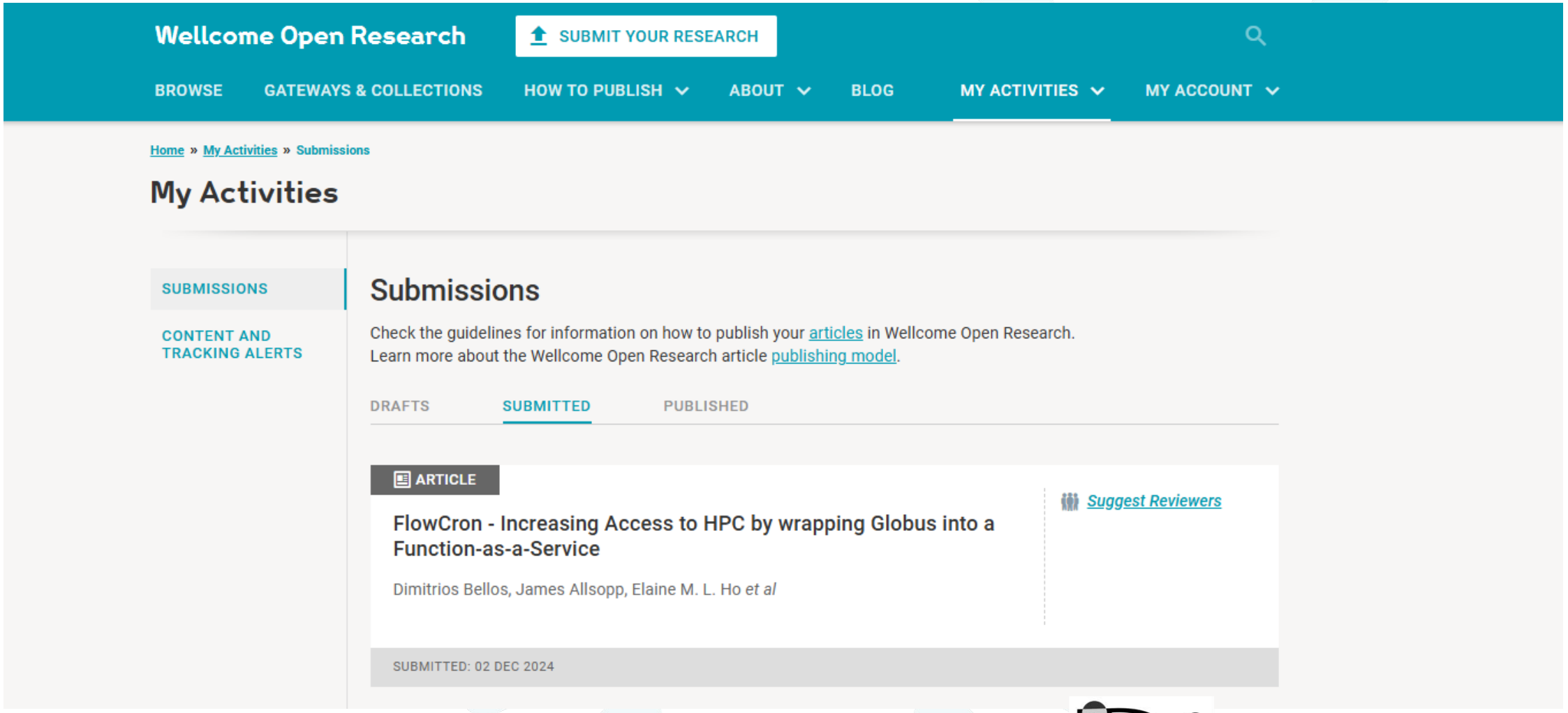
Under normal conditions

	Process	Size of data	Compute Resources	Time
1	3D classification	600GB dataset	2 GPUs	~25 hours

RELION

The Rosalind Franklin Institute

Manuscript already under consideration



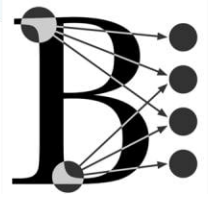
The Rosalind Franklin Institute

Conclusion

- FaaS to academic HPC
- Globus - fast data transfers
- Easy troubleshooting, maintenance, support
- Promotes reproducibility
- Ubiquitous software dependencies
- Simple Install & Easy Use

Future plans

- Globus Flow: Select different destination to return of processed data
- Globus Flow: Optional return of processed data
- Globus Flow: Sync slurm output files to destination collection
- Cron Job: Convert to use scron



The Rosalind Franklin Institute

Acknowledgement

Advanced Research Computing, Rosalind Franklin Institute

Dr Mark Basham

Dr Laura Shemilt

Dr Elaine M. L. Ho

Advanced Research Computing, Baskerville HPC

Dr Gavin Yearwood

Dr Warrick H. Ball

Dr Simon Branford

Announcement: New RCE positions for the ARC team in the Franklin coming soon!

Advanced Research Computing, Brain Sciences, University College London

Dr Tibor Auer

Dr James Allsopp

Platform Engineer, HPC and cloud infrastructure for Secure Data Environments, University of Sheffield

FlowCron Beta testers, Rosalind Franklin Institute

Gabryel Mason-Williams

Dolapo Adebo

School of Metallurgy and Materials, University of Birmingham

Prof Andrew J. Morris

The HPC Hardware Lab @Durham

Durham University

Alastair Basden, Peter Draper, Mark Lovell, Fawada Qaisir, Richard Regan, Paul Walker

Durham University
Institute for Computational Cosmology

DiRAC
High Performance Computing Facility

ExCALIBUR

COSMA

- HPC Hardware Lab hosted by:
  - the COSmology MAchine
    - Est. 2001
    - The DiRAC Memory Intensive service
      - primarily covering STFC science
    - Alongside the 3 other DiRAC systems
  - 3 generations currently in operation
  - Newest being COSMA8
    - ~70k cores, 0.5PB RAM, 20PB storage, 26PB tape
    - Bespoke design appropriate for main workloads

DiRAC

Alastair Basden
DiRAC / Durham University

HPC Hardware Lab @Durham

- Mission: Provide access to the latest HPC hardware to users and DRI/RTP professionals from across the UK
  - For code testing, performance tuning and debugging
  - To advise on purchase of future technologies
  - To allow informed decisions to be made whenever funding appears
- Accessible to all
  - Single authentication system to lots of novel hardware
- An emphasis on production/going live
  - No shelf-ware!

Alastair Basden
DiRAC / Durham University

History

- Came together almost by accident
  - 2019: Intel provide a 56-core 6TB Cascade Lake system
    - For testing Non-volatile DIMM performance (Apache Pass)
  - 2019: University funding for the DINE cluster
    - 16 node BlueField DPU test system
  - 2019: ExCALIBUR announced, with a H&ES component
    - Hardware and Enabling Software: $4.5m
    - Perfect timing for the hardware lab

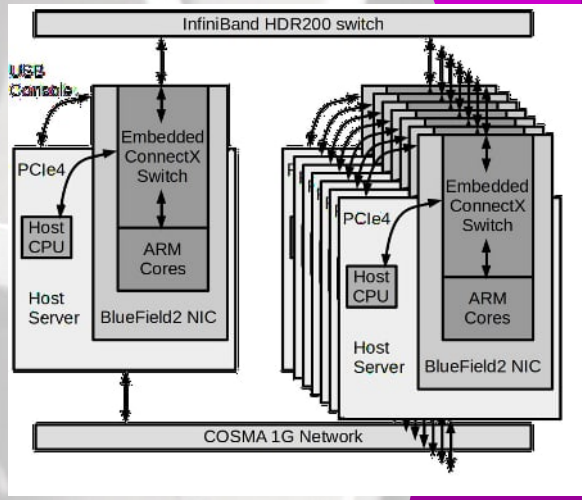
Hardware Lab Components

- DINE: Durham Intelligent Networking Environment
- DINE2: Durham Integrated Next-gen Environment
- CPU compute
  - X86, ARM
- GPU compute
  - NVIDIA, AMD, Intel
- Composability
- Networking
- Storage laboratory
- Environmental
  - Solar panel installation
  - Immersion cooling
  - Heat storage
  - Logging and awareness
- Quantum

Alastair Basden
DiRAC / Durham University

DINE

- Durham Intelligent Network Environment
  - A 24 node (initially 16) system for investigation of networking technologies
    - And other things
  - 32 cores and 512GB RAM per node, AMD ROME
- Has hosted (since 2019):
  - BlueField-1
  - BlueField-2 (host-separated and embedded modes)
  - Rockport 100G 6D Torus Ethernet network
- Investigating feasibility of charging by kWhr from April
- Funded by Durham, DiRAC, ExCALIBUR



Alastair Basden
DiRAC / Durham University

DINE-2

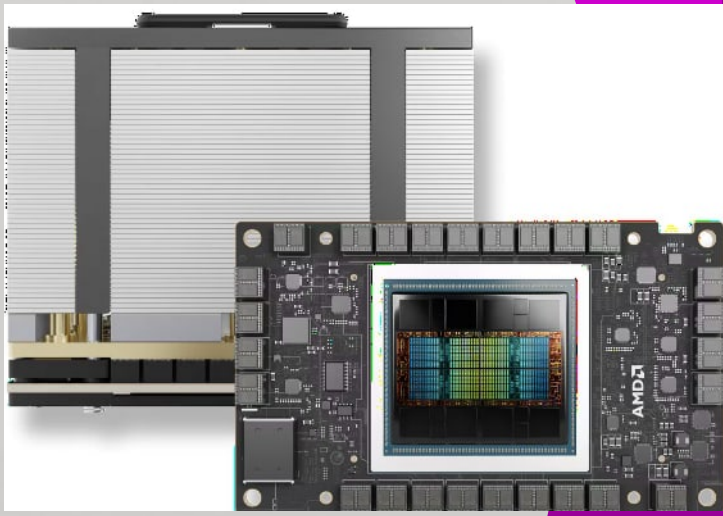
- Durham Integrated Next-gen Environment
  - 8 node Intel Sapphire Rapids system
    - 64 cores, 2TB RAM per node
  - Currently hosts a CerIO composable PCIe-5 fabric
    - 8x NVIDIA GPUs, assignable in any number to any host
  - Coming soon: 400G Ethernet fabric
- Also part of the kWhr charging/allocation study
- Funded by DiRAC, IRIS and SKA

CPU compute

- Providing users with access to cutting edge CPU technologies:
- Coming soon: AMD Turin
- AMD Genoa and Bergamo
- NVIDIA Grace
- Intel Emerald Rapids, Sapphire Rapids, Ice Lake
- AMD Milan-X (extreme cache version: 768MB L3 cache)
- AMD Milan, Rome
- Intel Cascade Lake (with Apache Pass RAM, 6TB)
- Funded by OEMs, Intel, AMD, Dell, DiRAC, ExCALIBUR

GPU Compute

- Access to small numbers of latest (and not-so latest) GPUs
- AMD MI300X
  - 8x MI300X, cache coherent global address space
  - And coming soon MI300A
- AMD MI210, MI100, MI50
- NVIDIA H100 (Grace-hopper system)
  - PCIe version also coming soon, X86
- NVIDIA A100, A30, V100
- Intel Ponte Vecchio
- Direct ssh (including Jupyter) and queue-based access
- Funded by Dell, AMD, Intel, DiRAC, ExCALIBUR, IRIS

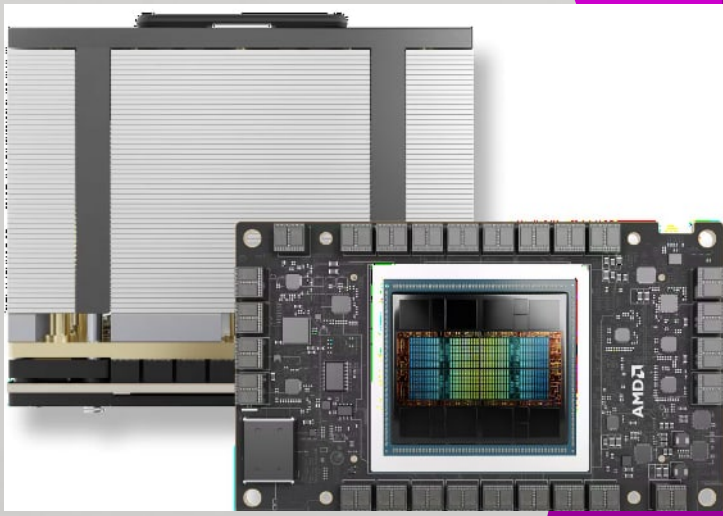


Credit: AMD

Alastair Basden
DiRAC / Durham University

GPU Compute

- Access to small numbers of latest (and not-so latest) GPUs
- AMD MI300X
  - 8x MI300X, cache coherent global address space
  - And coming soon MI300A
- AMD MI210, MI100, MI50
- NVIDIA H100 (Grace-hopper system)
  - PCIe version also coming soon, X86
- NVIDIA A100, A30, V100
- Intel Ponte Vecchio
- Direct ssh (including Jupyter) and queue-based access
- Funded by Dell, AMD, Intel, DiRAC, ExCALIBUR, IRIS



Credit: AMD

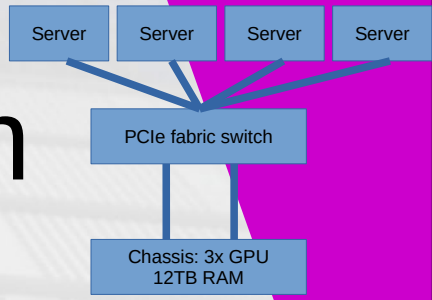
Alastair Basden
DiRAC / Durham University

Composability

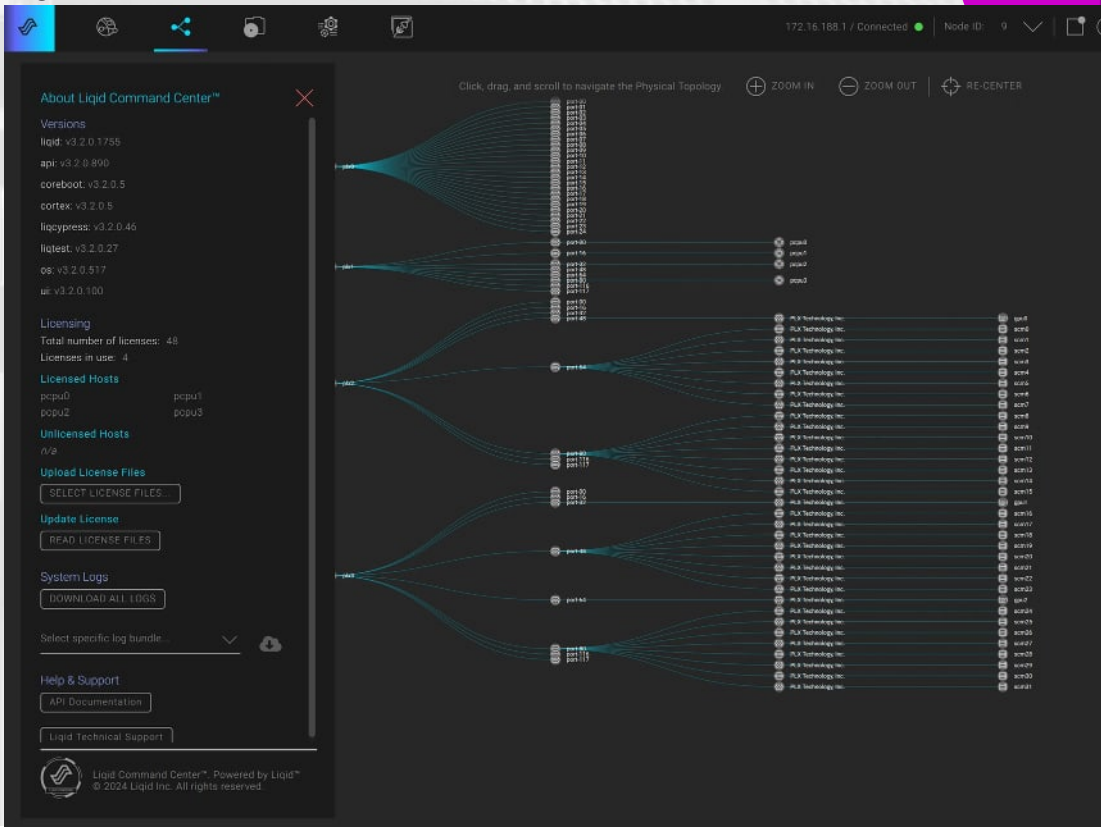
- Infrastructure-as-a-service
  - The ultimate goal for commercial cloud-type systems
  - Is it relevant for HPC?
    - How does performance suffer?
    - Is it stable?
    - What are the use cases?
- 2 composable systems
  - And a Gen-Z test system
    - Defunct: technology transferred to CXL

Alastair Basden
DiRAC / Durham University

Liqid composable system



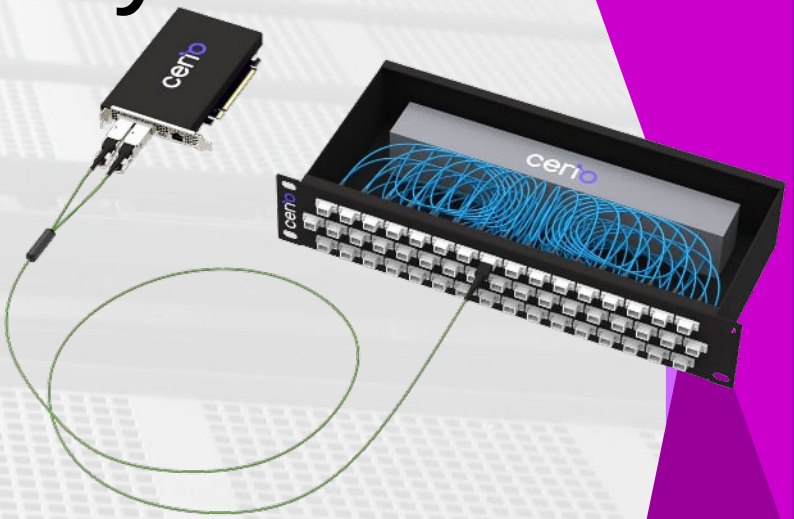
- Installed 2021
- A PCIe4-based composable fabric: 4 lanes to each server
- 3 A100 GPUs shared between 4 servers
  - Including a login node
  - GPUs per node can be changed in a few clicks
- 12TB RAM shared between these servers
  - Can be changed and reconfigured
- Positives: It works, can allow high RAM/GPU configs
- Negatives:
  - Bandwidth is shared
  - GPUs sometimes fail requiring a full stack reboot
  - RAM/kernel issues (and no Rocky9 support yet)
  - Rack-scale limitations
  - Bottlenecks
- ExCALIBUR funded



Alastair Basden
DiRAC / Durham University

CerIO composable system

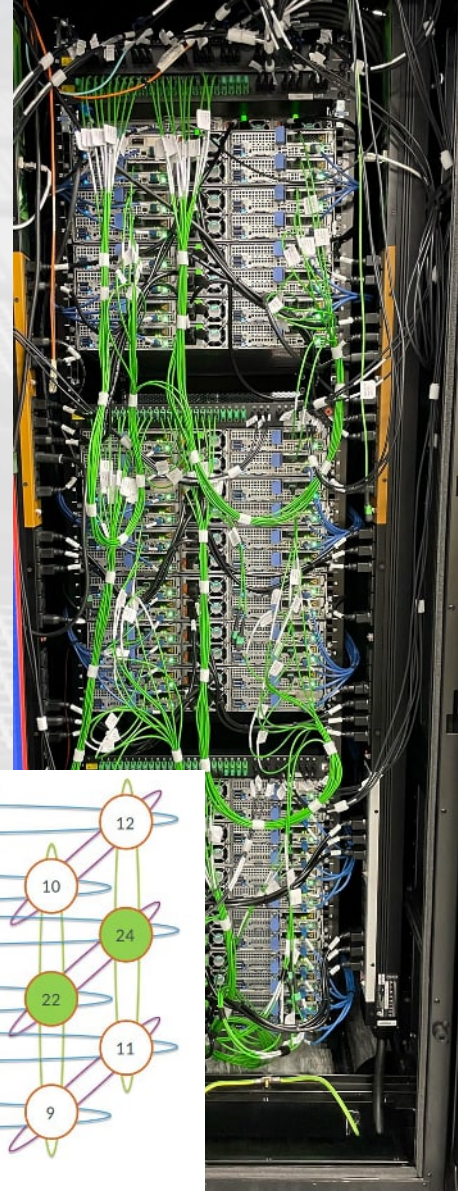
- Installed 2024
- PCIe5-based fabric
  - No central switch: Uses a flit-based torus topology
  - Full data centre scalability
  - 200Gb/s to nodes
  - 300Gb/s inter-card bandwidth
- 8 compute nodes, 8x A30 GPUs, 16TB RAM
- Will allow networking and composability within a single fabric
  - Exciting for future compute designs
- IRIS/SKA/ExCALIBUR/DiRAC

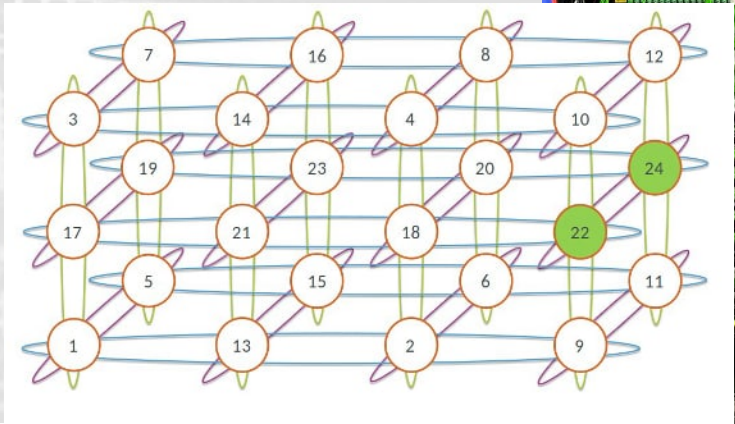


Alastair Basden
DiRAC / Durham University

Network fabrics

- Rockport 6D Torus network
  - A “switchless” fabric for 100G Ethernet
  - Trailed on DINE in 2021
  - Installed on COSMA7 in 2022
    - 224 nodes (half the cluster) replaced IB
    - Allows direct comparison of fabrics
    - At full HPC problem-size scale
  - Works well
    - Performance comparable to InfiniBand
    - For real workloads
    - Handles congestion well
  - ExCALIBUR/DiRAC funded
- 800G Ethernet fabric coming soon





