Science and Technology Facilities Council

Conference Proceedings
STFC-CONF-2024-001

Computing Insight UK 2023

Manchester Central Convention Centre, UK
7th-8th December 2023

G Lomas, D Jones (editors)

February 2024

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.2024001

ISSN 2753-5800

Computing Insight UK (CIUK) 2023 took place on and 7th and 8th of December 2023 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.

GOLD Sponsors	SILVER Sponsors	BRONZE Sponsors
BOSTON Servers | Storage | Solutions	Stack HPC	ddn
AMD together we advance_	aws	OCF
VANTAGE DATA CENTERS	Quobyte	NAG
VAST	BIOS IT	KAO DATA
	FUJITSU	CSI
	intel	PANASAS
		WEKA
		CORNELIS NETWORKS
		ALTAIR
		Quantum
		HPD DATAVITA
		ATEMPO
		KALRAY THE POWER OF MORE
		SCAN Business

Computing Insight UK 2023 Introduction

Computing Insight UK (CIUK) 2023 was the 34th 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 7-8 December at Manchester Central Convention Complex and attracted a record crowd of over six hundred and fifty attendees.

The theme for this year's conference was "Productive Supercomputing (Accessible and diverse large scale computing)" with sub-themes including "Industry Use of HPC", "Making HPC More Accessible", "User Portals", "Co-Design of Supercomputing Services and Systems" and "Supercomputing into the Future". There was also a session on the "Future of Compute in the UK" that included presentations on the UKRI Digital Research Infrastructure and the Department for Science, Innovation and Technology future of compute review.

CIUK 2023 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 Jessica Gould (UKRI-STFC) for her poster on "Code Coupling Library for Multiphysics CFD and Solid Stress Software for HPC Verification", and we also presented our annual Jacky Pallas Memorial Award, which this year was awarded to Muting Hao (University of Oxford) for her work on "Advancing Aviation Efficiency and Sustainability Through CFD". Muting presented her work as part of the main programme during the conference.

The CIUK 2023 Keynote Presentation was given by Melyssa Fratkin, Industry Programs Director from the Texas Advanced Computing Center. Melyssa's presentation was titled "Women in Advanced Computing: Leveling the Playing Field".

For the first time at a CIUK Conference we introduced a "Day Zero" - a pre-CIUK day that included the CoSeC Annual Conference, the Lustre User Group meeting and two hands-on training courses on "Using ReFrame for reproducible and portable performance benchmarking" and "Run the same code on CPUs, GPUs, and FPGAs with SYCL".

CIUK 2023 also saw the fourth edition of the CIUK Cluster Challenge competition with an incredible thirteen teams entering from University of Bath, Imperial College London, University of Birmingham, University of Strathclyde, Durham University, University of Bristol and Manchester University. The teams completed four online challenges leading up to the conference, followed by four challenges during the conference in Manchester. Team Isambards from the University of Bristol took the title after a closely fought competition and earned their place at the ISC'24 Cluster Challenge competition where they will represent CIUK against the best student teams from around the word.

Computing Insight UK 2023 *"Productive Supercomputing"*

CIUK DAY ZERO - Wednesday 6 December 2023

TIME	PRE-CIUK ACTIVITIES These activities will take place in the CIUK Breakout Room, the Gallery and the Meeting Rooms (upstairs)
From 08:00	**CIUK 2023 EXHIBITION SET-UP (The Gallery)**
09:30 - 10:00	
10:00 - 10:30	CoSeC Computational Science Centre for Research Communities **Annual Conference 2023** Wednesday 6 December @ CIUK 2023 **10:00 - 17:00** **(CIUK Breakout Room)** **Registration from 09:30** **www.scd.stfc.ac.uk/CoSeC**
10:30 - 11:00	code_saturne **Developer Meeting 2023** **Exchange Room 6 and 7 upstairs at Manchester Central** **Attendance By Invitation Only**
11:00 - 11:30	
11:30 - 12:00	
12:00 - 12:30	ReFrame **Using ReFrame for reproducible and portable performance benchmarking** **Exchange Room 2 and 3 upstairs at Manchester Central**
12:30 - 13:00	
13:00 - 13:30	
13:30 - 14:00	
14:00 - 14:30	LUG UK **Lustre User Group 2023** **Exchange Rooms 9 and 10 upstairs at Manchester Central**
14:30 - 15:00	SYCL **Run the same code on CPUs, GPUs, and FPGAs with SYCL** **Exchange Room 2 and 3 upstairs**
15:00 - 15:30	 **CIUK 2023 Student Cluster Challenge (The Gallery)**
15:30 - 16:00	
16:00 - 16:30	LOGICALIS Architects of Change
16:30 - 17:00	
17:00 - 17:30	
17:30 - 20:00	**CIUK 2023 Day Zero Networking Event - The Gas Works Brew Bar, 5 Jack Rosenthal St, Manchester M15 4RA** **(Open to all registered CIUK attendees and exhibitors - wrist band required for entry)**
Until 20:00	**CIUK 2023 EXHIBITION SET-UP (The Gallery)**