Alastair Basden
DiRAC / Durham University

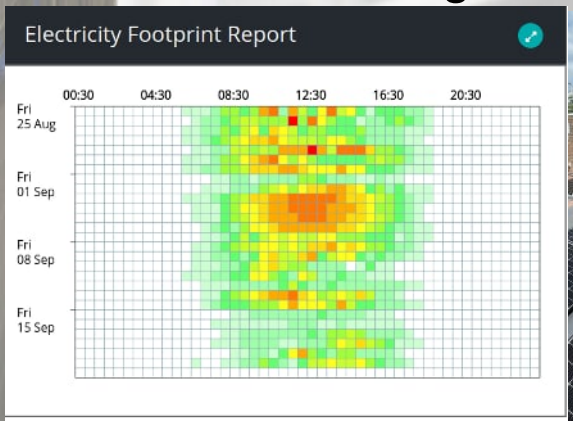
Storage sub-lab

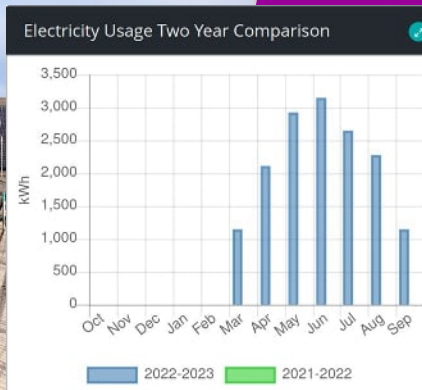
- Various different storage technologies
  - Most make it into production
- High-performance scratch Lustre (NVMe)
- DAOS (NVMe)
- Ceph
- StorJ private cloud
- VAST (NVMe)
- Globus (data transfer)
- Lustre (efficient bulk storage)
- Tape: Atempo Miria
- NVMe RAID test systems
- Funded by DiRAC/IRIS/SKA/ExCALIBUR

Alastair Basden
DiRAC / Durham University

Environmental-related

- HPC is a huge energy user
  - COSMA ~1MW at peak
    - (~5GWh/year, ~120T CO2, 70 people crossing the Atlantic and back)
  - Responsibility to keep this as low as possible
- 2023: Installation of ~£1m solar panels
  - Funded by DiRAC
  - Investigation into the interplay between supply and demand



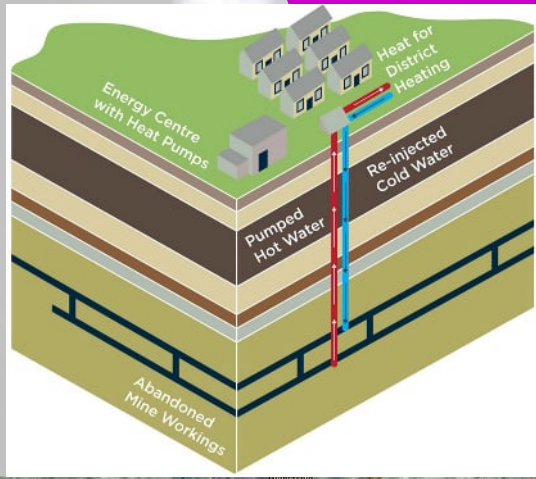


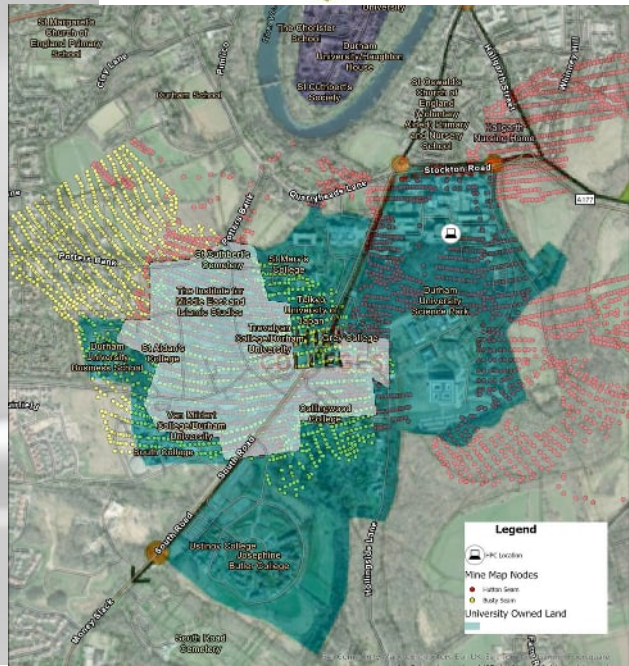
Cooling sub-lab

- A sequence of technologies:
  - Hot aisle (2006?)
  - Passive cooled rear doors (pre 2010)
  - Active cooled rear doors (2018)
  - Direct liquid cooling (2020)
  - Immersion cooling (2024-5)
    - As a national object-of-study
    - Support for visits to Durham for operators to learn this technology
    - Reduced operational and embodied CO2
    - EPSRC DRI funding
- Free air coolers: 12kW to cool 800kW load
  - Dependent on external temperature

Mine water heat storage

- HPC produces a lot of heat
  - We can heat buildings with this in the winter
  - What can we in the summer?
    - Store it underground for extraction in the winter
- Data centre is sitting on old, flooded mine workings
- Ongoing project to investigate feasibility
  - In particular, how fast does the water flow?
- EPSRC funded





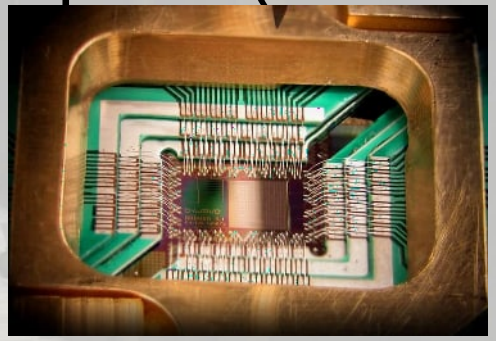
Alastair Basden
DiRAC / Durham University

Energy awareness

- Quarterly reporting to users around energy usage
  - Ability to query energy usage for each submitted job
- Monitoring of power on a node, rack, room and system scale
- Power-down of unused nodes
  - On quieter systems

Quantum computing

- No quantum compute in Durham (yet)
- However, the hardware lab administers national access to:
  - DWAVE quantum annealer
  - QuEra neutral atom quantum computer (imminent)
- ExCALIBUR funded



quantumcomputinguk.org

Alastair Basden
DiRAC / Durham University

Accessing the hardware lab

- Sign up on SAFE:
  - safe.epcc.ed.ac.uk/dirac
- Apply to join an appropriate project code: (durham.readthedocs.io)
  - do009: General purpose
  - do015: Cerio compsable system
  - do016: NVIDIA GPUs
  - do017: Intel GPUs
  - do018: AMD GPUs
- And feel free to arrange a visit to the data centre
- Availability: No plan to turn them off!

Alastair Basden
DiRAC / Durham University

Hardware Lab Outputs

- Key outputs from the hardware lab are:
  - Up to date knowledge of performance on new technologies
  - Experience profiling and optimising codes
  - Code preparation for future systems
  - Training on new technologies and tools
  - User awareness
  - Input into future system design

Alastair Basden
DiRAC / Durham University

Visiting

- Visits to the HPC lab:
  - During Durham HPC Days 2025
    - 2-6th June (week before ISC2025)
  - (and any other reasonable time)

Future plans*

- MI300A systems
- UntetherAI card
- Turin CPU system
- CXL composable systems
- Ultra-Ethernet fabric

* dependent on funding!

Alastair Basden
DiRAC / Durham University

Conclusion

- The Durham HPC Hardware Laboratory
  - Accessible for UK researchers
    - Single login
    - HPC environment
  - Cutting edge technologies
  - Let us know if there is something of particular interest
    - Always happy to host!
  - A steady funding line would be a good thing!
- Come and visit: Durham HPC Days 2-6th June

Alastair Basden
DiRAC / Durham University

CIUK 2024 Presentations – Jacky Pallas Memorial Award Winner

Lisa Lampunio (Imperial College London)

Advanced Modelling and Simulation for the Analysis of Novel Radiation Detection Technology and Thermal Fatigue Phenomena within the Nuclear Energy Sector

Abstract: A key aspect of this scientific computing research is to address challenging problems within the nuclear energy sector, such as the development of multiphysics modelling and simulations (M&S) of thermal fatigue phenomena within nuclear power plants (NPPs) and the design of novel radiation detection technologies. Thermal fatigue phenomena are widespread within many areas of engineering and can limit the operational life of components, cause power outages, and require costly periodic inspections. Therefore, high-fidelity predictive models enable engineers to improve the reliability of engineering components and optimise maintenance scheduling. The development of novel radiation detection technologies is critical not only in the nuclear power sector but also for nuclear security and safeguard. Both these fields require the application of advanced high-performance computing (HPC) M&S along with uncertainty quantifications (UQ), surrogate modelling (SM), and machine learning (ML) algorithms. This research work aims to develop computationally efficient and improved M&S methods as well as advanced, cost-effective, radiation detector designs with benefits to both the engineering community and the nuclear energy sector.

Bio: I hold a Bachelor and Master's degree in Nuclear Engineering from the University of Pisa in Italy and a PhD from the Nuclear Engineering Group of the Department of Mechanical Engineering at Imperial. My research work has been focused on developing high-fidelity multi-physics modelling and simulations and efficient data-driven machine learning surrogate models for critically important problems for nuclear power plant safety and control, such as thermal fatigue phenomena in T-junction pipes and design optimisation of neutron detectors within the reactor core.

CIUK 2024 Presentations

Alastair Basden (DiRAC / Durham University)

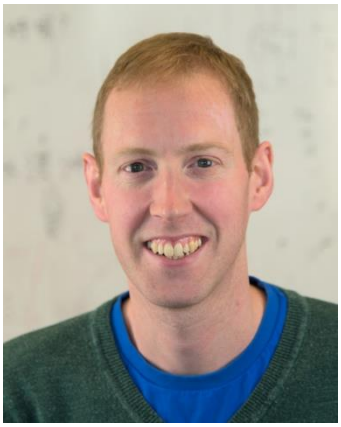
The Durham HPC Hardware Laboratory

Abstract: The Durham HPC Hardware Lab is hosted by the DiRAC COSMA HPC facility and provides UK researchers with access to cutting edge technologies and facilities, to allow testing and benchmarking of codes, software migration to newhardware, and study of new paradigms. This lab has grown in scope over the past few years, funded by ExCALIBUR H&ES, DiRAC, Durham and various UKRI grants. Of particular interest to many users is access to new GPU systems, novel networking topologies, composable infrastructure, and access to BlueField DPUs. This talk presents the Hardware Lab, including information about how it can be accessed.

Bio: Alastair manages the DiRAC Memory Intensive HPC service, COSMA, at Durham University.

CIUK 2024 Presentations

Nick Brown (EPCC at the University of Edinburgh)



RISC-V: The most exciting hardware innovation of recent times

Abstract: Powering over 20 billion devices since it was first released around a decade ago, the open, community driven, RISC-V Instruction Set Architecture (ISA) has enjoyed phenomenal growth especially in embedded computing. However, whilst RISC-V is yet to become widespread in HPC recent developments such as the high-core count SG2042 RISC-V CPU and several RISC-V accelerator cards have the potential to change that. From my perspective as leading the RISC-V HPC Special Interest Group (SIG) I will introduce RISC-V, explore the potential impact it could have on HPC and describe how people can get involved and be a part of potentially the most exciting hardware revolution of our lives.

Bio: Dr Nick Brown is a Senior Research Fellow at EPCC, the University of Edinburgh. His main interest is in the role that novel hardware can play in future supercomputers and is specifically motivated by the grand-challenge of how we can ensure scientific programmers are able to effectively exploit such technologies without extensive hardware/architecture expertise. Combining novel algorithmic techniques for new hardware, programming language & library design, and compilers, he has over 80 peer reviewed publications and has worked on a number of large scale parallel codes. He is chair of the RISC-V HPC SIG, leads EPCC's RISC-V testbed.

RISC-V: THE MOST EXCITING HARDWARE INNOVATION OF RECENT TIMES

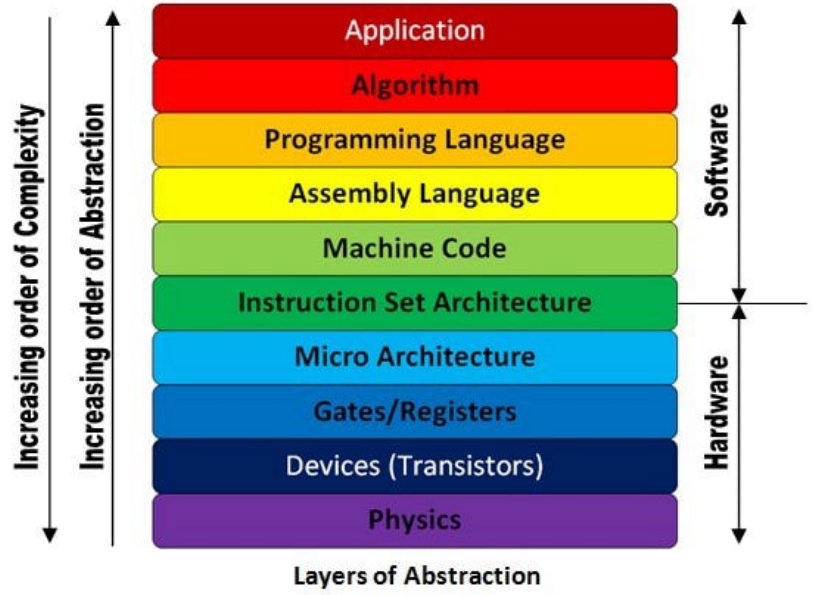
Nick Brown (EPCC)

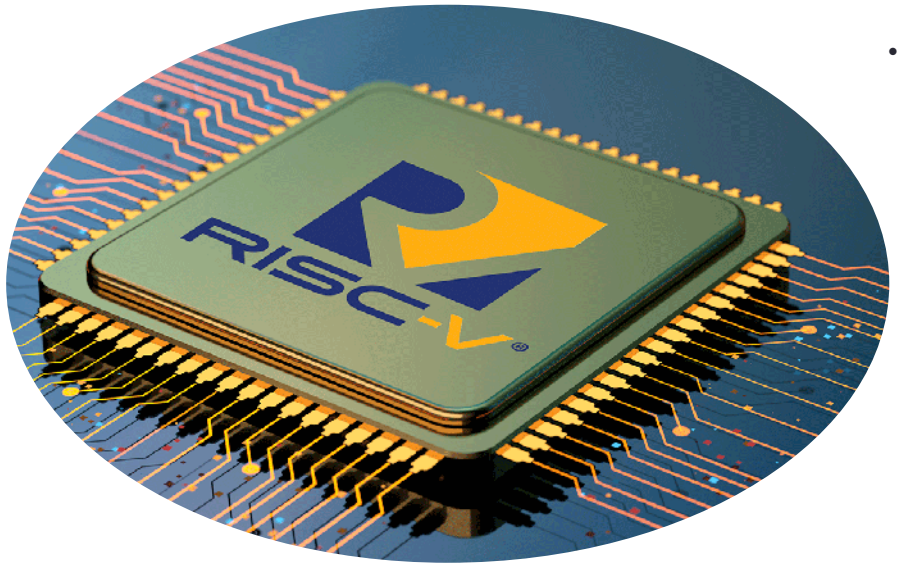
n.brown@epcc.ed.ac.uk

ExCALIBUR | epcc | THE UNIVERSITY OF EDINBURGH

What is RISC-V

- Started out by Berkley in 2012
- An open Instruction Set Architecture (ISA) which is overseen by RISC-V International
  - Standardisation activities driven by expert members
  - Numerous areas of focus ranging from HPC & ML to the data centre to embedded computing





- RISC-V is an Open Standard Instruction Set Architecture (ISA)
  - Software uses the ISA to tell the hardware what to do.
  - At the base level, the RISC-V ISA and extensions ratified by RISC-V International are royalty free and open base building blocks for anyone to build their own solutions and services on

RISC-V International is the global standards body

- **4k+ individuals in 60+ RISC-V work groups** and committees
- **330+ RISC-V solutions** online including cores, SoCs, software, tools, and developer boards
- **35 local RISC-V community** groups, with more than **6,800+ engineers**
- **Over 4200** members in **70+** countries

RISC-V®

Lots of serious players here

tenstorrent

Qualcomm

SiFive

RiVOS

InspireSemi™

ExCALIBUR

epcc

THE UNIVERSITY OF EDINBURGH

RISC-V HPC Special Interest Group (SIG)

- The group in RISC-V International that is concerned with driving the update of RISC-V in HPC
- Lots of members across a wide range of organisations and vendors
- Do a whole load of activities:
  - Organise workshops, panels etc at SC, ISC, HPC Asia
  - Marketing of RISC-V in HPC
  - Gap analysis to understand what is missing
  - Explore how other parts of the standard/activities can be leveraged for HPC

Why for HPC?

- Modularity and freedom to design bespoke hardware is the key advantage
- Especially as we see an increased focus on energy efficiency

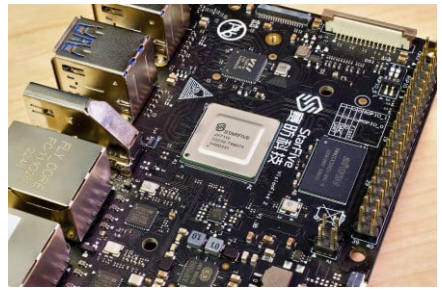
ISA	Chips?	Architecture License?	Commercial Core IP?	Add Own Instructions?	Open-Source Core IP?
x86	Yes, ***three*** vendors	***No***	***No***	***No***	***No***
ARM	Yes, ***many*** vendors	Yes, ***expensive***	Yes, ***one*** vendor	***No*** *(Mostly)*	***No***
RISC-V	Yes, ***many*** vendors	Yes, ***free***	Yes, ***many*** vendors	***Yes***	***Yes, many available***

- We have set up a RISC-V testbed for HPC, enabling free access for scientific developers to experiment with the technology for their workloads
- Looks/feels like any other HPC machine, with login node, shared filesystem, module environment and Slurm queue providing access to compute nodes

ExCALIBUR | epcc | The University of Edinburgh

Initially a scrabble to get hardware…..

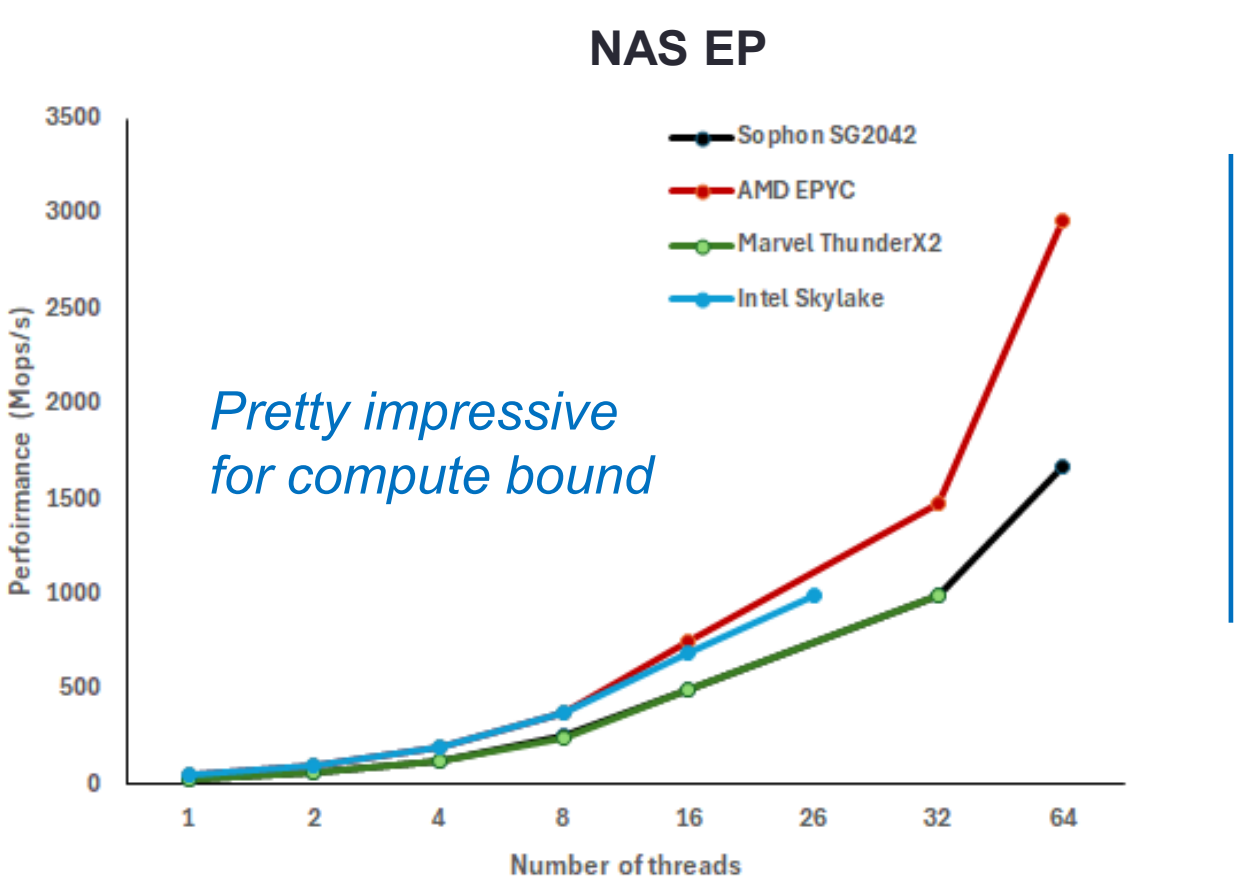
- When we initially set this up in early 2022, hardware availability was a big challenge
  - Had to rely on SBC designed more for embedded workloads

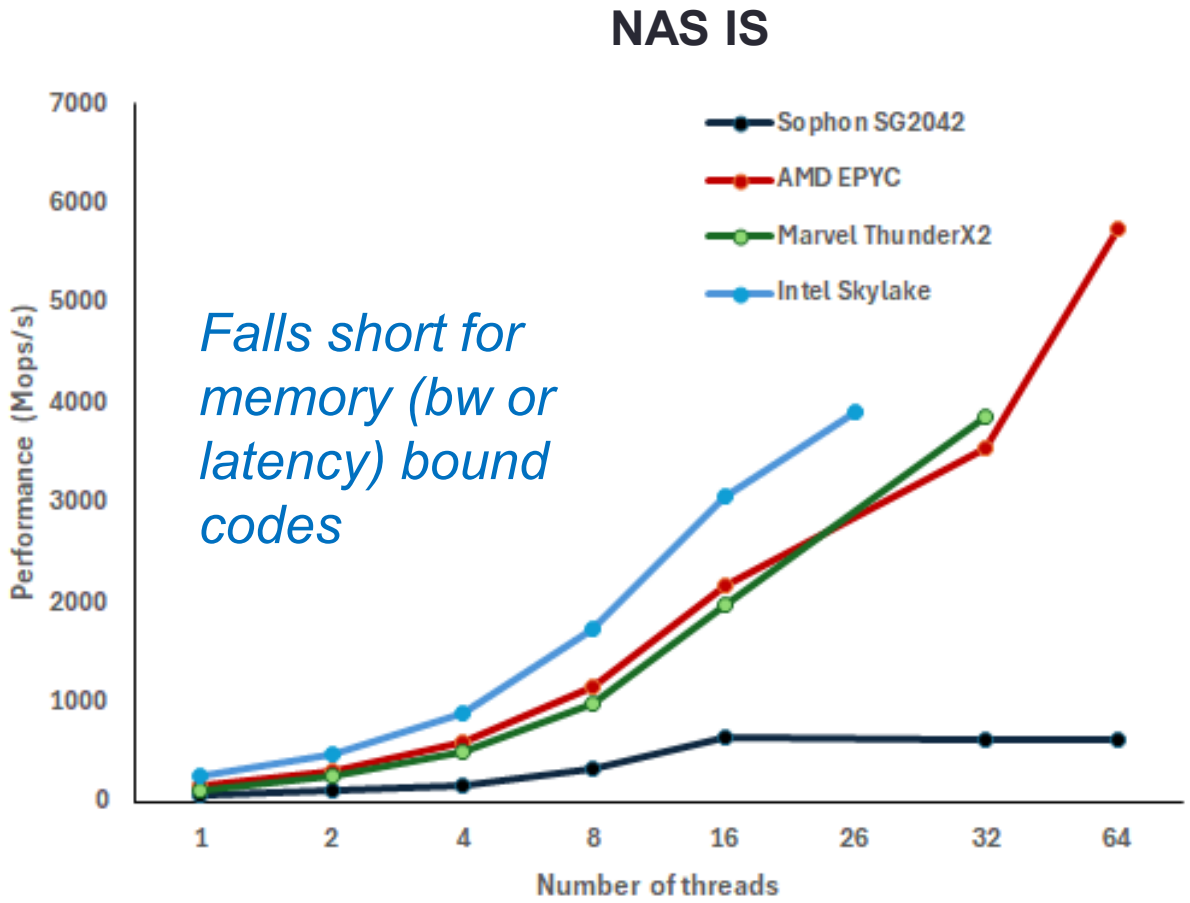


- In 2023 Sophon released their 64-core SG2042 RISC-V CPU designed for high performance workloads
  - A much more realistic proposition for HPC
  - Core for core, each for of the SG2042 is at-least three times as fast (and sometimes many times) what is in these SoCs

ExCALIBUR | epcc | The University of Edinburgh

How do RISC-V CPUs currently compare?

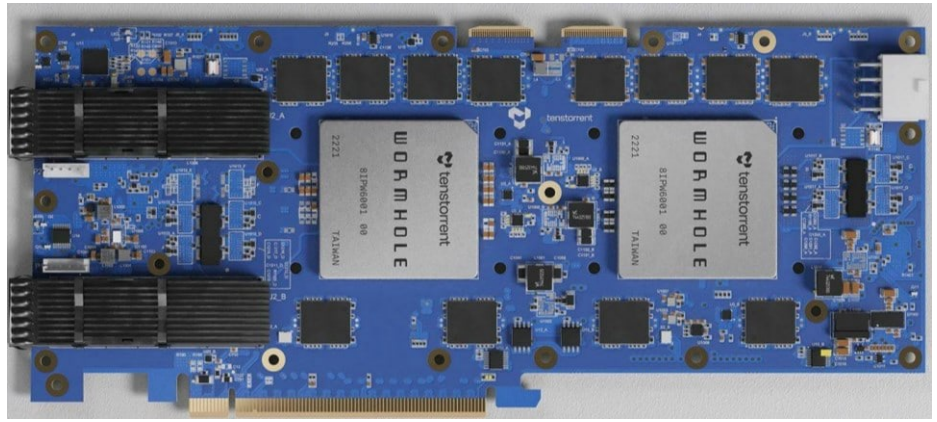




Lots more details at https://arxiv.org/pdf/2406.12394 and https://browse.arxiv.org/pdf/2309.00381.pdf

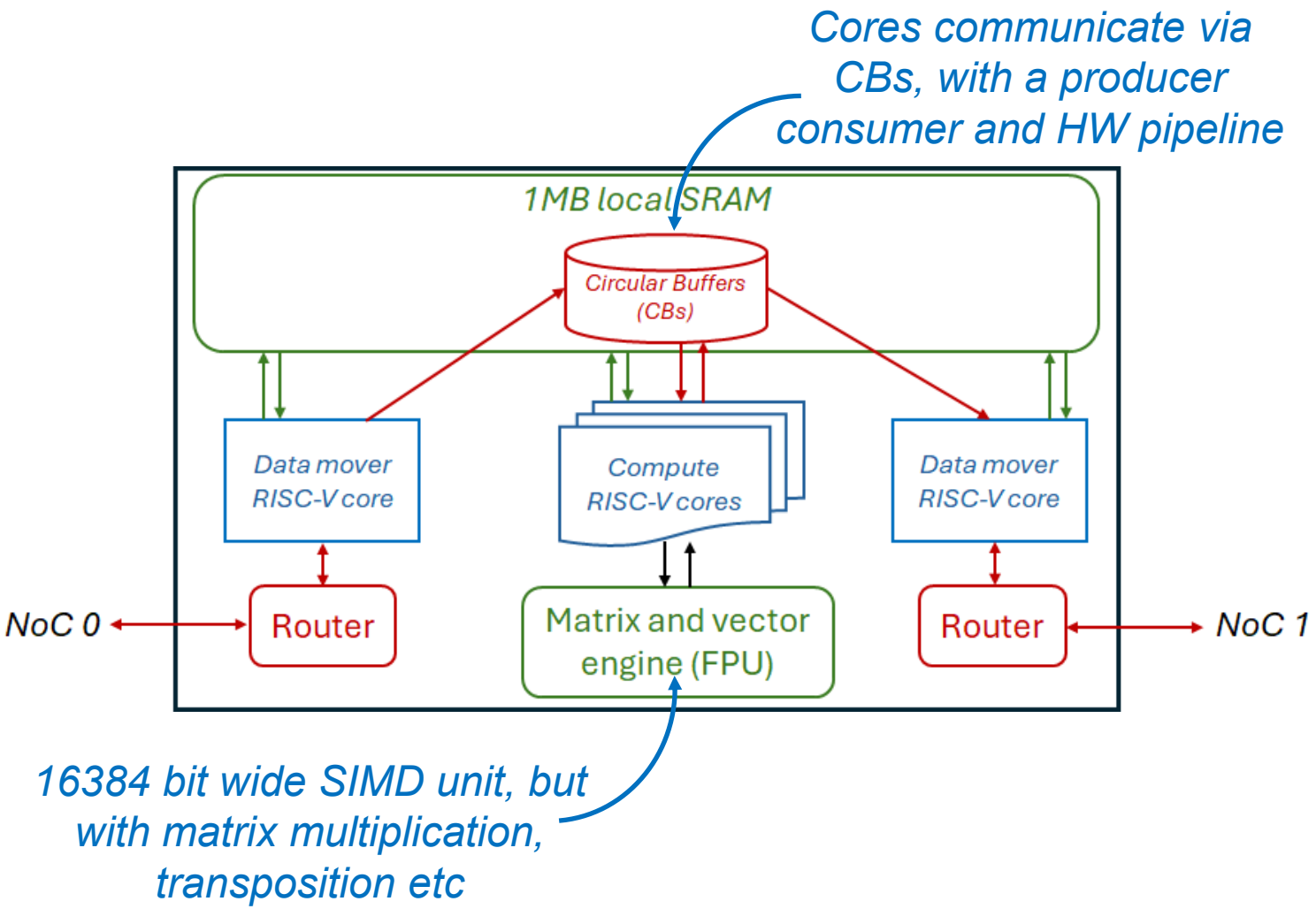
ExCALIBUR | epcc | The University of Edinburgh

RISC-V PCIe accelerator cards



ExCALIBUR 10 | epcc | The University of Edinburgh

Deep dive: Tenstorrent accelerator



- 120 of these Tensix cores in the Grayskull with 8GB of DDR4
- 128 of these Tensix cores in the Wormhole with 24GB of DDR6
- Goes back to the ability to specialise here
- RISC-V means a common compiler can be used, although need to use their specific API for programming

Experimenting with Tenstorrent Grayskull

- The architecture has been designed for AI inference, but TT develop an open source Metalium framework that enables direct programming

- We ported a stencil code to the accelerator
  - Stencils are ubiquitous in scientific computing, whilst this is a simple Jacobi solver for LaPlace's equation for diffusion it enabled us to explore the appropriate code techniques in detail
  - Initially, was way slower than a core of the CPU (Xeon Platiunum Cascade Lake)

Version	Performance (GPt/s)
CPU single core	1.41
Initial	0.0065

Initially a Tensix core was considerably slower than a CPU core!

Experimenting with Tenstorrent Grayskull

- We undertook a range of experimentation to understand the most appropriate approaches for data movement and where bottlenecks may lie
  - Considering how best to map the domain to the FPUs in the Tensix cores, and most effectively leverage the data movement capabilities

- Ultimately, slightly better performance but at five times less energy.

Based upon this optimisation, we were able to improve performance on a single Tensix core by around 160 times

Across the entire chip, we can slightly outperform the 24-core Xeon Platinum (Cascade Lake) but at five times less energy usage

Type	Total cores	Cores in Y	Cores in X	Performance (GPt/s)	Energy (Joules)
CPU	1	-	-	1.41	1657
CPU	24	-	-	21.61	588
e150	1	1	1	1.06	2094
e150	2	1	2	2.48	893
e150	4	1	4	2.92	744
e150	8	4	4	7.99	276
e150	32	8	4	9.20	240
e150	64	8	8	12.96	170
e150	72	8	9	17.26	128
e150	108	12	9	22.06	110

Much more detail at https://arxiv.org/pdf/2409.18835

ExCALIBUR | epcc | The University of Edinburgh

Conclusions & get involved!

- RISC-V is growing extremely rapidly
  - There is still plenty to do here, but I think lots of potential for scientific computing and HPC more generally. This is our opportunity to shape this technology to suit our needs.

- A wide range of activities are going on to push RISC-V more into HPC and mature the ecosystem

- Get involved!
  - https://riscv.epcc.ed.ac.uk
  - RISC-V International is free to join for individuals and academic institutions
    - HPC SIG meets virtually every month (and lots of other SIGs too!)

CIUK 2024 Presentations

Tina Friedrich (Snr HPC Systems Administrator, Advanced Research Computing, University of Oxford)

A first look at JADE 2.5

Abstract: Acting on behalf of the JADE consortium, the University of Oxford has recently purchased and provisioned the JADE 2.5 system - a technology pilot to investigate the suitability of AMD Instinct Accelerators for Artificial Intelligence and Machine Learning (AI/ML) research. The JADE

2.5 system is operated by the University of Oxford's IT Services Advanced Research Computing (ARC) team; in this talk, we give an overview of the system hardware and setup, as well as first experiences with the system.

Bio: Tina Friedrich is a Senior HPC Systems Administrator in the Advanced Research Computing (ARC) team at the University of Oxford, which provides access to High Performance Computing resources, support, and advice to researchers within the University of Oxford. Before joining Oxford, they worked at Diamond Light Source, managing HPC systems and contributing to various systems administration responsibilities. They hold a Master's in Physics from the University of Heidelberg.

A first look at JADE 2.5

Computing Insight UK 2024

Tina Friedrich / Oxford ARC team

Contents

What is JADE?

- ”Joint Academic Data Science Endeavour”
- EPSRC funded Tier 2 facility supporting advanced data science
- National GPU computing facility dedicated to advancing research in machine learning and artificial intelligence research in the UK.
- Operated by a consortium of UK universities and The Alan Turing Institute.

JADE 2

- Based on ”NVIDIA MAXQ Deep Learning System”.
- 63 servers, each with 8 Tesla V100 GPUs, connected by NV link.
- Supplied and integrated by ATOS Bull, run by the Hartree Centre.
- 1PB useable storage (plus some flash storage), EDR IB
- 600 active users; over 3 million GPU hours every year.
- Typical monthly utilisation of ~85%.
- Now end of live (as of November 2024).

JADE 2.5

- ”Technology Pilot”
- Based on 24x AMD MI300 GPUs linked by infinity fabric.
- Small system — three servers (8 GPUs each).
- Primary objective is to asses the AMD GPUs’ suitability for this type of research, rather than directly conducting research.
- Time will be allocated — target is a month of time per JADE partner.
- 20% of the time will be allocated to EPSRC to be made available through the RAP process.
- Plan to also schedule ”student weeks”.
- Project allocations carry a requirement for feedback on system performance and behaviour.

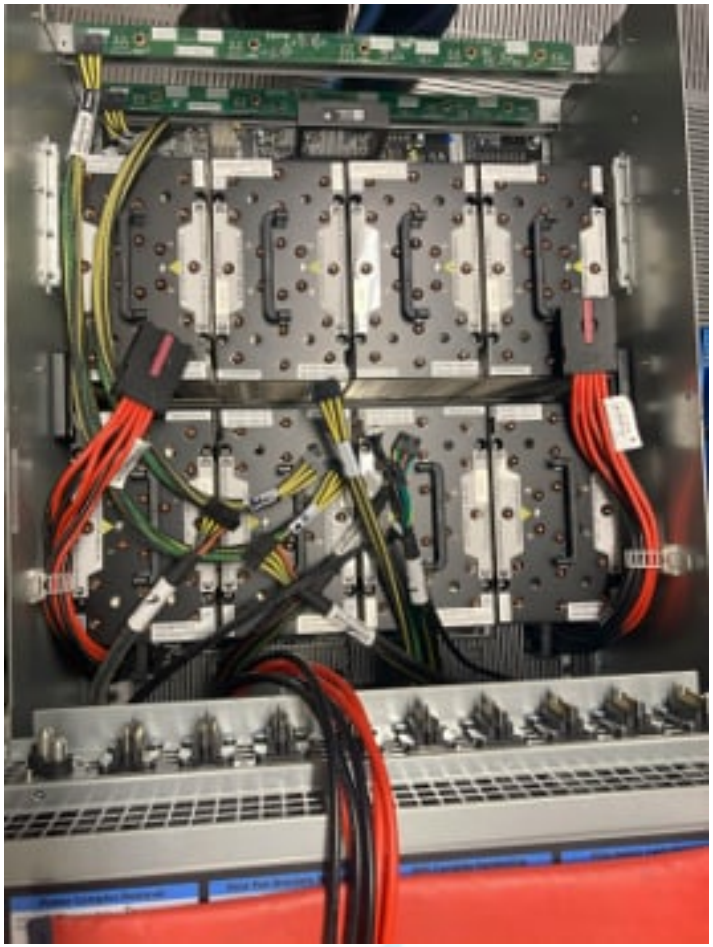
- Aim was to start onboarding users in November.
- First beta users created last week.
- Slippage mainly due to
  - Delayed shippping
    - System was supposed to be installed by 31st July.
    - System got delivered on 11. September,
  - Shortly after installation, Lenovo identified that the as built power distribution board could not adequately supply power to the GPUs with the latest AMD firmware.
    - Performance reduction of 5-10%
    - Required replacement of power distribution board.

Hardware

- 3 Lenovo ThinkSystem SR685a with:
  - AMD MI300X GPUs (192GB GPU memory)
  - 8 GPU Board, linked by Infinity fabric
  - 2x AMD EPYC 9534 (128 cores per server)
  - 2.25 TB RAM
  - 32TB node-local 'scratch' storage
  - 960GB M.2 (RAID 1) for OS
  - Networking:
    - NVIDIA ConnectX-7 NDR200 IB
    - Mellanox ConnectX-6 Lx 25GbE
  - 8x 2600W power supplies (2N redundancy)

- 8U server; weighting ~ 110kg.
- Requires 1200mm racks.
- Estimated power draw per server:
  - 6715W at 85% workload
  - 7946W at 100%

Notable installation — first systems of this type installed in the UK.



Setup

- Systems built using ARCs existing stack (PXE, kickstart, CFEngine).
- OS is AlmaLinux 9.4.
- ROCm installed via RPM (using AMD's repositories).
  - No significant issues during installation.
- Integrated into existing ARC infrastructure.
  - Dedicated SLURM controller and SLURM database.

- Shares ARC networking and storage infrastructure:
  - HDR fabric (interconnect, fast storage).
  - WEKA file system for non-local scratch.
  - Lenovo DM3010H for ”bulk” data storage.
- Software environment:
  - Managed with EasyBuild.
  - Minimal set of pre-installed toolchains.
  - Users are expected to bring (or build) their own software.

Theoretical performance

Performance figures as published by AMD and NVIDIA[1] [2].

benchmark	H100 SXM	MI300X
TF32 (TFLOPs)	989	1,307.4
FP16 (TFLOPs)	1,979	2,614.9
BFLOAT16 (TFLOPs)	1,979	2,614.9
INT8 (TOPS)	3,958	5,229.8
FP8 (TFLOPS)	3,958	5,229.8
FP64 vector	34	81.7
FP32 vector	67	163.4
FP64 matrix	67	163.4

- GPU memory
  - H100 SXM: 80GB (bandwidth 3.35 TB/s)
  - MI300X: 192GB (bandwidth 5.3 TB/s)
- power
  - H100 SXM: 700W
  - MI300X: 750W

[1] https://www.amd.com/en/products/accelerators/instinct/mi300/mi300x.html

[2] https://www.nvidia.com/en-us/data-center/h100/

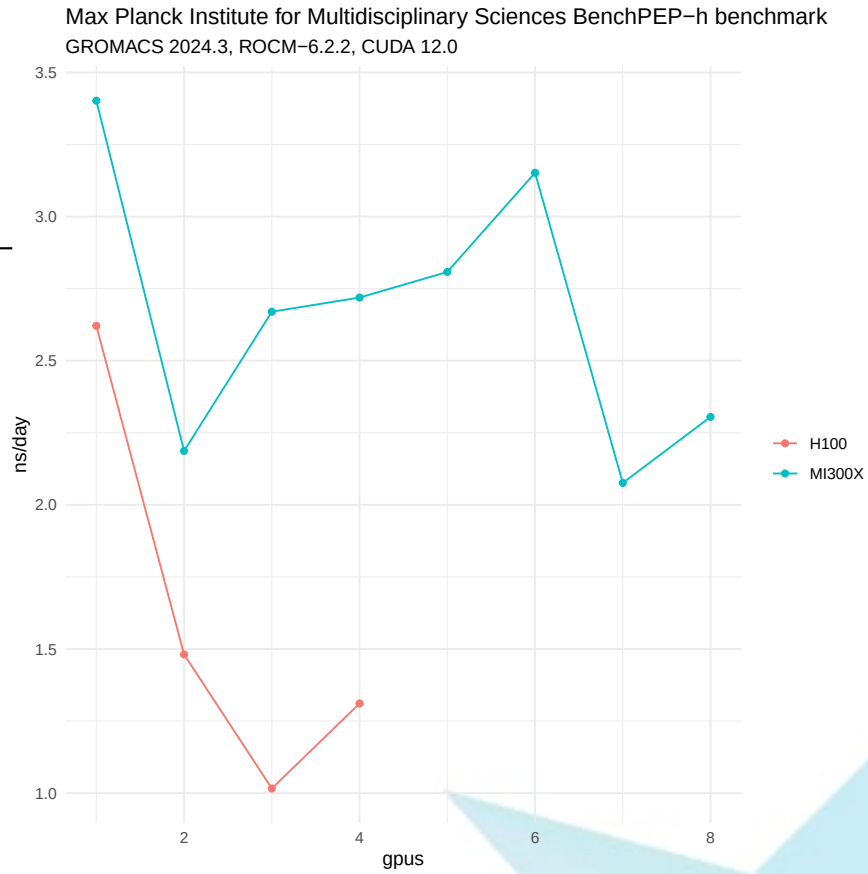
Functional tests

- TensorFlow container (AMD Infinity Hub):
  - Successfully run on system — no issues observed.
- ROCm installation/ROCm Validation Suite testing:
  - Installed latest ROCm; initial testing promising.
  - Running MI300X test configurations caused severe system crashes.
  - Setting a power cap helped.
  - Querying with AMD, we were told that these systems are not currently supported by latest ROCm and suggested this as root cause of the crashes.
  - Running a ”supported” firmware/ROCm combination, all is stable.

GROMACS benchmarks

GPUs	MI300X capped [ns/day]	MI300X uncapped [ns/day]	H100
1	1.791	3.402	2.621
2	1.472	2.187	1.481
3	2.325	2.67	1.016
4	2.826	2.719	1.311
5	2.224	2.808	
6	2.236	3.152	
7	2.498	2.076	
8	2.351	2.305	

Table: GROMACS benchPEP-h results



Recap

- ”Technology pilot” based on AMD MI300 GPUs.
- Access will be allocation based.
- Beta users on the system now.
- Successful set up — no major issues encountered during provisioning.
- Significantly delayed delivery.
- Minor issues encountered throughout setup, likely due to the newness of the system.
- The platform requires a very specific combination of firmware and driver versions (not made clear in Lenovo/AMD documentation). We do not yet understand the impact of this on user workloads.

Next steps

- More comprehensive testing and benchmarking.
- Collaborate with AMD and Lenovo to check, tune, and adjust BIOS settings for optimal performance.
- Assist first BETA users getting up and running on the system and get initial user feedback.

Visualising HPC Node Performance via Swift ReFrame Benchmark and Grafana

Gokmen Kilic

Durham University

High-Performance Computing systems are fundamental enablers of complex computational workloads across scientific, engineering, and industrial domains. Accurately measuring and visualizing the performance of HPC nodes is analytical for optimizing resource allocation, identifying performance bottlenecks, and improving overall system efficiency. This work presents an approach to monitor and visualize HPC node performance using the Swift ReFrame Benchmark framework in conjunction with the Grafana data visualization platform.The ReFrame Benchmarking Framework is a highly flexible and portable tool designed for writing regression tests for HPC systems. It focuses on scalability, ease of use, and seamless integration with various HPC environments. This work demonstrates ReFrame's capabilities to design customizable performance benchmarks that gather key metrics such as CPU utilization, memory bandwidth, and network latency across multiple HPC nodes. The benchmark results produced by ReFrame.

Visualising HPC Node Performance via Swift ReFrame Benchmark and Grafana

Gokmen Kilic *(Research Software Engineer, Durham University, ICC),* · **gokmen.kilic@durham.ac.uk** - **Alastair Basden** *(HPC Manager, Durham University, ICC)* · **a.g.basden@durham.ac.uk**

Durham University
Institute for Computational Cosmology

Introduction

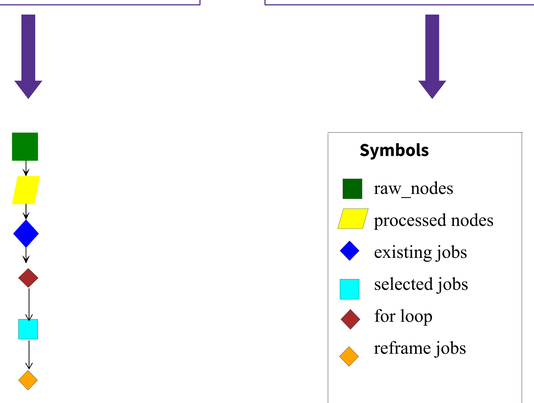
- Performance monitoring in High-Performance Computing optimizes resource management and system efficiency across scientific and industrial applications.
- The research integrates Swift ReFrame, Prometheus, and Grafana to create a flexible, economical approach to comprehensive HPC performance tracking and visualization.
- ReFrame's flexible framework enables precise performance assessment, dynamically measuring critical metrics like computational utilization, memory performance, and interconnect latency across diverse HPC infrastructures.
- Grafana's interactive dashboards provide real-time, comprehensive performance analytics, empowering sysadmins with actionable intelligence for strategic HPC system optimization.

Solution Overview

- **Node Management:** A bash script was developed to automate the selection and utilization of available HPC nodes for benchmarking.
- Extracts the list of available nodes while filtering out drained or non-usable states.
- Processes node ranges to create a usable format.
- Randomly selects nodes for ReFrame benchmarking tasks, ensuring no conflicts with ongoing jobs.

```
raw_nodes=

$(sinfo -h -o "%N %T"
-p cosma8 | awk '!/drain|alloc|mixed|down/
{print $1}')

processed_nodes=""
```

```
for i in {1..20}; do
    shuf_nodes=($(shuf -e "${NODE_ARRAY[@]}"))
    selected_nodes="m["
    for j in {0..1}; do
        selected_nodes+="${shuf_nodes[j]}"
        if [ $j -ne 1 ]; then
            selected_nodes+=","
        fi
    done
```



▲ The script integrated with the Swift ReFrame Benchmarking Framework, enabling on-demand execution of benchmarks with dynamic node allocation.