Computing Insight UK 2023 *"Productive Supercomputing"*

GOLD Sponsors: BOSTON, AMD, VANTAGE, VAST

SILVER Sponsors: Stack HPC, aws, Quobyte, BIOS IT, FUJITSU, intel

BRONZE Sponsors: ddn, OCF, JJAG, KAO DATA, CSI, PANASAS, WEKA, CORNELIS NETWORKS, ALTAIR, Quantum, HPC DATA, ATEMPO, KALRAY, SCAN Business

Main Programme Session Themes

	Session 1: Making HPC More User Accessible
	Session 2: User Portals
	Session 3: UKRI DRI and DSIT Updates

CIUK DAY ONE - Thursday 7 December 2023

TIME	MAIN PROGRAMME	BREAKOUT SESSIONS	CIUK CLUSTER CHALLENGE
From 08:30	**REGISTRATION OPEN (Charter Foyer) CIUK 2023 EXHIBITION OPEN (The Gallery)**		
09:15 - 09:30	*Welcome and Introduction* **Tom Griffin (Director, Scientific Computing, STFC)**		
09:30 - 10:00	**Dr Tom Deakin (University of Bristol)** *Performance Portability for Next-Generation Heterogeneous Systems*		
10:00 - 10:30	**John Garbutt (StackHPC)** *UKSRC simplify HPC using Azimuth*		
10:30 - 11:00	**Dr Owain Kenway (University College London, Centre for Advanced Research Computing)** *Bridging the Professional Services/Academic gap*	**Project Overview and Update 09:30 - 12:30 (Breakout Room)**	
11:00 - 11:30	**REFRESHMENTS**		**CIUK 2023 Student Cluster Challenge (The Gallery)**
11:30 - 12:00	**Matt Penn (King's College London)** *King's e-Research portal: making things easier, securely*		
12:00 - 12:30	**Dimitrios Bellos (Rosalin Franklin Institute)** *Accelerating Structural Biology: From PC to HPC*		
12:30 - 13:00	**Maria Fando (STFC, CCP4)** *Delivering HPC Power for Structural Biologists with CCP4 Cloud*	**Technical / SysAdmin Meetup 12:30 - 13:30 (Breakout Room)**	
13:00 - 14:00	**LUNCH**		
14:00 - 14:30	**Richard Gunn (UKRI DRI) Speaker TBC (Department for Science, Innovation, and Technology)** *Update on the UKRI Digital Research Infrastructure and the Future of Compute Review* **Followed by a panel discussion and Q&A session**	**Storage Scale User Group (GPFS User Group / Spectrum Scale User Group) 14:00 - 16:00 (Breakout Room)**	
14:30 - 15:00			
15:00 - 15:30			**CIUK 2023 Student Cluster Challenge (The Gallery)**
15:30 - 16:15	**REFRESHMENTS**		
16:15 - 17:00	**Martyn Guest (ARCCA, Cardiff University)** *Performance of Community Codes on Multi-Core Processors*		
17:00 - 18:00	**CIUK 2023 Keynote Presentation - Melyssa Fratkin (Industry Programs Director, Texas Advanced Computing Center, USA)** ***Women in Advanced Computing: Leveling the Playing Field***		
18:30 - 23:00	**CIUK 2023 Networking Event - Revolucion de Cuba, South Central, 11 Peter Street, M2 5QR** **(*CIUK 2023 lanyard and badge required for entry*)**		

Computing Insight UK 2023 *"Productive Supercomputing"*

GOLD Sponsors: BOSTON, AMD, VANTAGE, VAST

SILVER Sponsors: Stack HPC, aws, Quobyte, BIOS IT, FUJITSU, intel.