Visualising Node Performance via Grafana

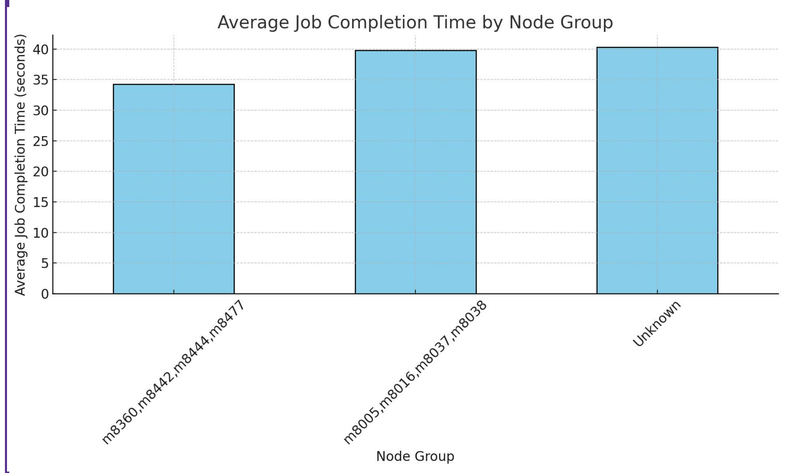
- **Real-Time Metrics Tracking:** Prometheus captures performance data, enabling dynamic dashboard updates for nodes, run time compilers etc.
- **Customizable Dashboards**: Node and cluster-level performance views, comparative node analysis.
- **Anomaly Detection:** Visualizes performance irregularities and potential system bottlenecks.
- **Integration Workflow:**
- ReFrame benchmarks generate the metrics.
- Prometheus stores performance data.
- Custom alerts monitor predefined for sysadmin.

Results

◀ The benchmark results collected from the SwiftBenchmark tests reveal variations in job completion times across different dates and configurations and over diffrent nodes. The results below from ReFrame perflogs that enable to handle node informations.

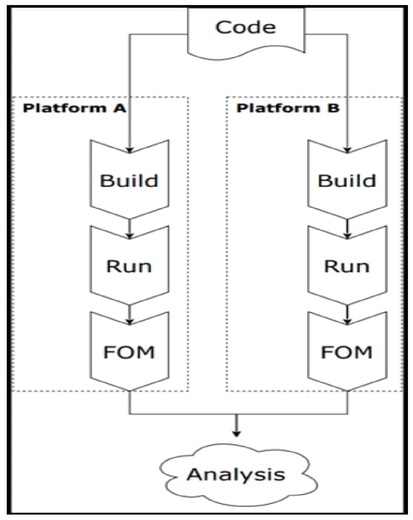
```
|4|8|1|null|40.0|seconds|50|None|0.2|swiftsim@0.9.0|SwiftBenchmark|cosma8|
|4|8|1|null|41.9|seconds|50|None|0.2|swiftsim@0.9.0|SwiftBenchmark|cosma8|
|4|8|1|null|40.4|seconds|50|None|0.2|swiftsim@0.9.0|SwiftBenchmark|cosma8|
m8360,m8442,m8444,m8477|4|8|1|null|34.0|seconds|50|None|0.2|swiftsim@0.9.0
m8360,m8442,m8444,m8477|4|8|1|null|34.5|seconds|50|None|0.2|swiftsim@0.9.0
```

◀ Due to the time and resources limitations, we only able to produce below bar chart to see node performances via Swift ReFrame benchmark.



- We can see from charth there is no anomaly between nodes and some of them a bit faster than others.
- The integration of Prometheus and Grafana enables real-time tracking and visualization of HPC node performance, providing actionable insights through interactive dashboards.

DiRAC COSMA



▲ Excalibur portable benchmarking framework introduces an innovative approach to evaluating parallel application performance portability, presenting a comprehensive software framework that integrates and enhances ReFrame and Spack environments for UK HPC systems.[1].

Github

https://github.com/ukri-excalibur/excalibur-tests

▶ A typical workflow for benchmarking of an application on different systems, adapted from a diagram from Pennycook et al. [2].

Limitations & Future Work

Limitations:

- Busy job queues with scaling large HPC clusters.
- Limited metric coverage.
- Overhead on HPC resources.
- Anomaly detection accuracy

Focus of future work:

- Optimizing the data pipeline to handle larger HPC clusters.
- Handling more metrics such as energy consumption.
- Introduce algorithms for anomaly detection and resolution.

Acknowledgements

This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

References

[1] Koskela, T., Christidi, I., Giordano, M., Dubrovska, E., Quinn, J., Maynard, C., Case, D., Olgu, K., & Deakin, T. (2023). Principles for automated and reproducible benchmarking. SC-W '23: Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis, 609–618. https://doi.org/10.1145/3624062.3624133

[2] S. John Pennycook, et al. 2023. Performance, Portability and Productivity Analysis Library.

CIUK 2024 Poster Competition Winner
Using PINNS to predict temperature due to neutronic heating fusion power plants

Zeyuan Miao

University of Manchester

Fusion energy offers a secure and sustainable solution to meet future energy needs. The design of fusion reactor components is highly complex, necessitating efficient exploration of extensive design spaces. Physics-Informed Neural Networks (PINNs) provide a powerful approach by storing knowledge to solve specific partial differential equations, enabling rapid and informed design analysis. This study presents a PINN surrogate model developed with Nvidia Modulus for heat transfer analysis in a vacuum vessel wall. Validated against finite element simulations and tested on various GPUs, the model demonstrates its potential for accurate and efficient thermal modeling in fusion reactor applications.

CIUK 2024

USING PINNS TO PREDICT TEMPERATURE DUE TO NEUTRONIC HEATING IN FUSION POWER PLANTS

Zeyuan Miao, William Smith, Raska Soemantoro, Muhammad Omer, Adam Barker, Anastasia Vasileiou, Hujun Yin, Lee Margetts

The University of Manchester

MANCHESTER 1824
The University of Manchester

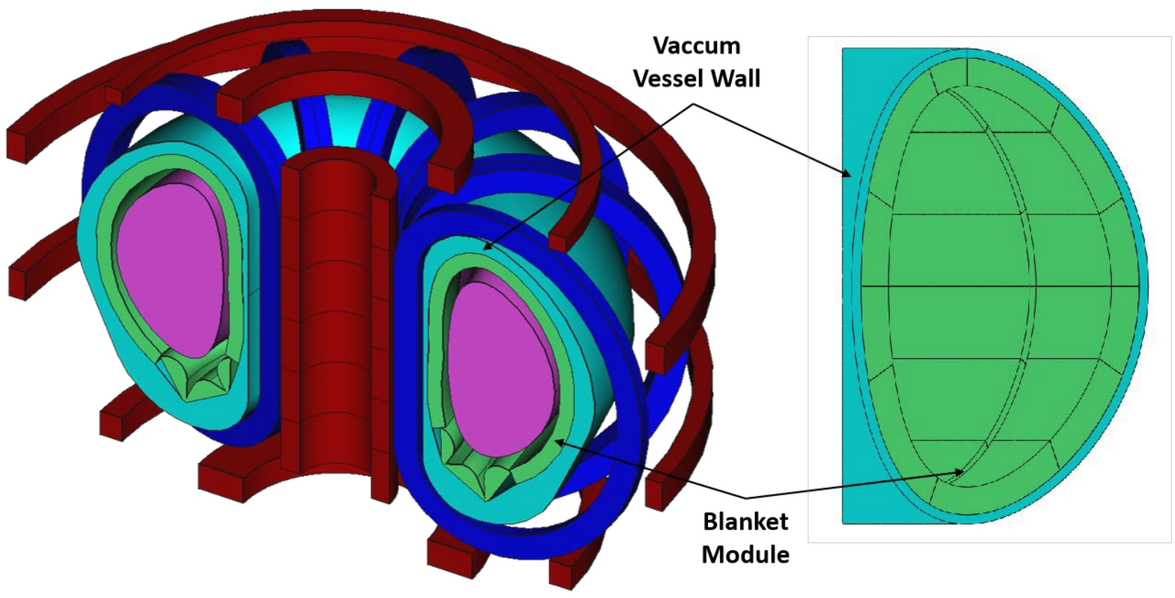
Introduction

Motivation

- Fusion energy, recognized for its potential to offer an almost limitless and secure power source, while mitigating environmental consequences and ensuring economic feasibility, stands as a promising solution for addressing our future energy requirements.
- Designing components for fusion reactors poses a complex challenge due to the need to consider a large design space and tight interdependence between different systems.
- Surrogate models and High-Performance Computing (HPC) are needed for quickly searching design space.

Objectives

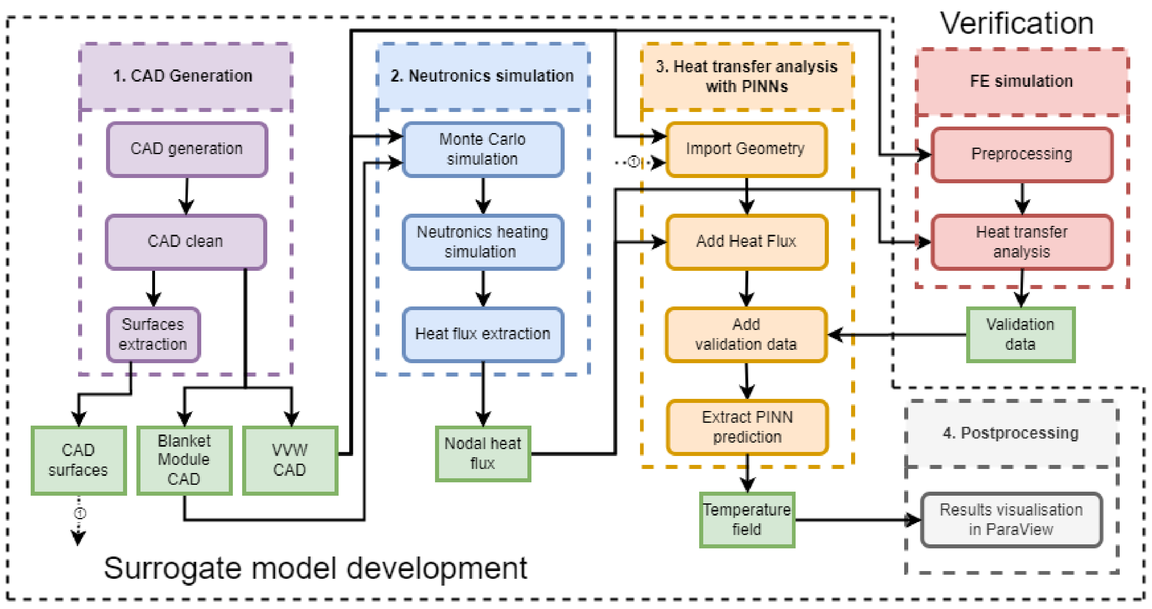
- Develop a Physics-Informed Neural Network (PINN) surrogate model for heat transfer analysis of vacuum vessel wall using Nvidia Modulus[1].
- Validate the PINN model by comparing the results with standard finite element simulations.
- Test PINN training on different types of GPU and utilize HPC resources to accelerate the training of PINN surrogate models.



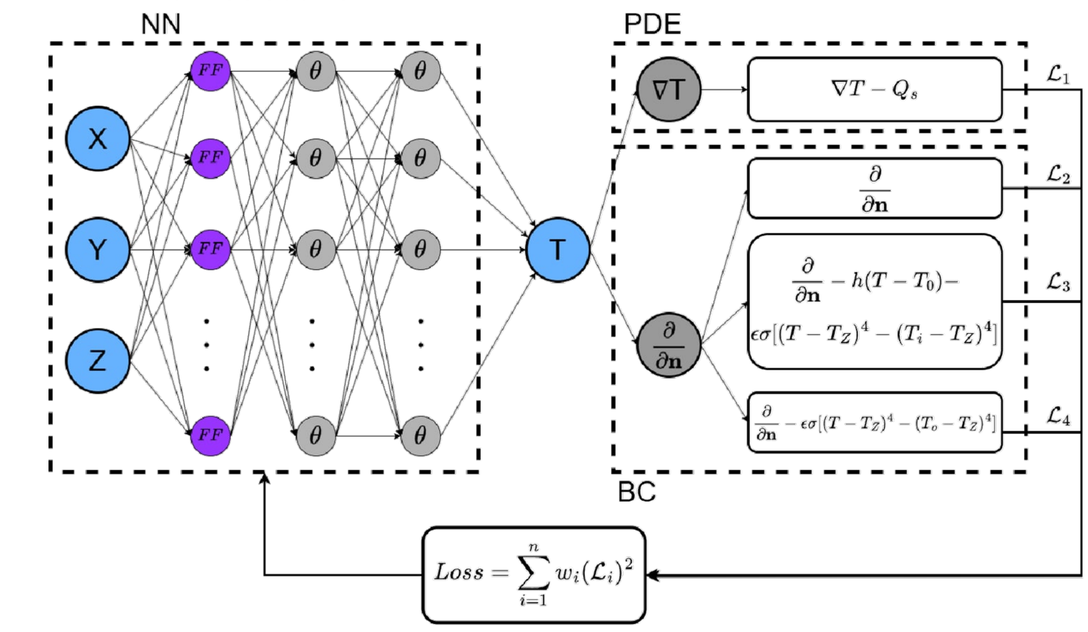
Simplified Blanket Module and Vacuum vessel wall used in a Tokamak[2]

Method and Key Results

Workflow



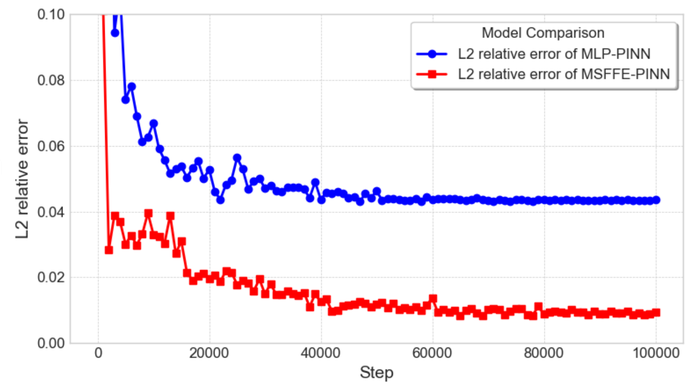
PINN Architecture



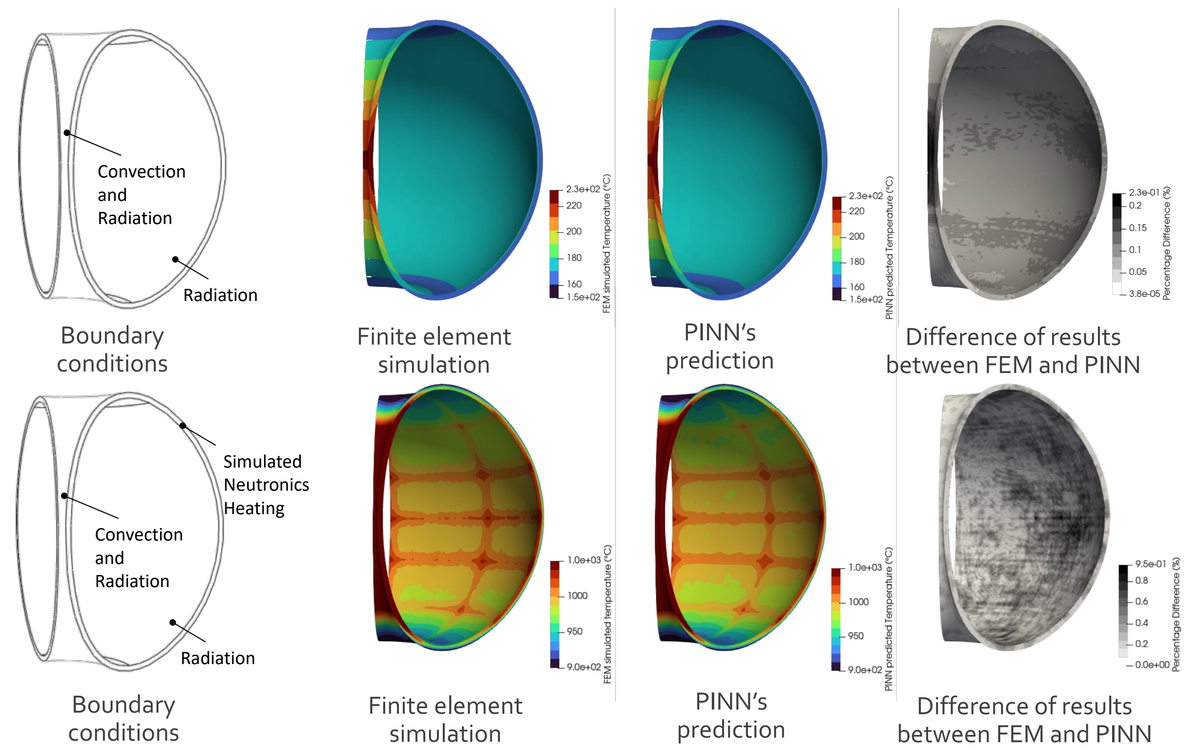
MLP-PINN (without purple layer): Standard PINN without Fourier Feature Extraction.
MSFFE-PINN (with purple layer): PINN with Multi-Scale Fourier Feature Extraction.

Learning Curves

- MSFFE-PINNs can do better prediciton on multi-scale problem
- The L2 relative error is 0.0093
- The training time of MSFFE is longer



FEM VS. PINNs



PINNs training on different type of GPUs

GPU Type	1 x L40 (Server)	1 x RTX 6000 (Workstation)	1 x GH200 (IsambardAI)
Training Time (s/per 100 steps)	19.4	18.1	9.8

Summary and Outlook

PINN surrogate models have been successfully developed using Nvidia Modulus to predict temperature fields due to neutronic heating in the vacuum vessel wall of a nuclear fusion power plant. The next step is to increase the complexity of the geometry of the problem and explore the benefits of using multiple GPUs on IsambardAI.

- Develop robust transfer learning strategies to adapt trained PINN models to new but similar cases, improving training efficiency and minimizing the need for extensive retraining, while enhancing knowledge retention across diverse simulations.
- While making a more general PINN will increase training time, deeper integration with HPC resources can be explored to efficiently manage larger-scale simulations.

Acknowledgments

This project has been supported by UK Atomic Energy Authority through the Fusion Industry Programme. The Fusion Industry Programme is stimulating the growth of the UK fusion ecosystem and preparing it for future global fusion powerplant market. More information about the Fusion Industry Programme can be found online: https://ccfe.ukaea.uk/programmes/fusion-industry-programme/. The authors acknowledge the use of resources provided by the Isambard-AI National AI Research Resource (AIRR). Isambard-AI is operated by the University of Bristol and is funded by the UK Government's Department for Science, Innovation and Technology (DSIT) via UK Research and Innovation; and the Science and Technology Facilities Council [ST/AIRR/I-A-I/1023].

Contacts

Email: Zeyuan.miao@postgrad.manchester.ac.uk

References

[1] Hennigh, O., Narasimhan, S., Nabian, M.A., Subramaniam, A., Tangsali, K., Fang, Z., Rietmann, M., Byeon, W. and Choudhry, S., 2021, June. NVIDIA SimNet™: An AI-accelerated multi-physics simulation framework. In *International conference on computational science* (pp. 447-461)

[2] Shimwell, J., Billingsley, J., Delaporte-Mathurin, R., Morbey, D., Bluteau, M., Shriwise, P. and Davis, A., 2021. The Paramak: automated parametric geometry construction for fusion reactor designs. *F1000Research*, *10*.

Accelerating High-Order Finite Difference Methods for Turbulence and Combustion Using CUDA Fortran

Terence Lobo

STFC- UKRI

High-performance computing (HPC) has become a cornerstone in advancing scientific and engineering research, particularly in fields requiring intensive numerical computations. This project focuses on accelerating an existing CPU-only high-order finite difference (FDM) solver for turbulence and combustion simulations by leveraging the computational power of Fortran and GPU programming.

Fortran is widely used in the scientific community due to its efficiency in handling numerical tasks. However, the increasing complexity and scale of simulations necessitate further optimisation to meet the demands of modern research. This project explores the integration of NVIDIA's HPC SDK (CUDA) with Fortran to enhance computational performance. By utilising CUDA Fortran, this project aims to parallelise all the computational tasks, thereby significantly accelerating numerical computations.

The primary objective is to demonstrate the feasibility and effectiveness of CUDA Fortran in parallelising tasks to achieve significant acceleration in numerical computations. Key results include a detailed comparison of performance metrics between traditional CPU-based computations and the optimised GPU-accelerated approach, using a single A100 GPU for comparison. The findings highlight the potential of CUDA Fortran to transform HPC applications, achieving up to 20x acceleration, which includes IO operations, making it a viable solution for large-scale simulations in turbulence and combustion research.

This poster will present the methodology, implementation details, and performance outcomes, providing insights into the practical benefits and challenges of using CUDA to adopt GPU programming for a high-order compressible flow solver. The results underscore the importance of leveraging modern HPC tools to push the boundaries of computational research.

Science and Technology Facilities Council

Scientific Computing

scd.stfc.ac.uk

Accelerating High-Order Finite Difference Methods for Turbulence and Combustion Research Using CUDA Fortran

Terence Lobo [1] Jian Fang [2]

[1]terence.lobo@stfc.ac.uk [2]jian.fang@stfc.ac.uk

Scientific Computing Department, Daresbury Laboratory, UK

Abstract

High-performance computing (HPC) is crucial for scientific and engineering research, especially for intensive numerical computations. This project aims to accelerate a serial implementation of the ASTR code, a high-order finite difference (FDM) solver for turbulence and combustion simulations, using CUDA Fortran. The ASTR code leverages Fortran and MPI for efficiency and scalability. However, this project focuses on a mini-app version, specifically a Taylor-Green Vortex (TGV) solver for a three-dimensional case, implemented in serial without MPI. By integrating NVIDIA's HPC SDK with Fortran, the project parallelises computational tasks, achieving up to 20x speedup using a single NVIDIA A100 GPU. This demonstrates CUDA Fortran's potential to transform HPC applications for large-scale turbulence and combustion simulations.

Test Case Description

The Taylor-Green Vortex flow is a test case used to study the transition to turbulence and the subsequent decay in a triply periodic cubic domain, which spans from 0 to $2\pi L$ in each direction [1]. By simulating this flow using the compressible Navier-Stokes equations, we can validate the accuracy and performance of the numerical methods employed. This helps us determine if the TGV solver is effective in predicting the behavior of turbulent flows. The Taylor-Green Vortex problem is defined with the following initial conditions:

$$
\begin{aligned}
u &= V_0 \sin(\frac{x}{L}) \cos(\frac{y}{L}) \cos(\frac{z}{L}) \\
v &= -V_0 \cos(\frac{x}{L}) \sin(\frac{y}{L}) \cos(\frac{z}{L}) \\
w &= 0 \\
p &= p_0 + \frac{\rho_0 V_0^2}{16} [\cos(\frac{2x}{L}) + \cos(\frac{2y}{L})][\cos(\frac{2z}{L}) + 2] \\
Re &= \frac{\rho U_0 L}{\mu} \\
M &= \frac{V_0}{\sqrt{\gamma R T_0}}
\end{aligned}
$$

As time progresses, the initial vorticity distribution undergoes vortex-stretching, generating smaller scales and eventually leading to turbulence. Since there is no external forcing to sustain the turbulent motion, a decay is observed after the transition. The incompressible problem is governed entirely by the Reynolds number Re, which is set to 1600. However, since the code solves the compressible flow equations, other dimensional parameters are adjusted to achieve nearly incompressible conditions, with a Mach number $M = 0.1$.

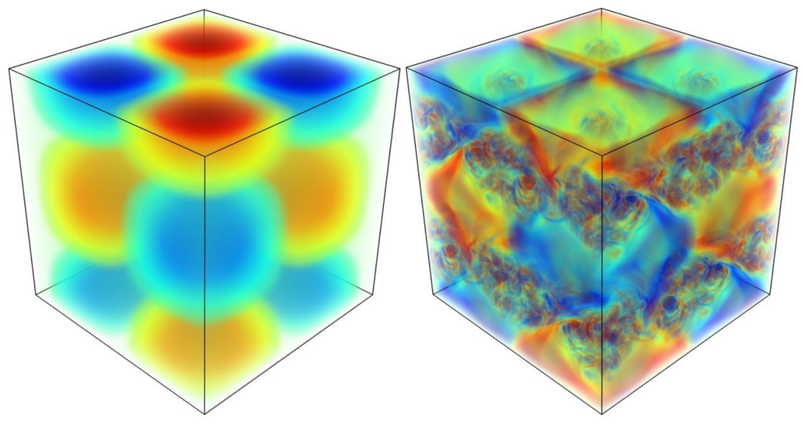


Figure 1: Taylor-Green Vortex: (Left) Flow at $t = 0$ showing inertial vortices. (Right) Later time showing transition to turbulence and small-scale structures. Figure taken from [2].

Objective

- This project explores NVIDIA's CUDA Fortran to optimise software and leverage GPU accelerators, aiming to identify performance bottlenecks in Fortran code and enhance performance through GPU programming techniques.
- The goal is to demonstrate CUDA Fortran's effectiveness in optimising Fortran applications and share insights to help developers accelerate their applications. Table 1 provides a detailed comparison of the key hardware components being utilised.

Characteristics	Intel Core i7-1365U	NVIDIA A100-SXM4-40GB
Type of Device	CPU (Central Processing Unit)	GPU (Graphics Processing Unit)
Clock Speed (Base vs Turbo)	1.8 GHz (P-cores), 1.3 GHz (E-cores) 5.2 GHz (P-cores), 3.9 GHz (E-cores)	1.095 GHz 1.410 GHz
RAM/Memory	32 GB LPDDR5-6400	40 GB HBM2e with 1.6 TB/s bandwidth
Cores/Threads	10 cores (2 P-cores, 8 E-cores) / 12 threads	6912 CUDA cores, 432 Tensor cores
Power Consumption	15W (base), 55W (max turbo)	400W
Compiler	GNU Compiler 11.4.0	NVFORTRAN Compiler 23.11

Table 1: Hardware Comparison

Results

Section	CPU Time (seconds)	GPU Time (seconds)	Acceleration
total time with IO	284.893	16.676	**17.08**
rk3 solver	284.444	15.794	**18.01**
intermediate steps	6.648	0.178	37.35
bchomo	9.431	0.012	785.92
rhscal	200.674	6.445	31.14
gradcal	38.011	0.008	4751.38
convection	76.157	2.468	30.86
diffusion	83.600	1.288	64.91
filterq	54.556	1.287	42.39
q2fvar	10.693	0.070	152.76

Table 2: Comparison of CPU and GPU Execution for a 128-3D test case

Section	CPU Time (seconds)	GPU Time (seconds)	Acceleration
total time with IO	2001.287	89.714	**22.31**
rk3 solver	1998.607	84.461	**23.66**
intermediate steps	47.334	1.195	39.61
bchomo	34.468	0.013	2651.38
rhscal	1433.342	17.275	82.97
gradcal	301.299	0.009	33477.67
convection	511.722	7.871	65.01
diffusion	600.195	5.234	114.67
filterq	388.198	4.977	78.00
q2fvar	76.878	0.477	161.17

Table 3: Comparison of CPU and GPU Execution for a 256-3D test case

- Using NVIDIA CUDA for GPU programming significantly improved performance. For the 128-3D case, execution time decreased from 284.893 seconds (CPU) to 16.676 seconds (GPU), achieving an acceleration factor of 17.08. For the 256-3D case, execution time reduced from 2001.287 seconds (CPU) to 89.714 seconds (GPU), achieving an acceleration factor of 22.31. (see Tables 2 and 3).
- The **gradcal** kernel was optimised by merging gradient calculations for the i, j, and k directions into a single CUDA kernel, leveraging parallel processing and shared memory.
- The **bchomo** kernel was optimised by parallelising boundary condition applications for the i, j, and k directions using CUDA, allowing simultaneous execution across multiple threads.
- The **convection** kernel was optimised by creating two separate CUDA kernels. However, the need to synchronise both kernels for operations in all three directions limits the acceleration factor.

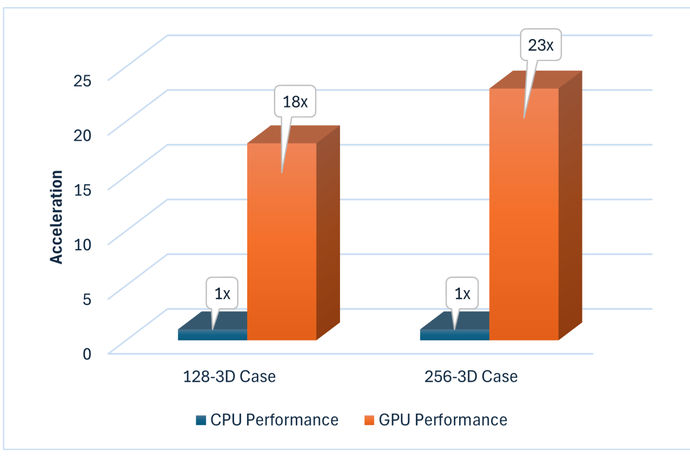


Figure 2: Acceleration over GPU

Future Work

- Enhance Fortran code portability by leveraging AMD ROCm's hipFORT to provide interfaces to the HIP API, enabling GPU acceleration on both AMD and NVIDIA GPUs, ensuring efficient performance and increasing flexibility.
- Supports various ROCm libraries (rocBLAS, rocFFT, rocRAND) to optimise performance for computational tasks, leveraging modern GPUs' full power.
- Boosts parallel processing capabilities of Fortran applications, paving the way for robust and scalable solutions in computational science and engineering.

References

[1] J. R. Bull and A. Jameson. "Simulation of the Taylor–Green Vortex Using High-Order Flux Reconstruction Schemes". In: *AIAA Journal*. Vol. 53. 9. 2015, pp. 2750–2761.

[2] Koen Hillewaert. "TestCase C3.5 - DNS of the transition of the Taylor-Green vortex, Re=1600 - Introduction and result summary". In: May 2013.

Tamil Spell Checker

Yazhmozhi Vasuki Murugesan

University of Dundee

The poster showcases the development of a spell checker for Tamil, which is an official language in Tamil Nadu, Pondicherry, Sri Lanka, and Singapore. This will provide a valuable resource for over 69 million Tamil speakers in India and will also aim to support Tamil writers who have dyslexia. This research examines how different language models such as LSTM, BiLSTM, mBERT, mT5, XLM-R and Llama perform with spelling error detection and correction of a highly inflectional language like Tamil. A systematic literature review of spell checkers for the major written Dravidian languages – Tamil, Telugu, Kannada, and Malayalam - has identified research gaps and challenges. This has led to the development of two Tamil language resources – "TamilCorp," a balanced corpus of Tamil written texts containing approximately 1.7 billion tokens across 16 genres and TamilSpell, a first-of-a-kind Tamil spelling error corpus with ~10 million entries to support the spell checker. TamilCorp will be used to train the aforementioned language models to detect and correct context sensitive errors. Both the test split from TamilCorp and context sensitive errors from TamilSpell will be used to evaluate the performance of the models. The performance of the deep learning models is compared to statistical language models as well. The TamilContextSpell, an open-source toolkit that features both a command-line interface and a graphical user interface, will be published along with the performance metrics of all models tested. A comparison will be made between dictionary-based hashing techniques and cross-lingual language models for the detection and correction of non-word errors. Additionally, a neurosymbolic approach will be used to detect and correct Sandhi errors in Tamil. This research also aims to demonstrate that even a small component of inclusion will make a significant difference in people's lives.

University of Dundee

Tamil spell checker

PhD Candidate: Yazhmozhi | Supervisors: Prof. Annalu Waller, Dr. Jacky Visser.

RQ1 - How can cross-lingual language models be effectively applied to improve spell checking in highly inflectional, low-resource languages like Tamil?

RQ2 - What is the impact on the accuracy and efficiency in comparison to traditional rule-based methods?

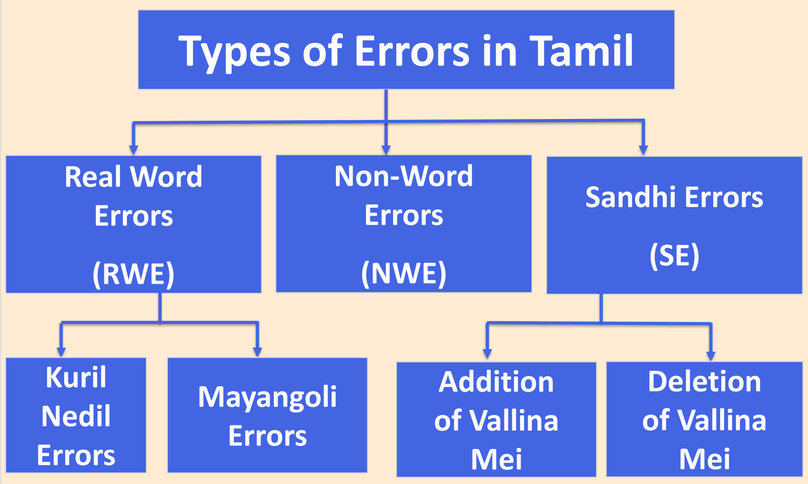
Tamil is a morphologically rich and highly inflectional language. With limited language resources, there are no fully functional Tamil spell checkers that cover all the errors that occur in the language.

Systematic Literature Review

Spell checkers for written Dravidian languages – Malayalam, Telugu, Tamil, and Kannada.

56 studies were identified from **9** databases.

44 studies were selected.



Research Gaps

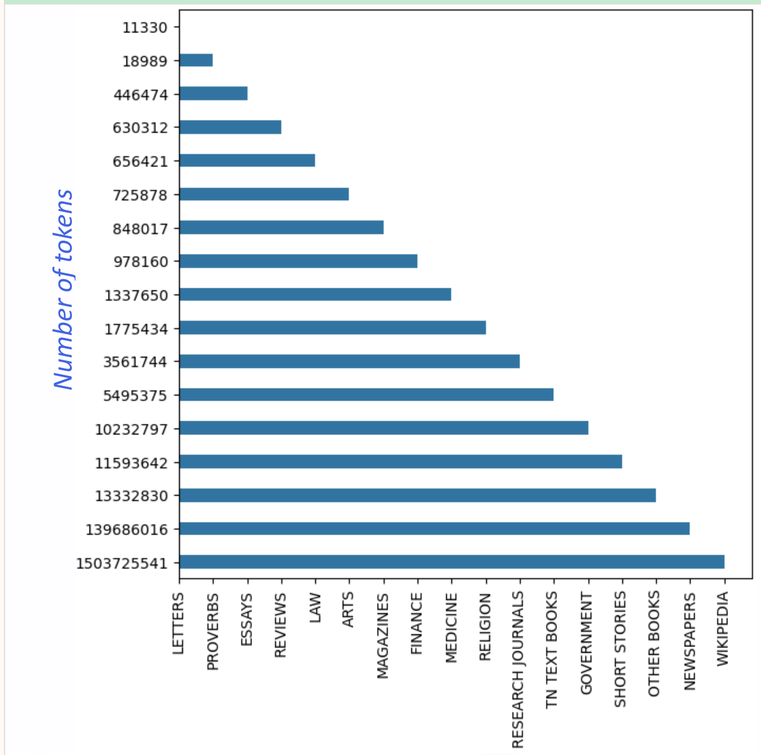
1. **Real-word errors and sandhi errors** need more research.
2. Minimal research using LLMs and other **ML /DL techniques.**
3. Scarcity of massive **publicly available datasets**.
4. No **benchmark test sets** with spelling errors.
5. No **suggestion adequacy** measures.

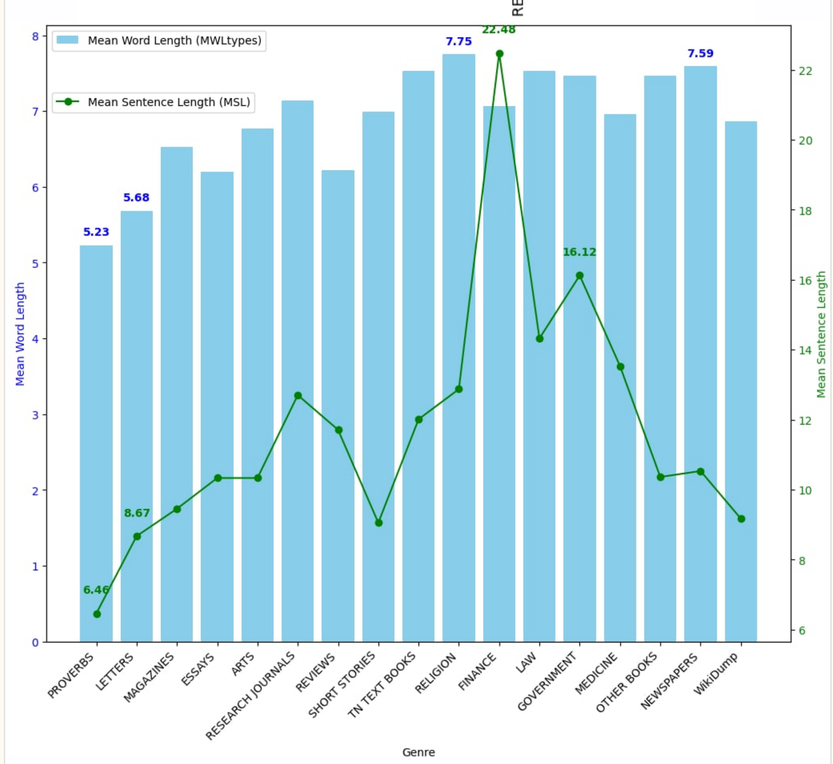
Corpus Construction

STEPS:

1. Identification of genres and sources
2. Data collection
3. Pre-processing — **17 genres**
4. Adding metadata
5. TEI XML conversion — **1.69 billion tokens**

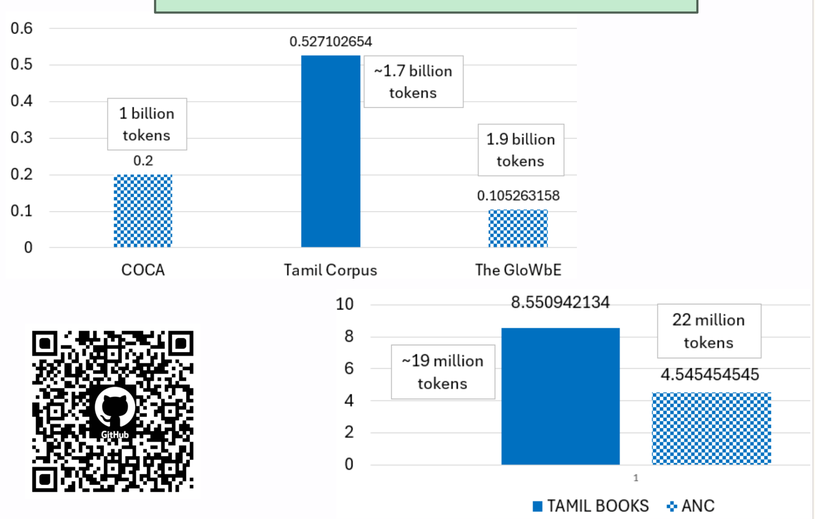
Corpus Analysis





Average word and sentence length

Tamil is highly inflectional?

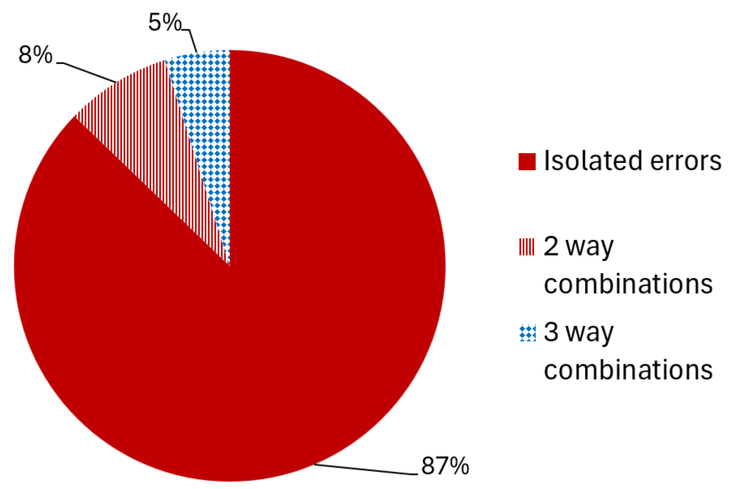


1. **Hypothesis -** 'Is there a significant difference in lexical diversity between Tamil and English?
2. Carried out **six independent two sample t-tests** and fond that the difference is significant with a very strong effect.
   1. Brown TTR and TamilCorp TTR
   2. COCA TTR and TamilCorp TTR
   3. COHA TTR and TamilCorp TTR
   4. Brown MTLD and TamilCorp MTLD
   5. COCA MTLD and TamilCorp MTLD
   6. COHA MTLD and TamilCorp MTLD
3. Open-source language resources published – Tamil dictionary containing around **8 million** Tamil words and word embeddings.

Tamil Spell

- A Tamil spelling error corpus.
- 10,157,662 entries.
- A four-column corpus < sentence with errors, correct sentence, error code, edit distance>

ERROR TYPE	CODE
Real word errors	RWE_M and RWE_KN
Non word errors	NWE_INS, NWE_DEL, NWE_SUB, NWE_TPN
Sandhi errors	SE_VA, SE_VD
X = {KN, M}, Y = {VA, VD} and Z = {INS, DEL, SUB, TPN}	RWE_X + NWE_Z
	RWE_X + SE_Y
2-way combinations	SE_Y + NWE_Z
3-way combinations	RWE_X + SE_Y + NEW_Z



Error distribution

ONGOING / NEXT STEPS …

To detect and correct real-word errors,

- XLM models to be finetuned with the corpus mBERT, mT5, and XLM ROBERTa
- Other models – LSTM, BiLSTM, and Llama.
- Comparative performance analysis with statistical bigram and trigram models to answer the research questions.
- Publish **TamilContextSpell,** an open source toolkit (6 models) with a CLI and GUI.

Future Work

Sandhi errors that require a neuro-symbolic AI approach will be a great area for future work.

To Contact me

SCAN ME

MetaWAAM: Real-Time Digital Twin Architecture for Direct Energy Deposition-Arc Manufacturing

Jinjiang Li

University of Manchester

Digital twins (DT) are increasingly used in manufacturing for real-time monitoring, process optimization, and fault prediction, but challenges remain in achieving seamless physical-virtual integration and real-time system optimization. This study leverages NVIDIA Omniverse™ to develop a DT architecture for a Direct Energy Deposition-Arc (DED-Arc) manufacturing unit, enabling real-time monitoring of robotic arm movements and welding states. The architecture integrates edge computing and finite element analysis (FEA) of laser-scanned 3D models, allowing on-site data processing and efficient transmission to high-performance computing (HPC) systems for rapid simulation. This scalable DT system provides immediate feedback to the physical manufacturing process, supporting high-quality additive manufacturing. As the first study to achieve real-time DT monitoring in DED-Arc manufacturing, it offers a foundation for future advancements in simulation and process optimization.

MetaWAAM: Real-Time Digital Twin Architecture for Direct Energy Deposition-Arc Manufacturing

Jinjiang Li[1**], Robin. C. Laurence[2], M.J. Roy[2], Lee Margetts[1]

1. Introduction

❖ What is a Digital Twin ?

Digital Twins (DT) bridge physical and virtual worlds, enabling applications such as real-time system monitoring, virtual simulation for process optimization and verification.

❖ What is direct energy deposition?

Direct energy deposition (DED) for metals, as realized with an electric arc, selectively melts a wire feedstock forming a molten pool which is deposited along a predetermined path to build a desired geometry.

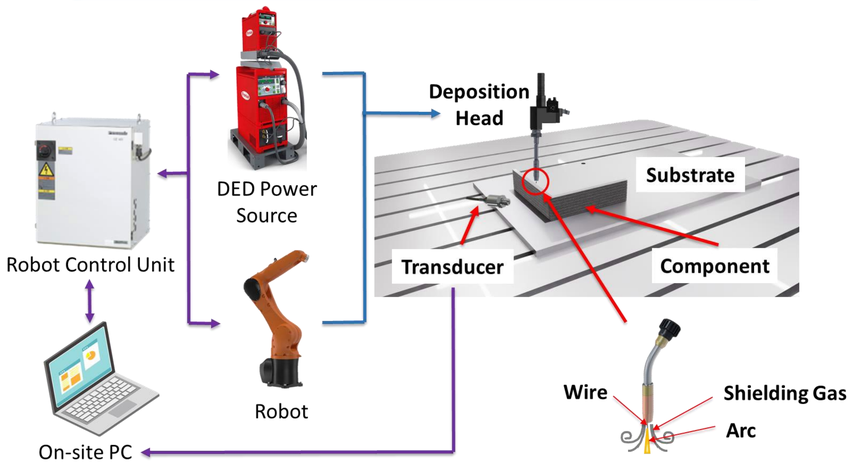


Figure 1. Introduction of DED-Arc manufacturing cell.

2. Challenges

- ❖ The simultaneous coordination of energy input, material deposition, and motion of robotic arms requires precise control to ensure consistent quality.
- ❖ Quality assurance and non-destructive testing techniques in-process are difficult to implement.
- ❖ Monitoring the deposition parameters in real-time is challenging due to the high speed and complexity of the process [1] .
- ❖ Simulation tools can be computationally intensive, and integrating them for real-time predictions in feedback lools have yet to be implemented.