BRONZE Sponsors: ddn, OCF, IAG, KAO DATA, CSI, PANASAS, WEKA, CORNELIS NETWORKS, ALTAIR, Quantum, HPC DATA, ATEMPO, KALRAY, SCAN Business

Main Programme Session Themes

	Session 4: Industry Use of HPC
	Session 5: Co-Design of Supercomputing Services and Systems
	Session 6: Supercomputing into the future

CIUK DAY TWO - Friday 8 December 2023

TIME	MAIN PROGRAMME	BREAKOUT SESSIONS	CIUK CLUSTER CHALLENGE
From 08:30	**REGISTRATION OPEN (Charter Foyer)** **CIUK 2023 EXHIBITION OPEN (The Gallery)**	WHPC WOMEN IN HIGH PERFORMANCE COMPUTING **Women in HPC Breakfast** **08:30 - 10:30** **(CIUK Breakout Room)**	
09:30 - 10:00	**Ubaid Ali Qadri (STFC Hartree Centre)** *Industrial Use of High-Performance Computing and Artificial Intelligence: A Hartree Perspective*		
10:00 - 10:30	**Ben Rixon and Neil Martin (The Manufacturing Technology Centre)** *Automated Deployment of Manufacturing Use-Cases through OpenStack HPC*		 **CIUK 2023 Student Cluster Challenge (The Gallery)** OCF BRILLIANCE. TOGETHER.
10:30 - 11:00	**Dr Rosemary Francis (Altair)** *Presenting the biggest change to HPC in 20 years*		
11:00 - 11:30	**REFRESHMENTS**		
11:30 - 12:00	**Alastair Basden (DiRAC / Durham University)** *A foray into composable infrastructure for HPC*	**Examples, Challenges and Opportunities for Industrial use of High-Performance Computing and Scalable Artificial Intelligence** **11:30 - 13:30** **(CIUK Breakout Room)**	
12:00 - 12:30	**Dr Tim Bellerby (School of Environmental Sciences, University of Hull)** *The PM Programming Language : Developing Numerical Models on Distributed Systems*		
12:30 - 13:00	**Nick Brown (EPCC at the University of Edinburgh)** *Making HPC more accessible: Effective HPC programming via domain specific abstractions*		
13:00 - 14:15	**LUNCH**		
14:15 - 14:30	**Awards Presentation - The CIUK 2023 Student Cluster Challenge and CIUK 2023 Poster Competition**		
14:30 - 15:00	**The Jacky Pallas Memorial Presentation - Muting Hao (University of Oxford)** ***Advancing Aviation Efficiency and Sustainability Through CFD***		
15:00 - 15:30	**Mark Parsons (EPCC, University of Edinburgh)** *The UK's Exascale Supercomputer* **Dr Sadaf Alam (Bristol University)** *Isambard AI - a National AI Research Infrastructure*		
15:30 - 16:00	**Chris Edsall (Cambridge University)** *AIRR at Cambridge*		
16:00	**CIUK 2023 CLOSES - See you in 2024!!!**		

Welcome to CIUK 2023

“Productive Supercomputing“

Science and Technology Facilities Council

Scientific Computing

CIUK 2023

Welcome to CIUK 2023

"Productive Supercomputing"

Science and Technology Facilities Council
Scientific Computing

- Two Days of Presentations
- Exhibition of the Latest Technology
- Parallel Breakout Sessions
- Poster Competition
- CIUK Cluster Challenge
- Jacky Pallas Memorial Award
- Research Zone
- CIUK 2023 Evening Networking Event

CIUK 2023

Welcome to CIUK 2023

UKRI Science and Technology Facilities Council
Scientific Computing

CIUK 2023

Welcome to CIUK 2023

Science and Technology Facilities Council
Scientific Computing

In the interest of sustainability, we will not be printing a conference book for CIUK 2023.

Instead, we have made the book available to view and download online.

Download your copy here…

CIUK 2023

Welcome to CIUK 2023

Keynote Presentation

Thursday 7 December
17:00 – 18:00

Melyssa Fratkin
(Industry Programs Director, Texas Advanced Computing Center, USA)

Women in Advanced Computing: Leveling the Playing Field

UKRI Science and Technology Facilities Council
Scientific Computing

CIUK 2023

Welcome to CIUK 2023

Science and Technology Facilities Council
Scientific Computing

Jacky Pallas Memorial Award

Friday 8 December
14:30 – 15:00

Muting Hao
(University of Oxford)

Advancing Aviation Efficiency and Sustainability Through CFD

CIUK 2023

Welcome to CIUK 2023

Science and Technology Facilities Council
Scientific Computing

Want to ask a question at the end of a presentation? We are using slido…

slido

Join at
slido.com
#1012 106

CIUK 2023

Welcome to CIUK 2023

Science and Technology Facilities Council
Scientific Computing

archer2 · JASMIN · MMM Hub · cirrus Powered by epcc · GW4 · NI-HPC Northern Ireland High Performance Computing · DiRAC · JADE Tier 2 HPC · ExCALIBUR 10 · hpc sg · N8 CIR Computationally Intensive Research BEDE · CSD3 Cambridge Service for Data Driven Discovery · Bristol · Science and Technology Facilities Council Hartree Centre · iris