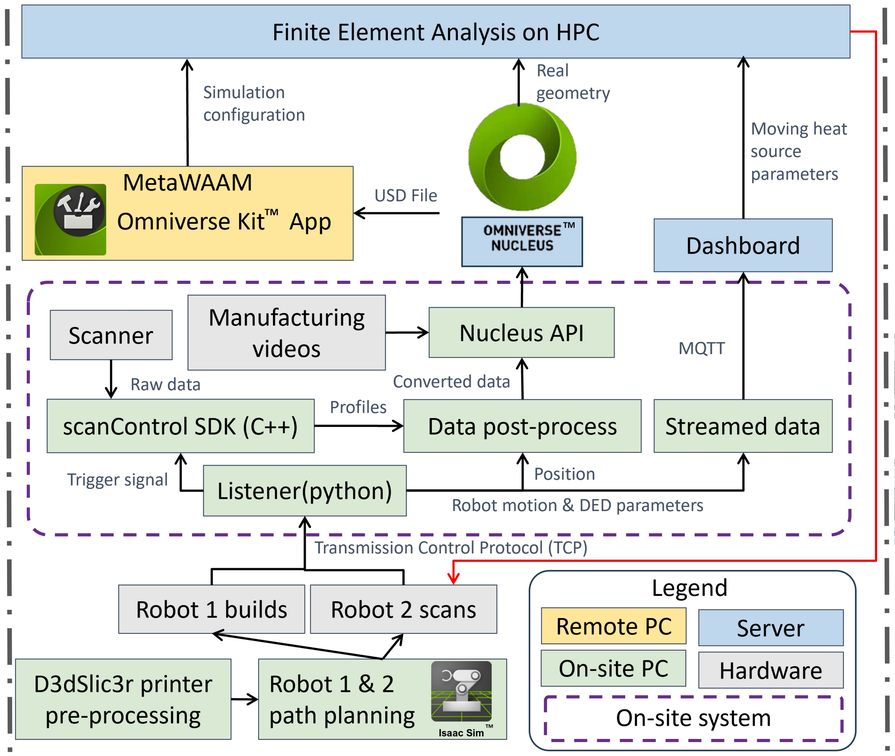


Figure 2. Architecture of the DT.

3. Methodology

- ❖ System architecture contains a multi-layer integrated physical, virtual and data management stack.
- ❖ Omniverse™ [2] includes software development kit (SDK), cloud services (Nucleus™), and a robot-based task development application (Isaac Sim™).
- ❖ Edge computing integration:
  - Slicer and the path planning generates robot instructions for the on-site PC.
  - The robot executes these instructions and publishes the realized manufacturing parameters via an industrial Ethernet connection.
  - High-frequency robot motion data is transferred through a MQTT protocol to a server.
  - File-based data is stored and logged on Nucleus™ server.
  - The developed app (MetaWAAM) monitors the manufacturing process and configures a simulation, which is then send to an HPC implementation.

4. Results

- ❖ Real-time monitoring a manufacturing process
- ❖ Physical manufacturing interlinked with simulation and cloud services via a digital twin.
- ❖ The creation of historical database on Nucleus™ to assist with data driven methods [3].
- ❖ Efficient edge computing integration allows distribution of the computational tasks.

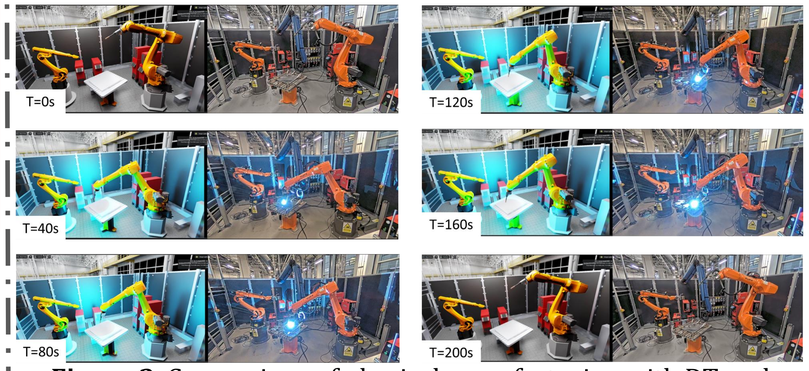


Figure 3. Comparison of physical manufacturing with DT real-time monitoring.

5. Extensibility

- ❖ This DT architecture is scalable to the verification of multiple scenarios.
- ❖ The connection with HPC system can handle complex simulation to provide predictive means for process improvement.
- ❖ The simulation result may be able to feedback real-time control.

References

.Mu, H., He, F., Yuan, L., Commins, P., Wang, H., & Pan, Z. (2023). Toward a smart wire arc additive manufacturing system: A review on current developments and a framework of digital twin. Journal of Manufacturing Systems, 67, 174-189. doi:10.1016/j.jmsy.2023.01.012
.NVIDIA. (2024).NVIDIA Omniverse. Retrieved from https://www.nvidia.com/en-gb/omniverse/
.Zhou, Z., Shen, H., Liu, B., Du, W., Jin, J., & Lin, J. (2022). Residual thermal stress prediction for continuous tool-paths in wire-arc additive manufacturing: A three-level data-driven method. And Physical Prototyping,(1), 105–124. https://doi.org/10.1080/17452759.2021.1997259

Acknowledgment: We would like to thank the Henry Royce Institute for Advanced Materials, supported by EPSRC grants (EP/R00661X/1, EP/S019367/1, EP/P025021/1, and EP/P025498/1), for providing the facilities used in this work. This project was also supported by the UK Atomic Energy Authority through the Fusion Industry Programme, which aims to stimulate the growth of the UK fusion ecosystem and prepare it for the future global fusion powerplant market. More information about the Fusion Industry Programme is available at https://ccfe.ukaea.uk/programmes/fusion-industry-programme/.

1 Department of Mechanical and Aerospace, School of Engineering, University of Manchester, Manchester, M13 9PL, UK | 2 Henry Royce Institute, University of Manchester, Manchester, M13 9PL, UK
**jinjiang.li@manchester.ac.uk

CIUK 2024

Accelerating Research with HPC: A Collaborative Study of Space Charge Compensation in H- Ion Sources

Kiran Jonathan

STFC- UKRI

Particle accelerators such as the ISIS Neutron and Muon Source (ISIS) have been crucial in probing our understanding of fundamental physics and have a wide range of applications across academia and industry. One challenge faced in the design of such accelerators is the space charge effects exhibited in the low-energy beam transport (LEBT) region, where charge repulsion can cause a defocussing of the beam, reducing the overall transmission and energy efficiency. These effects can be reduced via a process known as space charge compensation (SCC), through the addition of factors such as a neutral background gas. This project is a collaborative effort between the Scientific Computing Department (SCD) and ISIS, which aims to use a combination of real-world beam diagnostics and high-fidelity particle-in-cell simulations to develop a deeper understanding of the SCC process and its key parameters, ultimately informing design decisions for current and future iterations of the ISIS H- ion source.

Science and Technology Facilities Council

Scientific Computing

scd.stfc.ac.uk

Accelerating Research With HPC:

A Collaborative Study of Space Charge Compensation in H^- Ion Sources

Kiran Jonathan[1], Benzi John[1], David R. Emerson[1],
Erin L. Flannigan[2], Olli Tarvainen[2], and Daniel Faircloth[2]

[1]Scientific Computing Department, STFC Daresbury Laboratory, Warrington WA4 4AD, United Kingdom
[2]ISIS Neutron and Muon Source, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, United Kingdom

Abstract

Particle accelerators such as the ISIS Neutron and Muon Source (ISIS) probe our understanding of fundamental physics, with wide-ranging applications across academia and industry. One design challenge for such accelerators is the space charge effects exhibited in the low-energy beam transport (LEBT) region, where charge repulsion can cause beam defocussing and reduce transport efficiency. These effects can be reduced via space charge compensation (SCC), through the control of factors such as the neutral background gas pressure. This collaboration between the Scientific Computing Department (SCD) and ISIS aims to use a combination of real-world beam diagnostics and high-fidelity particle-in-cell simulations to develop a deeper understanding of the SCC process and its key parameters, ultimately informing design decisions for future iterations of the ISIS H^- ion source region.

ISIS Neutron and Muon Source

ISIS (Fig. 1) is a particle accelerator based at the STFC Rutherford Appleton Laboratory in Oxfordshire which produces neutron and muon beams. These beams have many applications across the physical sciences and can be thought of as non-destructive 'super-microscopes'. With around 3000 national and international users yearly, ISIS has world-wide scientific, socio-economic, and environmental impacts [2].

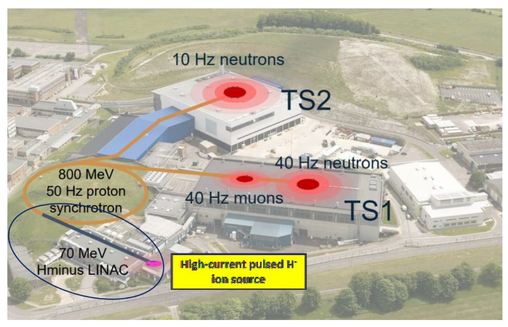


Figure 1: An overview of the ISIS facility. H^- ions are generated and accelerated in the linear region, before having their electrons stripped to produce protons. These protons are accelerated around the synchrotron, hitting target stations to produce neutrons and muons.

The Problem: Space Charge Effects

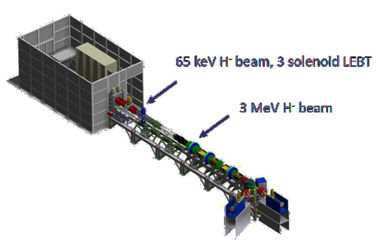


Figure 2: A diagram of the Front-End Test Stand (FETS) at ISIS, comprised of a Low-Energy Beam Transport (LEBT) and Radio-Frequency Quadrupole (RFQ) region.

This project is a collaboration with the Low-Energy Beam Group at ISIS, which focuses on the beam dynamics in the ion source and following low energy regions, such as the start of the LINAC region of the ISIS beam (Fig. 1), or in the Low-Energy Beam Transport (LEBT) region of the Front-End Test Stand (FETS) at ISIS (Fig. 2).

As the linear region of ISIS is where all of the ions, and therefore protons, originate, it has huge impacts on the overall beam quality and energy efficiency of the accelerator. As such, it is imperative that this region provides a focussed beam with minimal ion losses, in order to promote transport efficiency. Unfortunately, as the ions all have the same charge, they electrostatically repulse one another, causing the beam to diverge and lose ions to the wall (Fig. 3).

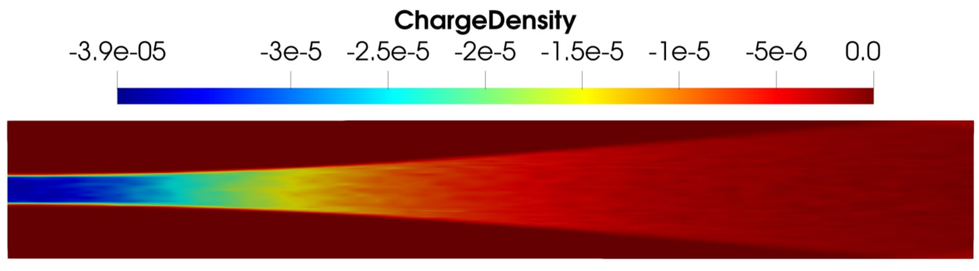


Figure 3: Simulation results for a beam without space charge compensation. Space charge repulsion between the ions in the beam causes it to diverge, losing ions to the accelerator walls.

kiran.jonathan@stfc.ac.uk

This work was supported by the Ada Lovelace Centre

The Solution: Space Charge Compensation

Space charge effects can be reduced through space charge compensation (SCC). Two common SCC methods are being studied in this project: external magnetic fields and additional background gas pressure.

Magnetic fields apply a force perpendicular to an ion's motion, which helps to curve their path and deter them from shooting out towards the walls of the accelerator. This helps to focus the beam and reduce losses.

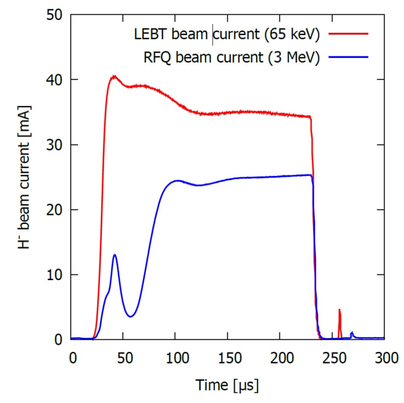


Figure 4: Measured H^- Beam Current in the Low-Energy Beam Transport (LEBT) and Radio-Frequency Quadropole (RFQ) regions of the FETS at ISIS.

Meanwhile, introducing a neutral background gas such as H_2 helps to reduce the repulsive force itself. It does this by introducing secondary positive ions, H_2^+, through an electron-stripping reaction: $H^- + H_2 \longrightarrow H^- + H_2^+ + e^-$. Over time, the beam generates more and more of these positive ions, and they act to reduce the overall charge density and therefore the repulsive force. In Fig. 4, you can see how the beam current in subsequent regions, such as the RFQ, increases over time as these ions are produced and act to focus the beam.

Leveraging HPC: Particle-In-Cell Methods

As we look to build more efficient accelerators, a deeper understanding of SCC and its key parameters is required. Computational methods such as the particle-in-cell method are extremely valuable for this as they allow us to quickly and cheaply explore many different parameter configurations. As the particle-in-cell method also simulates the individual trajectories of ions in the beam, it also allows for extremely detailed diagnostics to fully explore the problem from a theoretical standpoint.

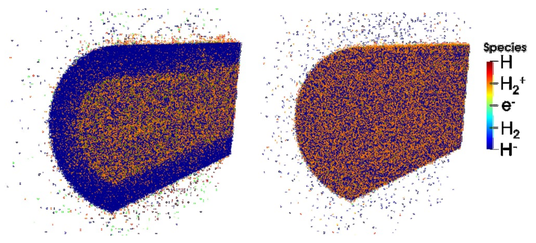


Figure 5: Simulated ion distributions at steady-state SCC for beam without (left) and with (right) an external magnetic field.

In this project, we make use of the open-source MPI-parallelised particle-in-cell code PICLas [1], developed at the University of Stuttgart. Fig. 5 shows an example of how we have used it to study the effect of an external magnetic field; the simulations show us that a magnetic field can prevent clustering of secondary ions in the centre of the beam.

Whilst these computational methods are very powerful, especially when leveraging HPC systems such as ARCHER2 and SCARF, the value of their results can only come from verification against real-world diagnostics. As such, with this unique collaboration between ISIS and SCD, we hope to make the most of both real-world data and high-fidelity simulations to fully explore the problem of SCC.

Future Developments

- **Study effects of background gas density and magnetic field strength**
- **Further comparisons between diagnostic results and simulations**
- **Simulation of oppositely charged H^+ ion source**
- **Collaboration with Indian Department of Atomic Energy to produce diagnostic results for H^+ ion source**

References

[1] Stefanos Fasoulas et al. "Combining particle-in-cell and direct simulation Monte Carlo for the simulation of reactive plasma flows". In: *Physics of Fluids* 31.7 (2019), p. 072006.
[2] *ISIS About — isis.stfc.ac.uk*. https://www.isis.stfc.ac.uk/Pages/About.aspx. [Accessed 01-12-2024].

Plasma Kinetic Code PICLas STUTTGART

ada lovelace centre

Science and Technology Facilities Council
Scientific Computing

MD-based spectroscopic simulations in Py-ChemShell

Maitrayee Singh

STFC - UKRI

Spectroscopic technologies, such as UV-vis and IR, are essential in molecular and materials science for uncovering molecular structures and interactions. However, interpreting experimental spectra accurately requires simulations that account for molecular dynamics. This poster introduces a new workflow in Py-ChemShell for MD-based spectroscopic simulations, designed to enhance spectral analysis by integrating molecular conformational fluctuations.Through the addition of a "Spectroscopy" task class, this workflow supports UV-vis spectral simulations on molecular snapshots from MD trajectories, offering a more realistic depiction of molecular behaviour over time. Key features include data averaging and convolution techniques for refining spectral outputs. Preliminary benchmarking shows that these simulations improve alignment with experimental spectra, making Py-ChemShell a more powerful tool for researchers.

Science and Technology Facilities Council

Scientific Computing

scd.stfc.ac.uk

MD-Based Spectroscopic Simulations in Py-ChemShell

Maitrayee Singh, Dr. You Lu
maitrayee.singh@stfc.ac.uk
Scientific Computing, Science and Technology Facilities Council

Introduction

Spectroscopic technologies such as UV-vis, IR/Raman, NMR, and EPR are essential tools in modern molecular and materials science. While these experiments provide critical insights into chemical structures, they are often combined with computational methods to interpret observations at the atomistic level. Py-ChemShell is a highly scalable, multiscale computational chemistry software package [1,2] that integrates quantum mechanics (QM) and molecular mechanics (MM) methods for studying complex structures and systems. It is capable of handling multiple independent computational tasks in parallel, i.e. task-farming parallelism (see **Figure 1**), by distributing them across multiple workgroups [2]. Specifically,

- A two-level parallelisation framework allows parallel execution within and across workgroups.
- MPI processes are divided into workgroups with independent tasks, and results are collected after all tasks are completed.
- Each workgroup has a master process that spawns a child Py-ChemShell lifecycle.

In this work, we present an automated workflow in Py-ChemShell that simulates spectra by accounting for the conformational fluctuations in molecular dynamics (MD) simulations. This new spectroscopic workflow significantly benefits from task-farming parallelism, providing excellent scalability.

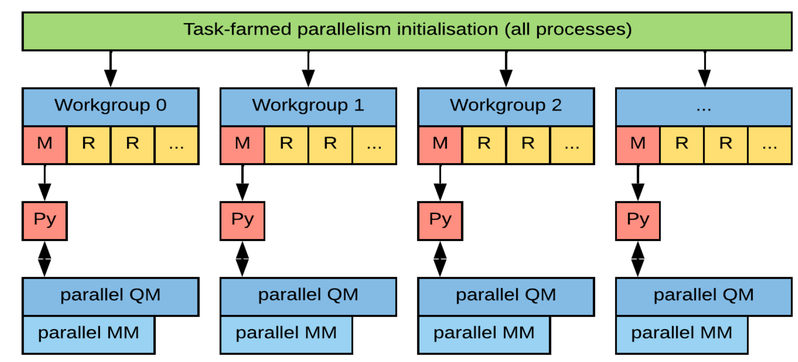


Figure 1. The framework of taskfarming parallelism in Py-ChemShell. M: Master; R: Replica; Py: Python interpreter [2].

Workflow of the project

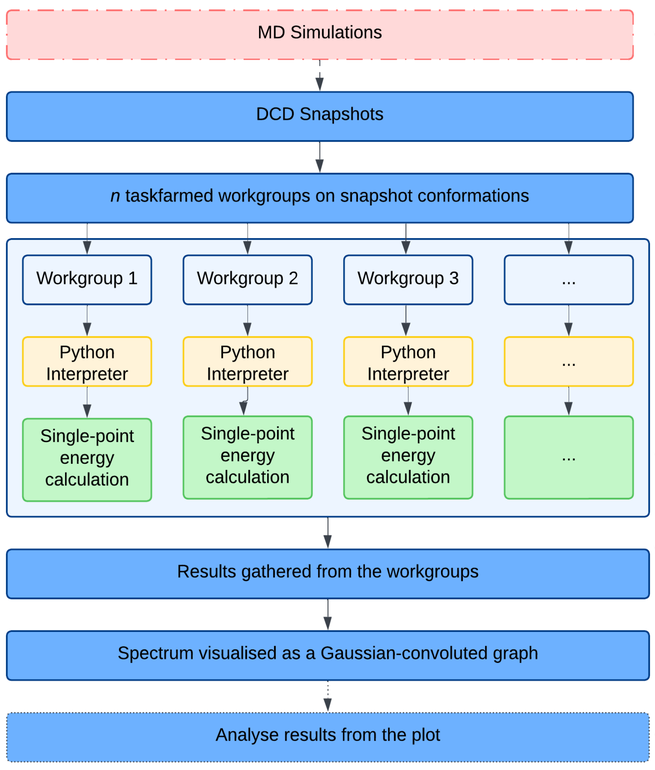


Figure 2. The implemented workflow for spectroscopic simulations in Py-ChemShell.

Figure 2 illustrates the workflow for conducting spectroscopic calculations based on MD conformations leveraging taskfarming for efficient computation. Py-ChemShell extracts snapshots evenly during a pre-run MD simulation's trajectory (typically in the DCD format) which represent conformations over the simulation time. The workflow sets up single-point spectroscopic tasks to run in parallel across workgroups. Each workgroup is assigned with a specific task managed by their individual Python interpreters.

Within each workgroup, single-point calculations are performed on the assigned snapshot structures to obtain the contribution to the spectroscopic properties. Once all workgroups complete their tasks, results are compiled from each workgroup and are used to generate a data series. These are typically convoluted using a Gaussian or a Lorentzian function, allowing visualisation of the overall simulated spectrum with *matplotlib*. This workflow is compatible with any theoretical levels that are available for spectroscopic calculations, either QM or QM/MM. A further analysis could be carried out to shed more light on the chemical system (see below).

Future work

We aim to develop a toolset for analysing the correlation between geometrical parameters and spectral evolution over time during molecular dynamics trajectories, along with enhanced plotting features. We will employ the new workflow in an application study.

References

[1] https://chemshell.org
[2] Y. Lu, et al., "Open-Source, Python-Based Redevelopment of the ChemShell Multiscale QM/MM Environment", J. Chem. Theory Comput., 15, 1317-1328 (2019)
[3] D. Bressanini, et al., "The infrared spectrum of L-proline: A comparison of experiment and theory". J. Am. Chem. Soc, 118(40), 6076-6086 (1996)

Results

Proline molecule was chosen to demonstrate and validate the implemented workflow. TD-DFTB (time-dependent density functional tight binding) is a fast quantum chemistry method derived from DFT. It simplifies calculations by utilising precomputed parameters, making it particularly effective for large-scale simulations, such as molecular dynamics or materials science, as it captures key properties like bonding, charge transfer, and excited states. DFTB+ is a specialised program designed for DFTB methods. As shown in **Figures 3** and **4**, DFTB+ was utilised to perform the demonstrative absorption spectroscopy simulation for a proline molecule in the gas phase. Multiple geometries of proline were extracted from an existing molecular dynamics (MD) trajectory in the DCD format.

```
# Defining coordinates of Proline atom from PDB files
proline = Fragment(coords="proline.pdb")

# Defining theory for the Spectroscopy task
dftb = DFTBplus(frag=proline, charge=1, skf_path="mio-1-1/", maxiter=2000, restart=False, nstates=5)

# User input for Spectroscopy task defining all arguments
my_spec = Spectroscopy(theory=dftb, trajectory='proline.dcd', type='absorption', nsnapshots=20, nsigma=20)

# job.run()
my_spec.run()
```

Figure 3. Example user input excerpt used for the demonstrative proline absorption spectrum.

The simulated spectrum was derived by applying a Gaussian convolution to the individual spectra of the snapshot geometries taken from MD trajectory. When compared to the experimental result, the simulated spectrum qualitatively reproduces the fingerprint absorption of proline, especially at 800-900 cm^{-1}. The deviation is reasonable due to the quality of DFTB; however, the accuracy can be improved by applying higher level quantum chemistry methods. The workflow's taskfarming parallelism ensures efficient computation of single-point energy calculations.