CIUK 2023

Welcome to CIUK 2023

Science and Technology Facilities Council
Scientific Computing

Join the conversation…

@CompInsightUK
#CIUK2023

/in/computing-insight-uk
#CIUK2023

CIUK 2023

CIUK 2023 Presentations

Dr Tom Deakin (University of Bristol)

Performance Portability for Next-Generation Heterogeneous Systems

Abstract: There is a huge and growing diversity in the computer processors (CPUs, GPUs, and AI/ML accelerators) used to power the leading HPC systems and supercomputers. Despite their differences in how we as applications developers need to program them, these processors lie on a spectrum of architectural design. With multiple hardware vendors now offering competitive processors, we as a HPC community have more choice, but how do we write our applications to make the most of this opportunity in a productive way? This talk will share how to achieve Practical Performance Portability, and role programming languages, models and abstractions play in the shifting heterogeneous landscape.

Bio: Dr Tom Deakin is a Lecturer in Advanced Computer Systems at the University of Bristol researching the performance portability of massively parallel High Performance simulation codes. He develops both the theory and practice of performance portability, exploring parallel programming languages and designing and evaluating proxy applications. Tom contributes to a number of open standard programming models, including SYCL, OpenMP, OpenCL and ISO C++. He is Chair of the Khronos SYCL Working Group, and a member of the ISO WG21 C++ Standards Committee. Tom's book, with co-author Tim Mattson (Intel Labs), "Programming Your GPU with OpenMP", is now available from MIT Press.

University of BRISTOL

Performance Portability for Next-Generation Heterogeneous Systems

Dr Tom Deakin

Lecturer in Advanced Computer Systems

University of Bristol

Nov'23 Top500 Rank	System	Accelerator
1	Frontier	✅
2	Aurora	✅
3	Eagle	✅
4	Supercomputer Fugaku	❌
5	LUMI	✅
6	Leonardo	✅
7	Summit	✅
8	MareNostrum 5 ACC	✅
9	Eos NVIDIA DGX SuperPOD	✅
10	Sierra	✅

“Complex” cores

Instruction Level Parallelism

Deep cache hierarchy

NUMA

Wide SIMD

In-core accelerators

More “simple” cores

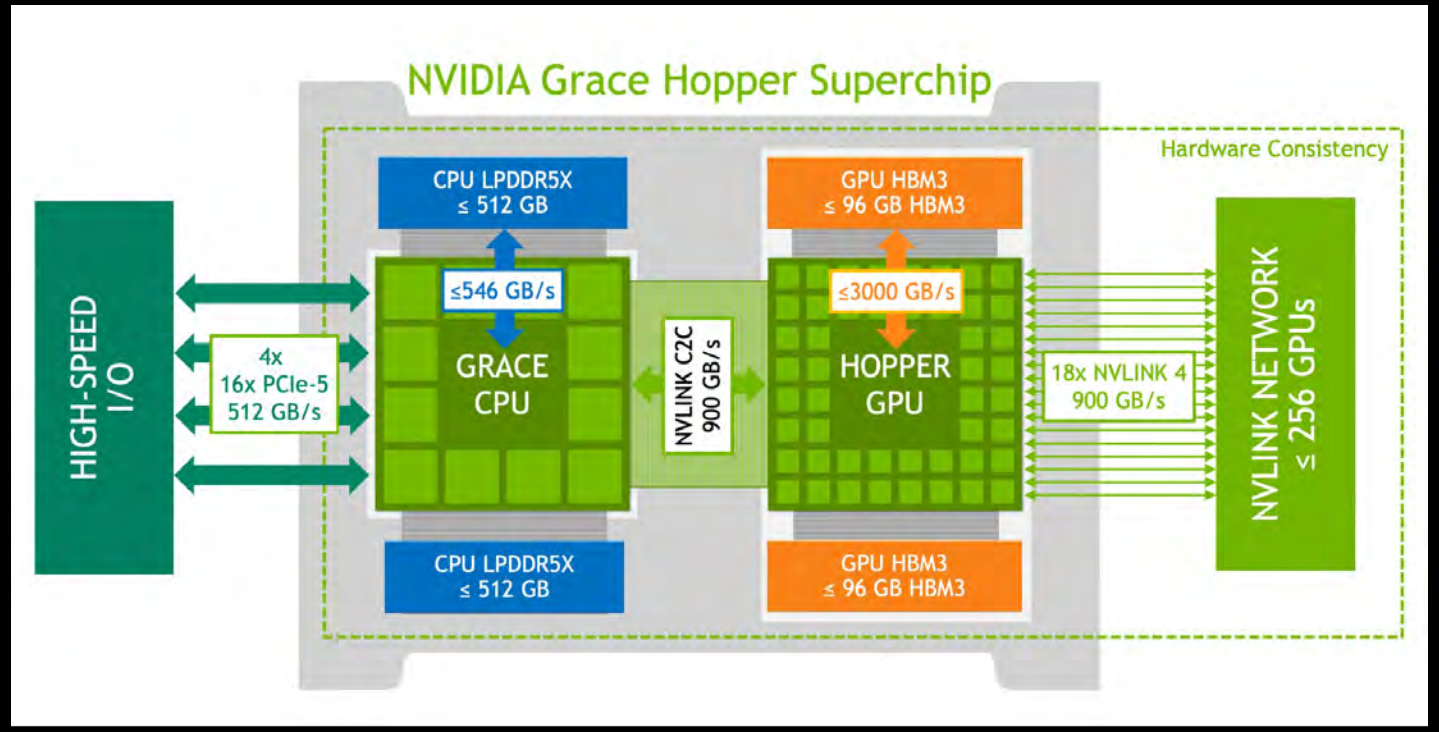
Very wide SIMD

Fast context switching

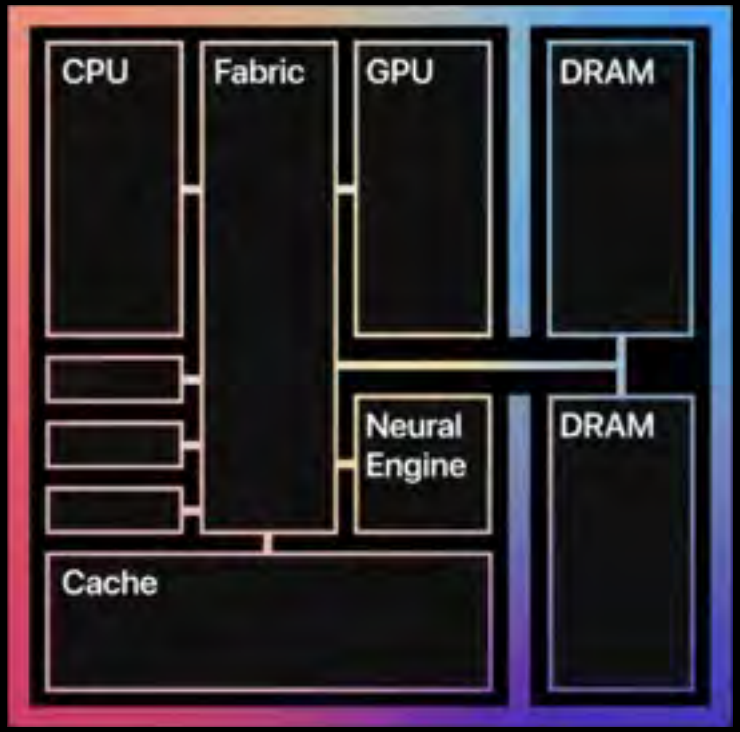
Programable memory hierarchy

Latest memory technology

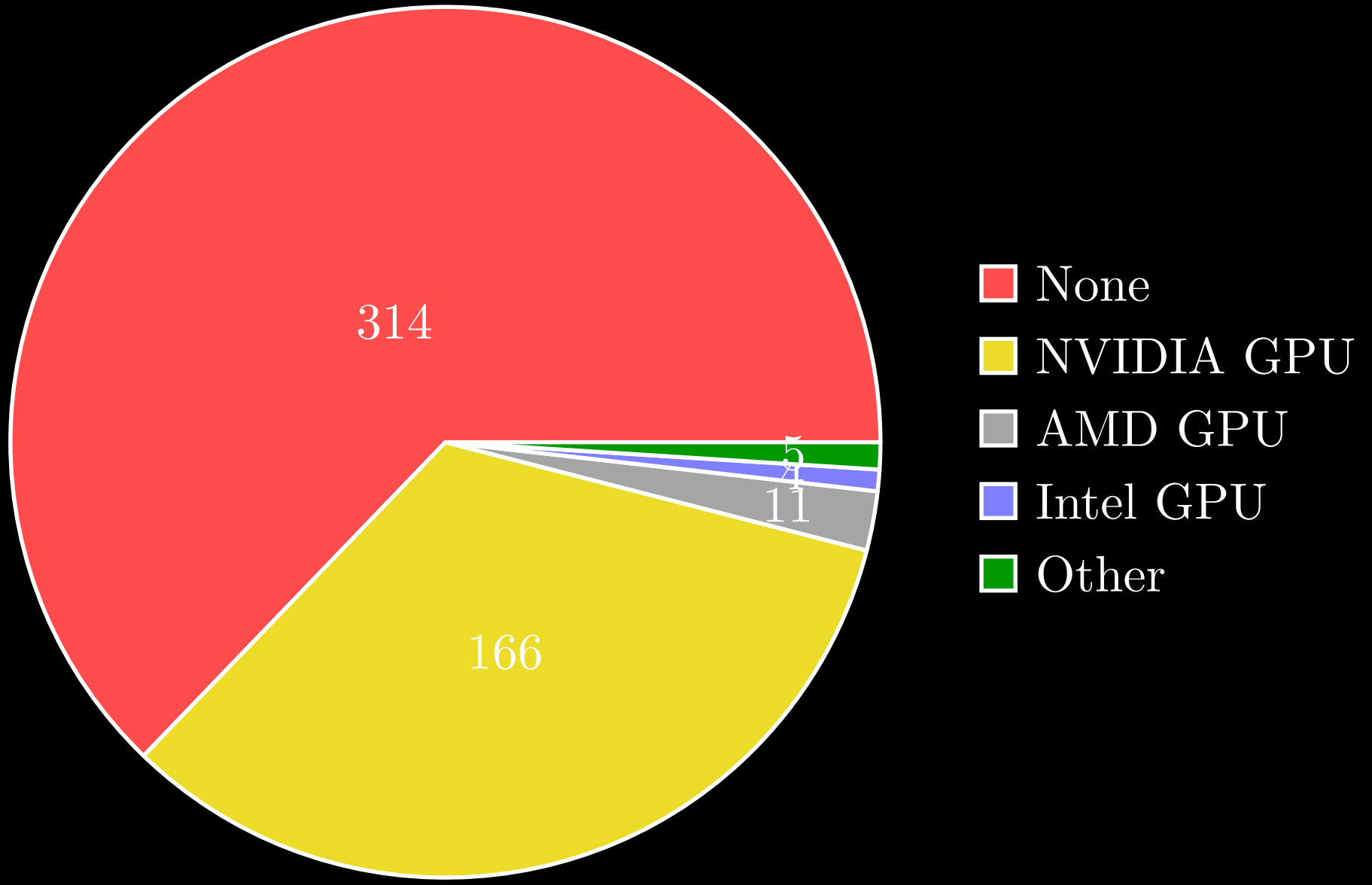
NVIDIA Grace-Hopper



Apple M1



Images belong to their respective owners



Data: TOP500 November 2023
Updated version of chart from: doi.org/10.1109/P3HPC56579.2022.00006

Tension between migrating to next system (which may be GPUs), and keeping running on current system

Performance, Portability, and Productivity

“A code is performance portable if it can achieve a similar fraction of peak hardware performance on a range of different target architectures”.