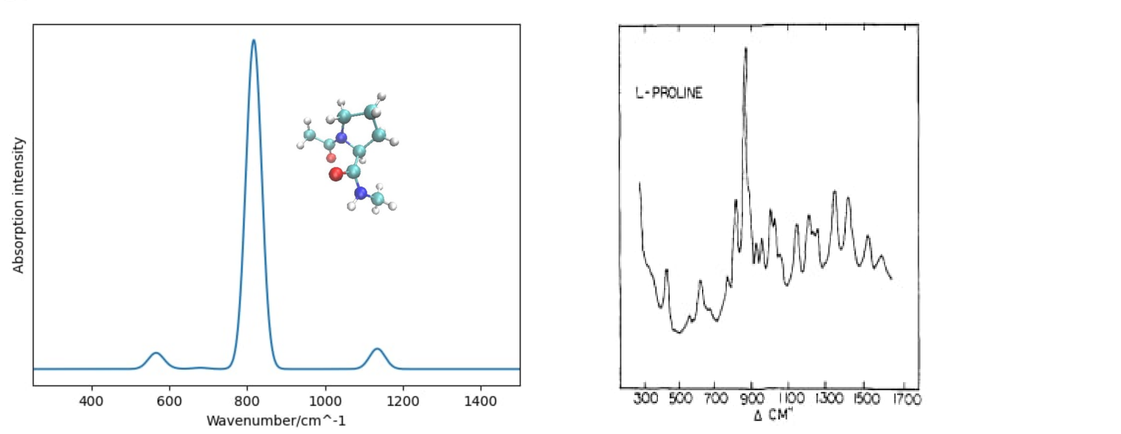


Figure 4. Left: Computed absorption spectrum of proline (example geometry shown as the inlet) in the gas phase using a Gaussian expansion. The calculation incorporated 10 snapshots sampled from the MD trajectory to capture the conformational change. 5 excited states are computed for each snapshot structure. Right: Experimental gas-phase absorption spectrum of proline [3].

ada lovelace centre

ChemShell multiscale computational chemistry

To learn more about ChemShell:

A case study of real-time collaborative design in FreeCAD and NVIDIA Omniverse

Raska Soemantoro

The University of Manchester

FreeCAD, a robust open-source CAD tool, is popular among small organizations and users seeking an efficient, bloat-free design platform, with many applications integrating it as a CAD backend. However, it lacks collaborative design features essential in modern engineering. This work introduces a FreeCAD connector for NVIDIA Omniverse, enabling real-time collaboration, version control, and photo-realistic rendering. Demonstrated through a case study of engineers designing a sensor for a stellarator fusion plant, the connector enhances FreeCAD workflows by connecting to Omniverse's Nucleus—a secure, central storage platform compatible with third-party tools. This integration offers FreeCAD users a collaborative, non-proprietary CAD solution, expanding design capabilities for engineers and creators within Omniverse.

A case study of real-time collaborative design in FreeCAD and NVIDIA Omniverse

Raska Soemantoro, Lee Margetts
raska.soemantoro@manchester.ac.uk

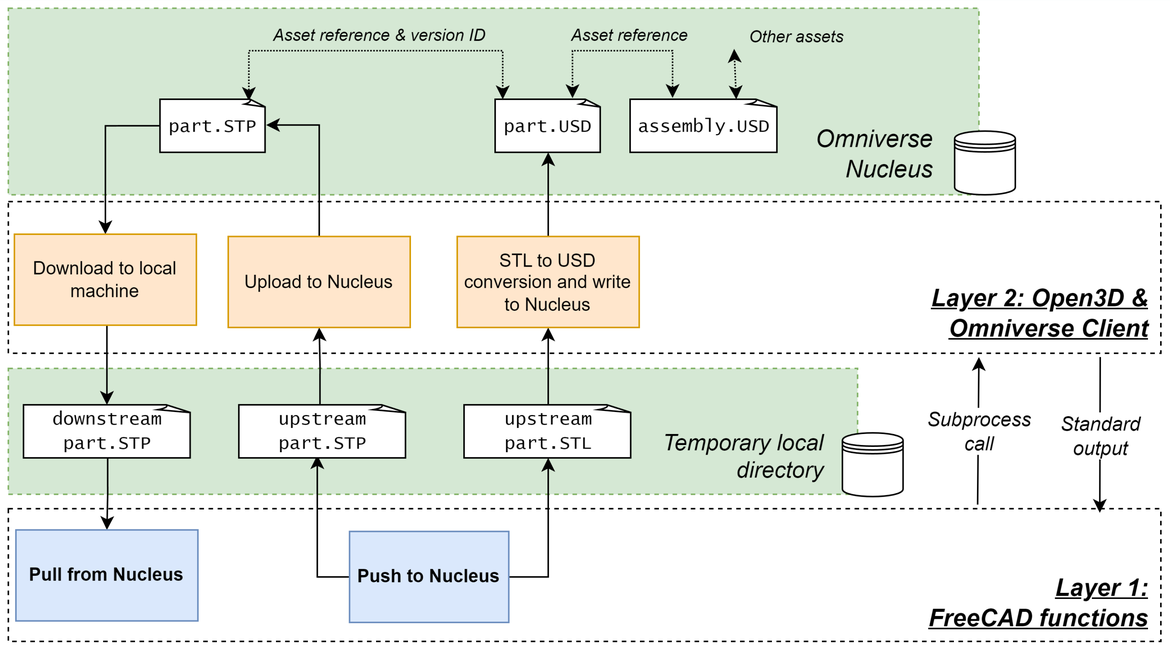
The University of Manchester

Background

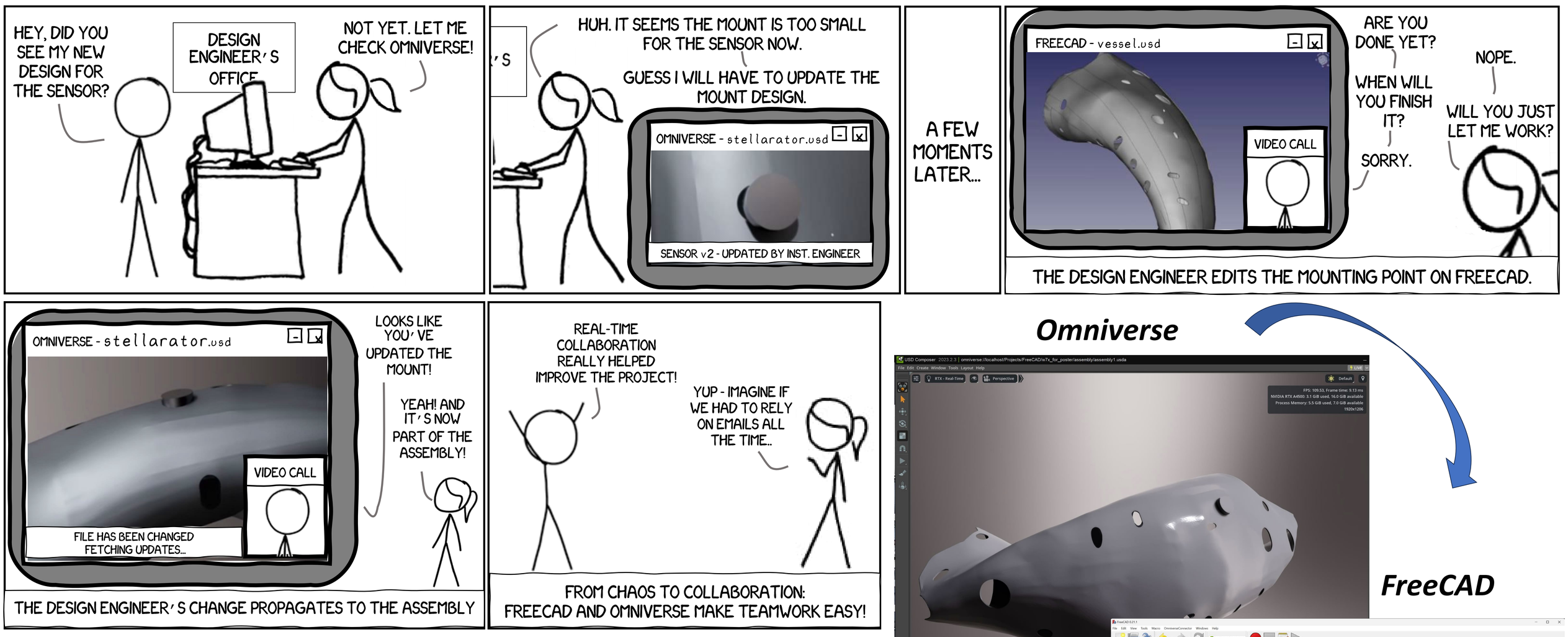
- Managing CAD file versions often feels like a game of tennis—but with no clear winner.
- Email-based workflows are slow, messy, and riddled with version control headaches.
- Without real-time collaboration, reviewing and aligning changes becomes a tedious process.
- A centralised, version-controlled platform can eliminate such issues, streamlining the workflow.

When FreeCAD met Omniverse

- FreeCAD, an open-source CAD tool, is popular those seeking a cost-effective, bloat-free platform.
- Despite its advantages, FreeCAD lacks native collaborative design features essential for modern, teamwork-driven engineering workflows.
- NVIDIA Omniverse is a platform that allows users to integrate OpenUSD, simulation, and rendering technologies into existing software tools.
- A connector tool has been developed, which introduces collaborative design functionality by connecting FreeCAD users to Omniverse's shared environment.



Collaboration made easy!



- FreeCAD users can now edit, review, and align CAD files in real time without endless back-and-forth emails.
- Omniverse's version control ensures everyone is always on the same page.
- Real-time collaboration allows engineers to interact within a shared scene.
- Integration with Omniverse enables access to simulation and photo-realistic rendering capabilities.

Engineering and Physical Sciences Research Council
UK Atomic Energy Authority

This project has been supported by UK Atomic Energy Authority through the Fusion Industry Programme. The Fusion Industry Programme is stimulating the growth of the UK fusion ecosystem and preparing it for future global fusion powerplant market. More information about the Fusion Industry Programme can be found online: https://ccfe.ukaea.uk/programmes/fusion-industry-programme/

FUSION CDT

CIUK 2024 Breakout Sessions

Technical/SysAdmin Meetup

A meeting for technical staff and systems administrators to discuss the latest topics in the field and share information and ideas. This will be an unconference-style event, where the technical staff are able to steer the discussion to a variety of topics of interest. This meeting has been very popular for the last two years, and we expect a similar level of interest this year. If you have any topics of particular interest you feel strongly about, feel free to register them in advance.

STEP-UP - developing HPC technical professionals:

The STEP-UP project aims to change the landscape for digital Research Technical Professionals (dRTPs) within the London and South East of England region, while sharing developments, approaches and findings with the wider community, across the UK and beyond. We class dRTPs as anyone undertaking work related to software, data and computing infrastructure / High Performance Computing (HPC) within the research community. In this session we'll focus on the HPC-related aspects of STEP-UP's activities. How did you become an HPC technical professional? How do you progress as an HPC technical professional? What's missing to support the next generation of HPC experts? Share your journeys and learn about the STEP-UP project and how we're looking to support the HPC community. The session will begin with a couple of short introductory presentations to set the scene. However, this will be an interactive session where we would like to hear from you about the routes you have taken your journeys into your roles, your learning journeys and your perspectives on career opportunities within the HPC community. The information shared in this session will help us in our identification of career and learning pathways for HPC professionals' careers and the skills and career professional development requirements. As part of its activities over the next few years, the STEP-UP project will also be setting up a programme to support short-term secondments to help technical professionals develop their specialist expertise and share skills, and creating a mentoring programme. We want to know how you think you could benefit from such opportunities and what you'd like to see within these schemes? What can we offer you to help further develop the HPC community, address challenges around skills shortages and recruitment, and provide more opportunities for technical engagement across institutions. We'll close with a summary, thoughts around next steps and an opportunity for participants to provide their contact details to engage further with us or receive updates on the work of STEP-UP.

Portable benchmarking and profiling using ReFrame

Testing and benchmarking scientific applications on HPC systems still often relies on manual effort and specialist knowledge. This makes studying application performance across HPC systems time consuming and error prone. In this session we show how building, running and measuring the performance of applications can be automated to develop benchmarks across HPC systems using ReFrame. ReFrame is widely used for regression testing by HPC services from small group clusters up to the largest supercomputers in the world (including the UK national supercomputer, ARCHER2). Together with Spack and lightweight profiling tools, we have used ReFrame to build an automated benchmarking pipeline for the ExCALIBUR Hardware and Enabling Software program for cross-platform benchmarking. The workshop will introduce using ReFrame to automate benchmarking and collecting performance data. Participants will learn to write a simple ReFrame test, run it on a HPC platform, measure its performance and compare against achievable peak performance of the system. Following the session, attendees should

have enough knowledge to start using ReFrame in their day-to-day activities. The workshop is aimed at attendees who have an understanding of basic HPC concepts and are interested in the performance of scientific applications. This includes both system administrators who run benchmarks to monitor the performance of a system, as well as application developers who are interested in performance benchmarks in their development workflow. Attendees will need to bring a laptop with an SSH client installed to access a remote facility to develop and run the ReFrame tests. A working knowledge of Python is required to develop tests in ReFrame. We will provide a repository with setup instructions and examples that will be communicated to attendees before the session starts and during the session. The workshop will be taught by the developers from the ExCALIBUR benchmarking project, and ARCHER2 service staff who have experience of using ReFrame in production. We will provide access to ARCHER2 and example applications to work on. The workshop is a mix of short presentations, carpentries-style live coding, and hands on sessions with helpers.

Cybersecurity and Federation for National DRI, AI and HPC Resources

The threat faced by the UK research and education community from cyberattack is persistent, with well publicised incidents against members of our community both nationally and internationally. Digital Research Infrastructures involving these organisations form a critical part of the supply chain of the UK and thus we must evolve a collective, collaborative approach to our defence in the face of this threat. Organisers of this session include representatives of the DRI Cybersecurity community including JISC and national HPC and AI Research Resource providers. With the HPC focus of CIUK, we seek to bring together a community of practice to share ongoing activities and engage with the wider CIUK participants In parallel, the research communities have ever- increasing needs for securely sharing data and accessing powerful computing resources within a federated ecosystem; the impact of the risks involved with these requirements must be carefully assessed so that we can provide secure, assured access to our national resources. A tentative outline of the session:

- Welcome and introduction
- Overview from the organisers (DRI Cybersecurity, JISC, and AIRR)
- Cybersecurity topics
- AAI topics
- Open discussion
- Close and next steps/events.

Storage Scale (GPFS) User Group / New Users Session

GPFS User Group CIUK 2024 Planned to be 2 hours, but with a new User session extending it to three hours. Continuing on from user groups at CIUK from a number of years, we want to run another Storage Scale User group at CIUK 2024. The user group will combine user talks, Vendor presentations and IBM engineer Talks around the user of and future directions of Storage Scale. The user group beings together representation from media, academia, finance, research, automotive, defence and pharmaceutical industries. We would also like to continue with the new user sessions that have been successfully run at previous user group meetings including at the event hosted at IBM in June 2024. These sessions have been well received and are helping to build expertise for the future to fill the admin roles required to run these systems. The User group continues to be run by users for users, and is not run by IBM, but we work closely with IBM. The Spectrum Scale User Group aims to:

- Bring together users of Spectrum Scale and Spectrum Scale with Spectrum Protect ILM (TSM HSM) into a collective environment
- Represent the needs of the Spectrum Scale User Group members

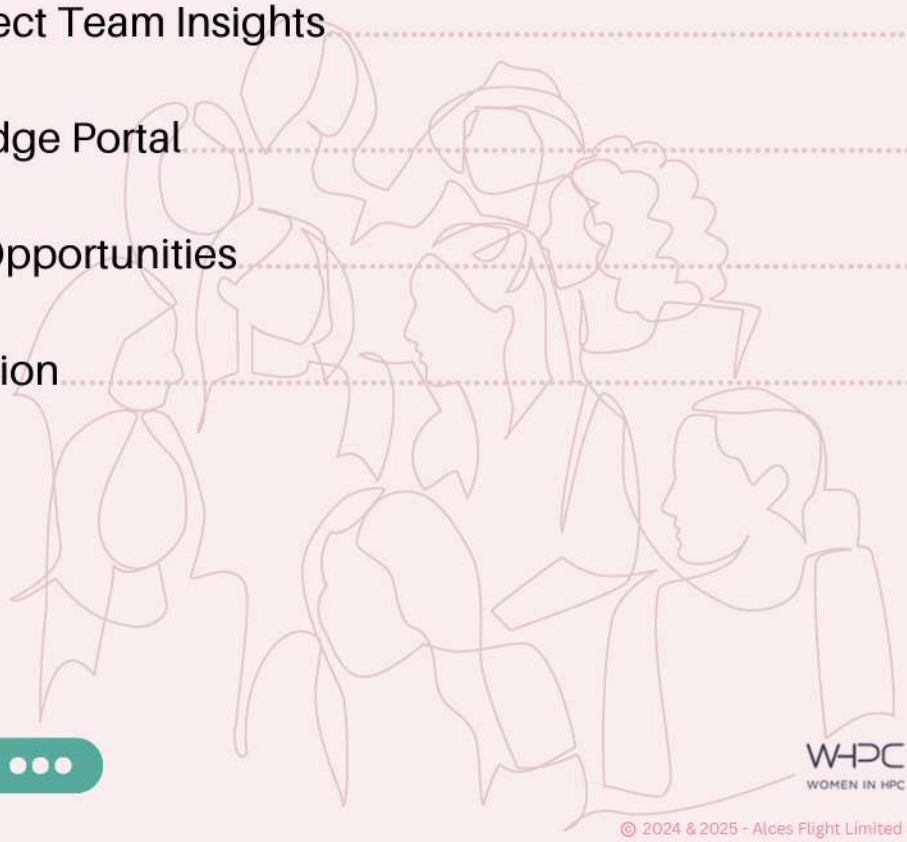
- Liaise with IBM and our User Group members to improve Spectrum Scale and Spectrum Scale with Spectrum Protect ILM
- Provide a “localised” community for support and knowledge sharing
- Raise awareness of Spectrum Scale as a capable data management platform I can provide a more detailed description if needed.

MOVE THE NEEDLE

2024 Project Report

WHPC
WOMEN IN HPC
x
alcesflight

TABLE OF CONTENTS



INTRODUCTION

The 2024 Move the Needle project, founded by Alces Flight, aimed to foster a positive and inclusive culture in High-Performance Computing (HPC) by promoting equality, diversity, inclusivity, accessibility (EDIA), and sustainability in the field.

To strengthen this effort, we partnered with Women in HPC (WHPC), an organisation dedicated to increasing the participation of women and underrepresented groups in supercomputing. WHPC's network of volunteers played a critical role in the success of this project.

Throughout 2024, we gathered insights on workplace inclusion initiatives and developed an accountability program to evaluate their impact on advancing EDIA. Here's what we've learned.

PROJECT OVERVIEW

The "Move the Needle" project, which took place from January to December, 2024, invited HPC students, enthusiasts, and professionals to commit to one to three actions aimed at advancing inclusion. We tracked individual and organisational efforts through surveys and online check-ins, with the goal of analyzing these findings and publishing the results at the Computing Insight United Kingdom (CIUK) conference in Manchester, UK, on December 5-6, 2024. Our aim was to demonstrate achievements, celebrate successes, and address barriers.

Initially, the project focused on building an Accountability Team (A-Team), who submited three surveys detailing their goals, progress, and lessons learned. Online meet-ups supplemented these surveys, though participation was optional due to the global distribution of A-Team members.

The A-Team began with ten core members, including three leading team projects, supported by two project managers from Alces Flight. After the second survey, one member left, leaving nine members who successfully completed the project.

PROJECT OVERVIEW

Due to community interest, the project expanded in 2024 to include a Knowledge Portal and a series of in-person meet-ups and sessions. The Knowledge Portal showcases video interviews and blog posts from members of the HPC and tech communities, highlighting their efforts to advance inclusion. In total, ten of these interview projects were completed across a range of subjects.

We also participated in several in-person and online events, supported five UK-based training grant applications for 2025 (two of these have already been successful), and celebrated as a member of the A-Team received an award for his contributions to fostering inclusive communities at his university.

Project At-A-Glance

- **Duration:** 12 months (January - December 2024)
- **Participants:** HPC students, enthusiasts, and professionals
- **Actions:** Commitment to 1-3 actions aimed at advancing inclusion
- **Tracking:*** Surveys, events and online check-ins for individual/organizational efforts
- **Goal:** Analyze findings and publish results at the CIUK conference (December 5-6, 2024, Manchester, UK)
- **Purpose:** Demonstrate achievements, celebrate successes, address barriers

* This project centered on the unique goals and needs of its participants, which may not be fully captured through standard surveys. As such, we encouraged and have accepted alternative reporting methods. By allowing participants to contribute in ways that suit them best this supports our aim of fostering an inclusive and equitable environment, respecting individual preferences, and enriching the project with diverse perspectives and insights.

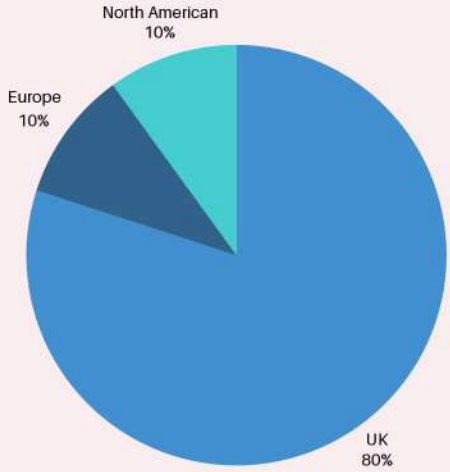
ACCOUNTABILITY TEAM

The Accountability Team, or "A-Team," consisted of nine members plus two project managers. Voluntarily recruited between December, 2023 and February, 2024, these individuals pledged up to three goals that they agreed to track over the project cycle.

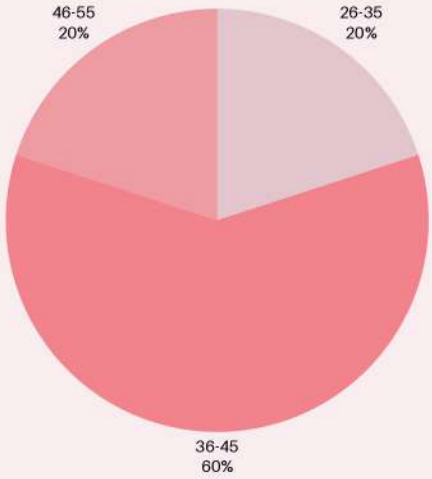
Their surveys, plus the online and in-person meet-ups, form the foundation of this report.