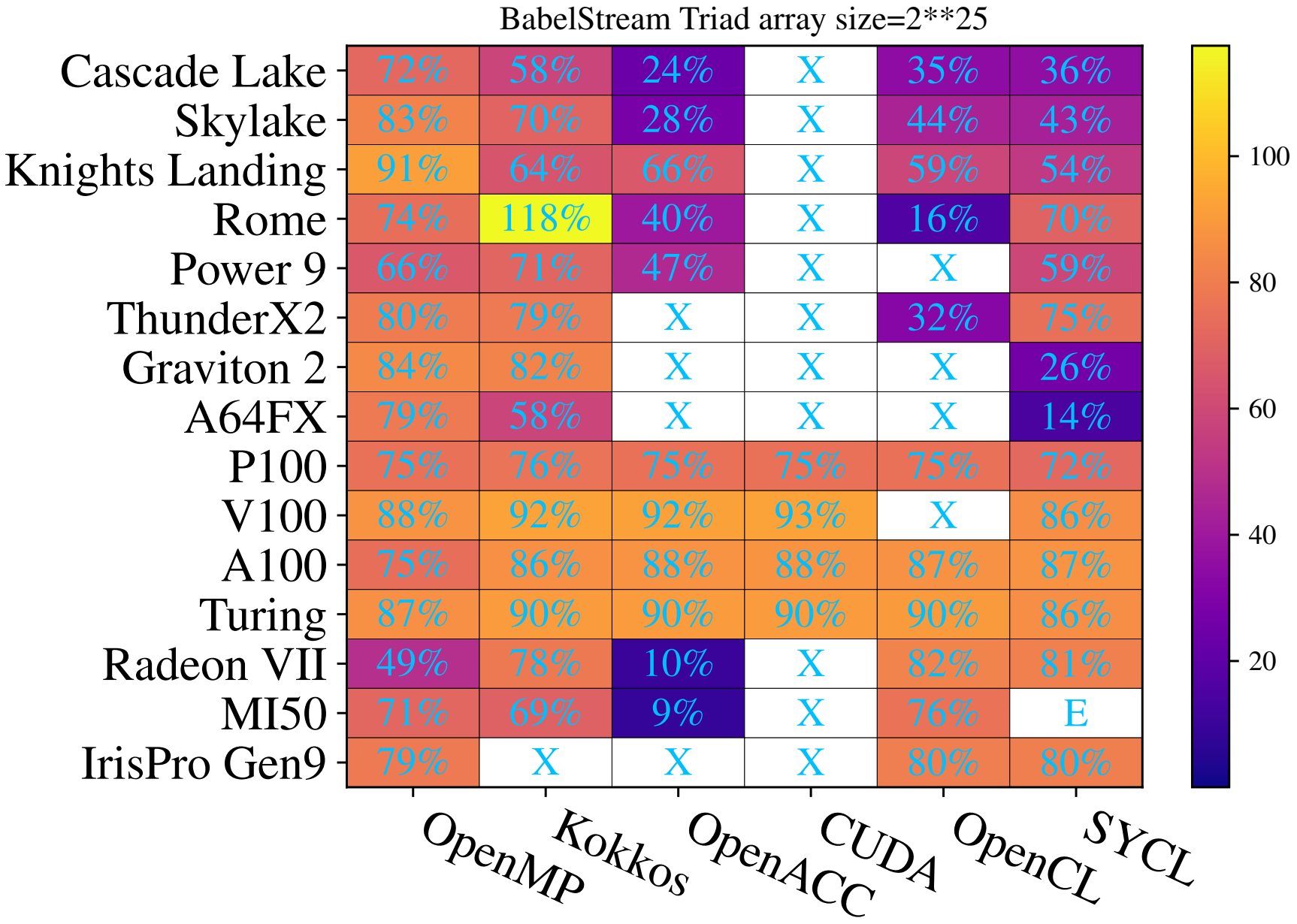
Problem

Application

Platform

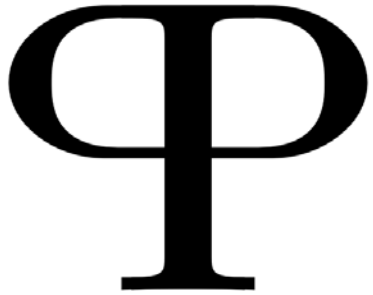
Efficiency

More details in doi.org/10.1109/P3HPC51967.2020.00007



From doi.org/10.1109/P3HPC51967.2020.00006

https://github.com/ukri-excalibur/excalibur-tests

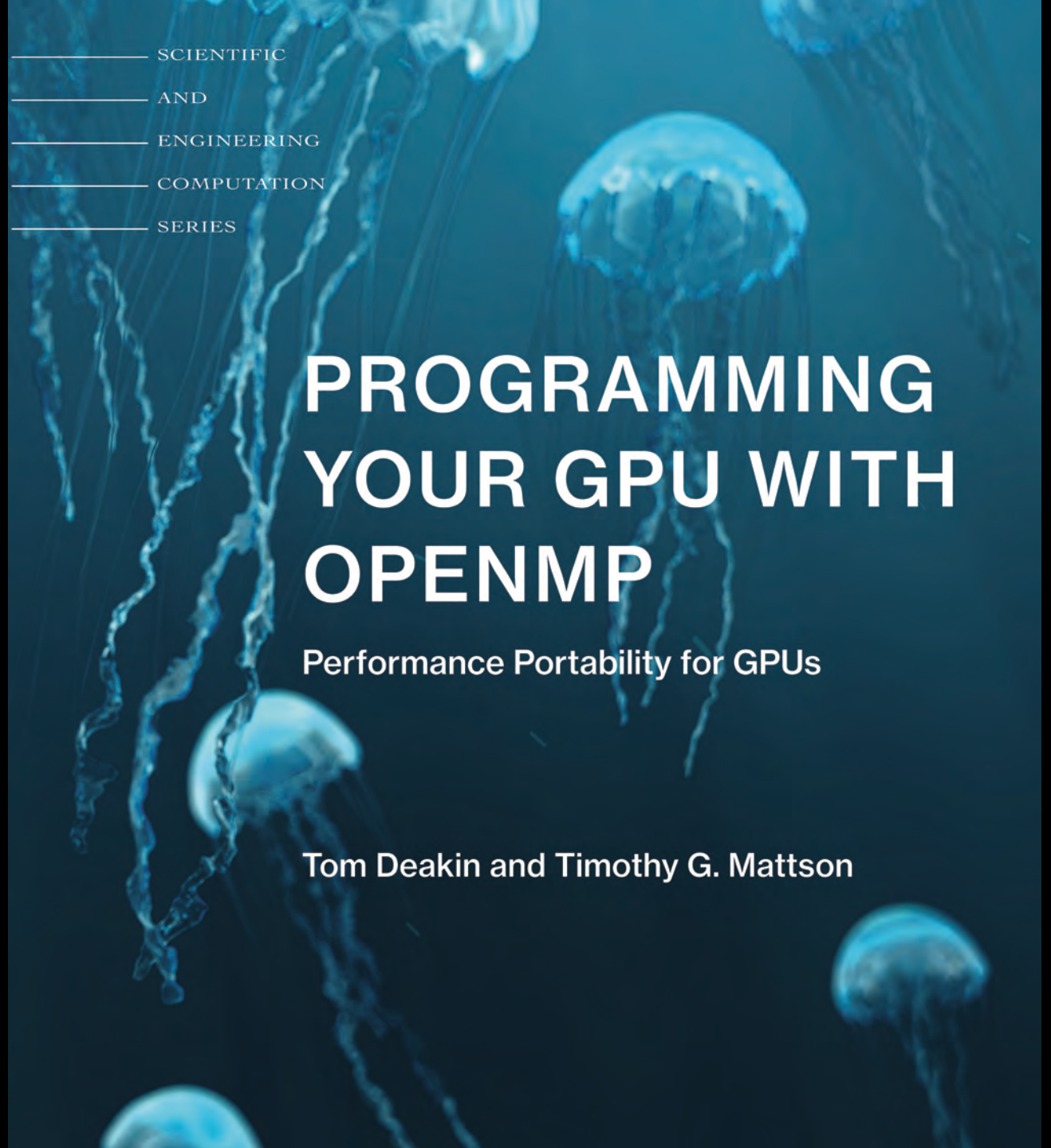


Spack

This work was supported by the Engineering and Physical Sciences Research Council as part of ExCALIBUR Hardware & Enabling Software [EP/X031829/1]

Logos belong to their respective owners

Prompt: A fight between parallel programming languages
Generated with AI on Microsoft Bing Image Creator · 6 December 2023 at 11:58 am


SCIENTIFIC
AND
ENGINEERING
COMPUTATION
SERIES
PROGRAMMING YOUR GPU WITH OPENMP
Performance Portability for GPUs
Tom Deakin and Timothy G. Mattson

Develop with P3 in mind with Standard Parallelism

Use open-standards as confluent off-ramp to be productive today

Express all concurrent work asynchronously

Build in tuning parameters

Test all compilers & runtimes, on all systems

Tell your vendor

CIUK 2023 Presentations

John Garbutt (StackHPC)

UKSRC simplify HPC using Azimuth