A-Team participants kindly submitted basic demographic information about themselves. This survey concluded with responses from the initial 10 individuals, with an equal gender distribution of 50 % male and 50 % female. The dataset encompasses various aspects of participants backgrounds, highlighted in the following page:

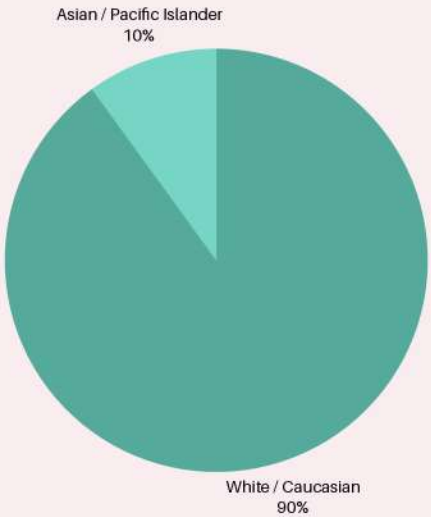
A-TEAM DEMOGRAPHICS



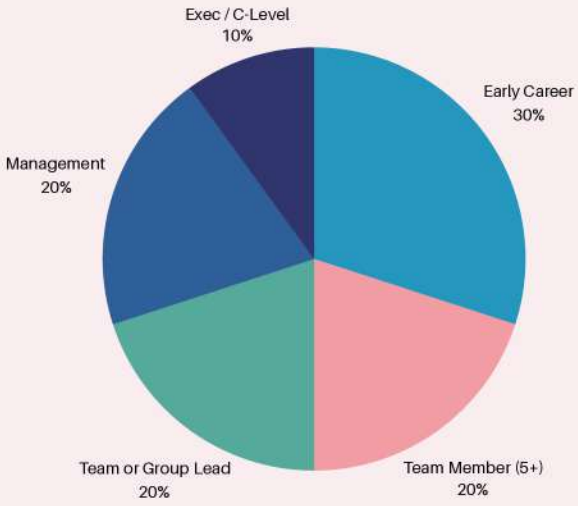
Geographical Location



Age Range of Participants



Race and Ethnicity



Career Level

A-TEAM TOPIC FOCUS

CORE AREAS

In February, 2024, our A-Team committed to goals in specific areas of focus. Utilizing surveys submitted, along with alternative reporting methods, our A-Team primarily focused on three key areas:

1. **Staffing:** Emphasizing hiring practices and improving overall staff retention.
2. **Engagement:** Fostering mentorship, advocacy, allyship, collaborations, partnerships, and coaching.
3. **Long-Term Impact**: Ensuring business and institutional continuity and promoting environmental sustainability.

We are pleased to report that none of our A-Team members needed to make significant changes to their goals or areas of focus during the project period. While some barriers emerged, these proved to be an aid to developing best practices, which can be seen in our "Lessons Learned" section.

STAFFING

Focusing on:

- New role creation
- Updating roles + responsibilities
- Clear HR policies and practices

ENGAGEMENT

Focusing on:

- Collaborations and Partnerships
- Mentorships, Advocacy and Allyship

LONG-TERM

Focusing on:

- Legacy and Institutional Evolution
- Environmental Sustainability

NOTED ACHIEVEMENTS

Over 2024 the A-Team, as well as the project at large, has exceeded initial expectations in several areas. This includes:

STAFFING

- Successfully utilising EDIA checklists to improve job advertisement and hiring practice.
- Re-establishing core role responsibilities to help define better career trajectory.
- Researched, wrote and received approval for HR strategy around defining and improving policy around leaves of absence.

ENGAGEMENT

- Mentorship opportunities in the HPC space located, explored, and prioritised.
- Elevation of 'bridge building roles' in HPC/AI (Research Software Engineer, Community Manager) commenced.
- Intersectionality interests noted and definitions created for establishing better guidelines in event/conference planning.
- Improved news and events distribution and information for those interested in accessing or working with HPC.

LONG-TERM

- Established review process on ecological sustainability and impact on job roles.
- Established review cycle of company culture and building sustainable work environments.
- Successful buy-in across A-Team projects for leadership involvement (low to high level).

AWARDS & FUTURE

We are thrilled to note that over the 2024 cycle the A-Team was able to achieve stretch goals beyond the intial boundaries of the project:

- Successful launch of WHPC - University of Cambridge Chapter (November, 2024).
- Successful application for UKRI Training Grant, HAI-End - Dr. Tobias Weinzierl
- Successful application for training grant: DisCouRSE: Developing a Community of Leaders - University College London (UCL)
- Successful establishment of award for female student excellence in HPC - nAG
- Award Recipient for Creating an Inclusive Community, University of Cambridge - Dr. Tom Meltzer.

The A-Team concluded 2024 with foundational goals to continue their work within this space. Noted areas of exploration include:

- Continued improvement of hiring practices.
- Continued engagement on improving retention practices - focus on job support.
- Support of training grants incorporating insights from Move the Needle.
- Project foundations utilising intersectionality established for:
  - Organisation around day and short conferences.
  - Software and Workflow development and feedback process.

LESSONS LEARNED

BEST PRACTICES

Through the work achieved across the 2024 project cycle, we've identified six essential best practices that are critical for the success of any EDIA or sustainability initiative. These best practices serve as key considerations to guide and inform meaningful, lasting impact in the HPC community. They are:

1

Clarify Your Motivation

Determine whether you're pursuing this initiative because you're genuinely committed or simply obligated. Authentic motivation fosters stronger engagement and better outcomes.

2

Set Clear and Adaptable Goals

Define specific, measurable goals that align with the needs of your team or organization. Be prepared to adapt them as circumstances change to stay on track.

3

Secure Leadership Support

Ensure leaders are actively involved and supportive. Their engagement significantly increases the likelihood of achieving your goals.

LESSONS LEARNED

BEST PRACTICES, CONT.

While not wholly representative of every goal set explored by our Move the Need A-Team, at their core these best practices represent a solid set of key components necessary in securing a successful outcome.

4

Ensure Inclusive Representation

Involve all relevant stakeholders. Excluding the right voices can undermine the initiative and limit its success.

5

Build Accountability Mechanisms

Assign ownership for each goal and establish systems to track progress, report outcomes, and access support when needed.

6

Invest in Resources and Rewards

Allocate sufficient resources, whether through budget, time, or tangible rewards (e.g., monetary incentives, events, or recognition). Investment underscores commitment and motivates success.

Ready to set your own goals?

We built a Pledge and Goal Guide you can download!

https://alces-flight.com/move-the-needle-knowledge-base/

S.C.A.L.I.N.G.

CHEAT SHEET FOR CHANGE!

In addition to Best Practices and Project Insights, we've developed a cheat sheet to help others kick-start their own Move the Needle journey. You'll notice that the first two letters of our acronym focus inward—this reflects a key lesson from our team: meaningful change often begins with taking stock of what already exists. This approach helps establish goals and metrics that align with your organization or group, creating a more positive and effective path forward.

Size of organisation (= **Speed** of change!)

Culture of your institution, department, group

Ally Engagement

Leadership Buy In

Investment (Time/Money or Both?)

Needs and Motivation

Goals and Accountability

PROJECT TEAM INSIGHTS

MARTA CAMPS SANTASMASAS

Lecturer, University of Salford

Goals focus: *Mentorship, Community Engagement and Personal Development*

Coming into the Move the Needle project, I had clear goals for what I hoped to gain but grappled with the question, "What can I contribute back?" I set goals centred on the individual experience, and what I discovered was having a supportive team behind me made those personal goals not only achievable but far easier to accomplish. Being in a supportive environment meant that not only could I accomplish my goals - but I could help other A-Team Members achieve theirs. It was a win-win contribution!

GOALS ACHIEVED:

- Mentored my first early career speaker.
- Engaged with two EDI groups at the University of Salford (Athena Swan and Women in CSE) and contributed to organise the Insight Summer School.
- Contributed to two online A-Team Meetings and served as Distinguished Speaker at Sustainable Reality and the CIUK24 Move the Needle Finale.

LESSONS LEARNED:

- Being part of a support/accountability team made it far easier to achieve goals.
- Starting small in mentorship (reviewing posters, helping with presentations) is an easy way to give back - and build confidence to do more.
- Being part of different communities is a great way to develop personal and professional talents.
- Personal development is not only achieved through attending courses, but also through getting involved with new projects and developing relationships.

PROJECT TEAM INSIGHTS

TOBIAS WEINZIERL

Professor, Head of Scientific Computing Research Group and Director of the Institute for Data Science at the University of Durham

Goals focus: *Personal development, Community and Educational development*

My interest in the Move the Needle project stemmed from a desire to make equity, diversity, and inclusion (EDI) as well as accessibility central to my work. As scientist and developer behind major scientific code development projects and PI of research consortia, I know the importance of fostering inclusivity and want it to be built into my work. Otherwise, we will not be able to get the best science out of our team, will not find the right team, and we will not disseminate our work in the best possible way.

The project challenged me to ask critical questions about how we integrate EDI aspects into our day to day software development, collaborate with a supportive team, and lay the groundwork for future initiatives.

GOALS ACHIEVED:

- Balanced event coordination across gender - Durham HPC Days
- Critically explored the role of inclusivity within scientific software development workflows – a write-up of lessons learned (similar to the recipes in eXtreme Programming) is to be published soon
- Built EDI considerations into the design and dissemination of HPC training – see the newly started Digital Research Infrastructure grant HAI-End

PROJECT TEAM INSIGHTS

TOBIAS WEINZIERL, CONT.

LESSONS LEARNED:

- Build Event Momentum by Expanding Perspectives:
  - Invite early-career speakers to events and consider gender balance.
  - Create sessions that appeal to complementary fields.
  - Involve community groups (e.g. WHPC) right from the start in event planning.
- Make Inclusivity a Continuous Effort:
  - Establish accountability measures to stay focused.
  - Don't be afraid to ask questions and seek fresh ideas from your network—it's how innovation happens.

Photos from previous Durham HPC Days

PROJECT TEAM INSIGHTS

TOM MELTZER

Senior Research Software Engineer, Institute of Computing for Climate Science, University of Cambridge

Goals focus: *Hiring and Retention, Community Engagement*

I joined the Move the Needle project to better understand what it truly means to be an ally in advancing EDIA. Initially, I worried about making mistakes, but I've learned that showing up and contributing meaningfully is what matters most.
I'm incredibly proud of the team we've built here at Cambridge and the progress we've made in raising the profile of women and underrepresented groups in our community. It was an unexpected honor to have my efforts recognized this year with the university's Creating an Inclusive Workplace award—a testament to the collective work and shared commitment of our team.

GOALS ACHIEVED:

- Launch of Women in HPC Chapter: Cambridge and East Anglia.
- Reviewed and identified improvements in hiring policy.
- Successfully established a Move the Needle Working Group with mix of senior and emerging leadership.

LESSONS LEARNED:

- Tips on being a good ally:
  - Show up: Start quietly - Listen first before acting.
  - Seek permission: Want to contribute? Ask where help is needed.
  - Shine the light: Allies have the power to lift people up - be sure you are bringing the team with you.
- Break big goals down into manageable pieces. Big task at hand? Don't be afraid to take small steps in order to find out how many targets you can hit.

KNOWLEDGE PORTAL

in 2024 we led discussions and activities with eleven experts in High-Performance Computing, Education, Coaching, Volunteering, Community Building, Staff Development and Mentorship. These interviews have provided insights into how organisations and individuals are advancing diversity and inclusion, as well as pushing boundaries in our field. Here is a brief summary of the knowledge we gained:

DAN OLDS AND **ADDISON SNELL**

Student Engagement in HPC

The Winter Classic Invitational Student Cluster Competition aims to address two key objectives: attracting young talent from diverse communities, particularly Black and Hispanic universities in the United States, and providing students with hands-on experience in High-Performance Computing.

This project is so large it took two interviews! Explore how project engages students with Dan Olds and to find out more about the project history watch our interview with Addison Snell.

VASHTI WHITFIELD

Coaching and Legacy Building

Vashti's insights illuminate the process of reframing challenges, establishing a lasting legacy, and the essential qualities to consider when selecting a coach. Her coaching methodology integrates psychological, neuroscience, and emotional intelligence aspects, emphasising the initiation of change from within.

Her interview covers her coaching style, and how it relates to building better technology, as well as social legacy.

KNOWLEDGE PORTAL

JAY LOFSTEAD

Mentorship and Advocacy

In his interview, Jay gave valuable insights into various types of mentorships, the significance of advocacy, and talked about his work in relaunching the WHPC Mentorship Programme. He highlighted the roles mentors can play in mentorship, sponsorship, coaching and advocacy.

Included in his interview are considerations for selecting the right mentor, details about WHPC Mentorship Programme, and advice on establishing a mentoring initiative.

NAGES SIESLACK

Creating Inclusive Events

Since 2011, the ISC High Performance Conference Series has actively pursued initiatives to foster an open and inclusive environment at their events. Central to their approach is the principle of 'Put People First', which guides their initiatives and has become integral to the conference's identity.

In her interview, Nages walks us through the history of ISC in regards to how they have and continue to build 'social sustainability' for the people of HPC.

https://alces-flight.com/move-the-needle-knowledge-base/

KNOWLEDGE PORTAL

LUDOVIC CAPELLI

HPC Education

What defines the 'ideal' HPC graduate? You might be surprised to learn that success hinges more on teamwork than technical skills.

In this interview, Dr. Capelli explores the human side of HPC, emphasising why creativity, communication, and a willingness to try—and fail—are essential for cultivating the next generation of technical talent.

AJ LAUER

Allyship and Coaching

What does it mean to be an ally in HPC? Dr. Lauer shares insights on how those in positions of power can foster a more inclusive supercomputing community.

Drawing on her expertise in EDIA, AJ highlights proven methods to empower teams and drive meaningful change. This discussion includes real-life case studies exploring why women and minorities leave the field—and what the HPC community can do to stem the tide and retain diverse talent.

KNOWLEDGE PORTAL

MARION WEINZIERL AND JEREMY COHEN

Community Volunteering

This interview is by volunteers, for volunteers! Whether you're already contributing to the supercomputing community or looking to get involved, Marion and Jeremy share invaluable insights you won't want to miss.

From identifying where you can make an impact to setting boundaries and organising your own volunteer group, this discussion offers practical advice and inspiration for anyone passionate about giving back.

KAREN LEWIS

Neurodiversity in the Workplace

Neurodiversity acknowledges that people experience and interact with the world in unique ways—there is no single "correct" way to think, learn, or behave.

In this special presentation, Karen Lewis of nAG explores the strengths neurodiverse individuals bring to supercomputing and offers insights on how our community can embrace their abilities to foster greater inclusivity.

https://alces-flight.com/move-the-needle-knowledge-base/

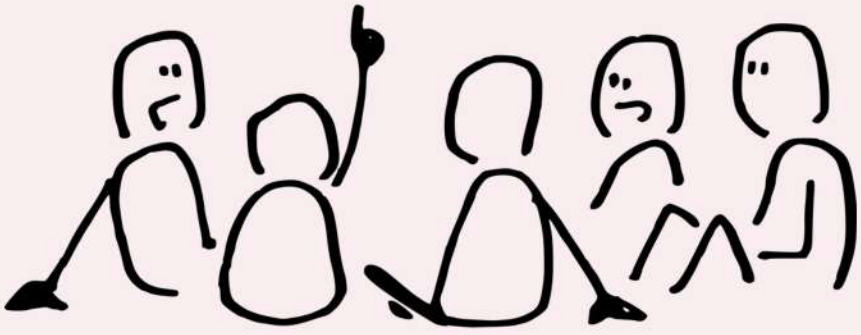
KNOWLEDGE PORTAL

ZARA BIRCH

Community Management

As HPC and AI continue to advance, the role of skilled community managers has become increasingly vital. These professionals act as bridges, fostering relationships, nurturing advocates, and channeling user feedback to drive product improvement and organisational growth.

In this special presentation, Zara Birch of NI-HPC explores the foundations of this role, highlighting its purpose in connecting the technical world and strengthening the broader community.



ONSITE OPPORTUNITIES

We are deeply grateful to the HPC community for offering us the opportunity to create sessions and present what we are learning for the Move the Needle project. Our 2024 sessions included:

UK KTN GREEN COMPUTING WORKSHOP

March 26, 2024 - Manchester Museum of Science and Industry

The workshop highlighted social sustainability, focusing on skills, staffing, engagement and long term impact. Hosted by UKRI, the event underscored the need for technical and soft skills, inclusive staffing, active engagement, and lasting impact to build a resilient HPC ecosystem. Highlights included bridging the skills gap, developing competency frameworks, the need for clear sustainability guidelines, and fostering a collaborative, risk-taking culture.

HPC DAYS - DURHAM

May 8 & 9, 2024 - University of Durham

Over two sessions on separate days, we engaged in discussions about skills, talent, and hiring; presented a draft of our six factors for consideration; and participated in a Lyceum, or open panel discussion. During these sessions, we explored problem-solving strategies and reframing approaches to address current issues in EDIA and HPC.

ONSITE OPPORTUNITIES

SUSTAINABLE REALITY

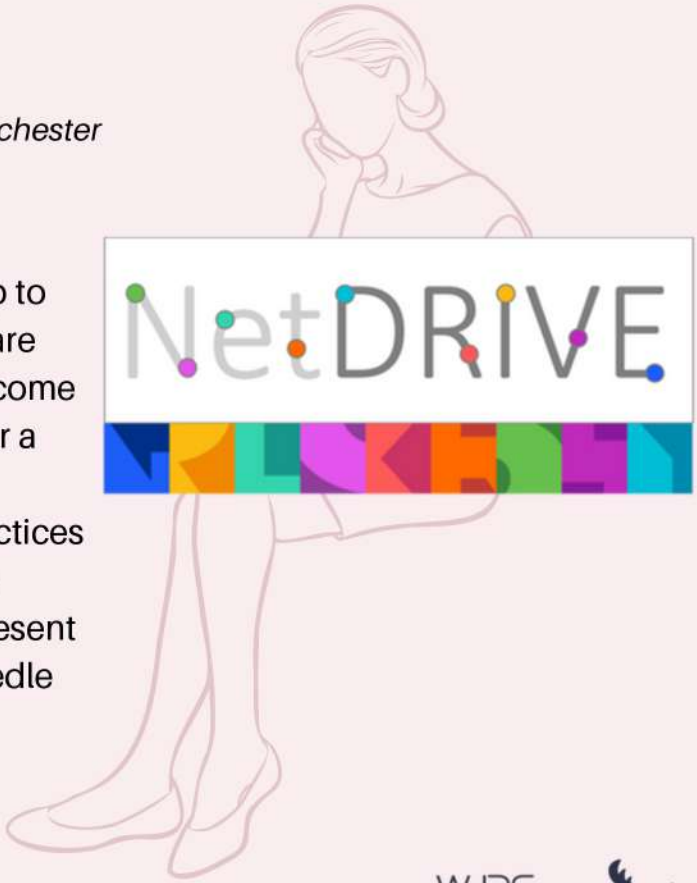
September 25, 2024 - Bletchley Park

This day conference explored the four pillars of HPC/AI: Hardware, Software, People, and Change. We were pleased to welcome two guest speakers for Move the Needle: Karen Lewis of nAG who presented on neurodiversity, and Zara Birch of NI-HPC who presented on community management. We were also pleased to have A-Team Member, Dr. Marta Camps Santasmasas present on her experiences with the project to date.

UKRI-NETDRIVE

October 21, 2024 - Daresbury Labs, Manchester

The Network for sustainable Digital Research Infrastructure Vision and Expertise (NetDRIVE) has been set up to provide a forum for managers, software engineers, academics and others to come together to build a common vision for a sustainable future and to incite a transition to sustainable working practices in the Digital Research Infrastructure communities. We were invited to present our findings to date on Move the Needle as part of their two-day workshop.



ONSITE OPPORTUNITIES

UNIVERSE-HPC

October 24, 2024 - Online

UNIVERSE-HPC or 'Understanding and Nurturing an Integrated Vision for Education in RSE and HPC' defines a training curriculum framework – spanning from undergraduate to continuing professional development level - for Research Software Engineers (RSEs) specialising in high performance computing (HPC). In an invited online talk we outlined our findings to date with the Move the Needle project, focusing on our A-Team's knowledge gained in job role support and retention.

CIUK 2024

December 5 & 6, 2024 - Manchester Central, Manchester

We were honored to host the finale of our Move the Needle project on December 6th during the Women in HPC Breakfast. During this session, three of our A-Team members—Tom Meltzer, Tobias Weinzierl, and Marta Camps Santasmasas—shared their full experiences with the project.

Following their talks, we facilitated an open forum, allowing attendees to engage directly with the speakers and the group through thoughtful questions and discussions.

CONCLUSION

UNEXPECTED RESULTS

When we started the Move the Needle project, our goal was to gather a group of individuals committed to tracking progress in EDIA. What we gained was not only a stellar "A-Team" but also the support of a generous HPC/AI and UK technical community, eager to share their experiences and provide platforms for our ideas.

The outcomes of this project were less clinical and data-driven than anticipated, instead evolving into an exploration of personal experiences and the development of actionable best practices. For those looking to create their own Move the Needle initiative, we hope the resources and foundational goal-setting materials in this report serve as a helpful guide in advancing equity in this field.

CONCLUSION

ACKNOWLEDGEMENTS

This project was made possible through the generous investment of the Alces Flight team who firmly believed that with dedicated exploration we could help improve (and give actionable proof!) to the narrative around how we can increase equity in our field.

Special thanks to Flight Crew Directors Wil Mayers, Steve Norledge, Mark Titorenko, and Michael Rudgyard for their invaluable support.

We also wish to extend a deep appreciation to Women in HPC, who - when proposed the project - joined with zero hesitation. Their willingness to co-host materials and work with their media supporters to get the word out about what was being done was the spark that led to the creation of the Knowledge Portal alongside of the core project itself.

Special thanks to WHPC Board Members George Beckett, Eleanor Broadway, Molly Presley, Lisa Arafune, Elsa Gonsiorowski, Mozhgan Kabiri Chimeh, and Mariann Hardey for their contributions.

To the WHPC Media Supporters who kindly published our initial call and halfway report - insideHPC (Stephanie C. Correra and Doug Black) and HPCwire (Tom Tabor, Lara Kisielewska, Doug Eadline) - we are deeply grateful.

To those who invited us to speak at events: Dawn Geatches, Alan Real, Clare Jenner, Marion Weinzierl, Dr. Martin Juckes, Eirini Zormpa, Martin Robinson, Damian Jones, and Georgina Ellis - you gave us opportunities to both reflect on what we've done and learn more about action being taken in our space.

And finally, to the project team and members of the A-Team who are happy to be publicly named: Tobias Weinzierl, Tom Meltzer, Marta Camps Santasmasas, Deepak Aggarwal, Cristin Merritt, Dominik Wojtak, Katie O'Hare and Louise Mitchell - we could not have done this without you.

THANK YOU

WHPC WOMEN IN HPC x alcesflight

CONTACT INFORMATION FOR FURTHER INQUIRIES

cristin.merritt@alces-flight.com

www.womeninhpc.org

www.alces-flight.com/move-the-needle

CoSeC Annual Conference 2024

As part of the CIUK Day Zero, the Computational Science Centre for Research Communities (CoSeC) ran its fourth annual event around cross-cutting computational science.

All presentations from the day were recorded and these, plus presented material, are available directly on the CoSeC website at www.cosec.stfc.ac.uk

The day saw speakers present work from across the sciences on topics ranging from "ChopMesh - massive parallel mesh generation for CFD/FES simulation on HPCs" (Liang Yang from Cranfield University) to "Data encoding for quantum pangenomics" (Hitham Hassan from Wellcome Sanger Institute) and "Improved viscosity calculations from machine learnt potentials"(Harvey Devereux from Queen Mary University London).

CoSeC's Gemma Poulter (STFC/UKRI) presented her work on the CCP-WSI research catalogue "CAT-WSI" with fellow CoSeC staff member Margaret Duff (STFC/UKRI) presenting on the "core imaging library".

The conference included a session dedicated to the 2024 cohort of CoSeC Fellows who gave a series of lightning talks that introduced the audience to their areas of research and their plans for the fifteen month term of their fellowships.

CoSeC were delighted to see the largest attendance so far at its annual conference and look forward to a further increase next year with the introduction of several new CoSeC communities as a result of recent funding exercises.


COMPUTING
INSIGHT UK 2025
Science and
Technology
Facilities Council
SAVE
THE
DATE
4-5 December 2025
Manchester Central, UK
SEE YOU IN 2025
www.ukri.org/CIUK