Abstract: Azimuth is an open source project that allows users to self-service a best fit platform for science and engineering. Azimuth started as a cloud portal at STFC JASMIN. STFC use cases remain core foci of the development of Azimuth in support of the UK SKA Regional Centre ambitions, IRIS federated cloud and the DiRAC community. Come along to learn how Azimuth can help you simplify deployment of on-demand desktops (aka a bigger laptop!), isolated JupyterHub with external users, Kubernetes clusters with GPUs and high-performance networking, isolated Slurm clusters and more.

Bio: John is a Principal Engineer at StackHPC, working on the convergence of HPC, AI and Cloud. He has been involved with OpenStack since December 2010. He was the OpenStack Nova Project Team Lead for the Liberty and Mitaka releases. Before joining StackHPC in 2017, John worked on OpenStack at Citrix and Rackspace (public cloud). John is currently leading various projects relating to OpenStack, Kubernetes, Large scale baremetal, high performance virtualization, federated clouds and Azimuth. John works from his home near Cambridge, UK.

UKSRC simplify HPC using Azimuth

John Garbutt, StackHPC
(Matt Pryor, StackHPC)
(John Taylor, StackHPC)

StackHPC


uk
SRC
SKAO Regional Centre United Kingdom


zimuth


UNIVERSITY OF
CAMBRIDGE
Research Computing Services

StackHPC

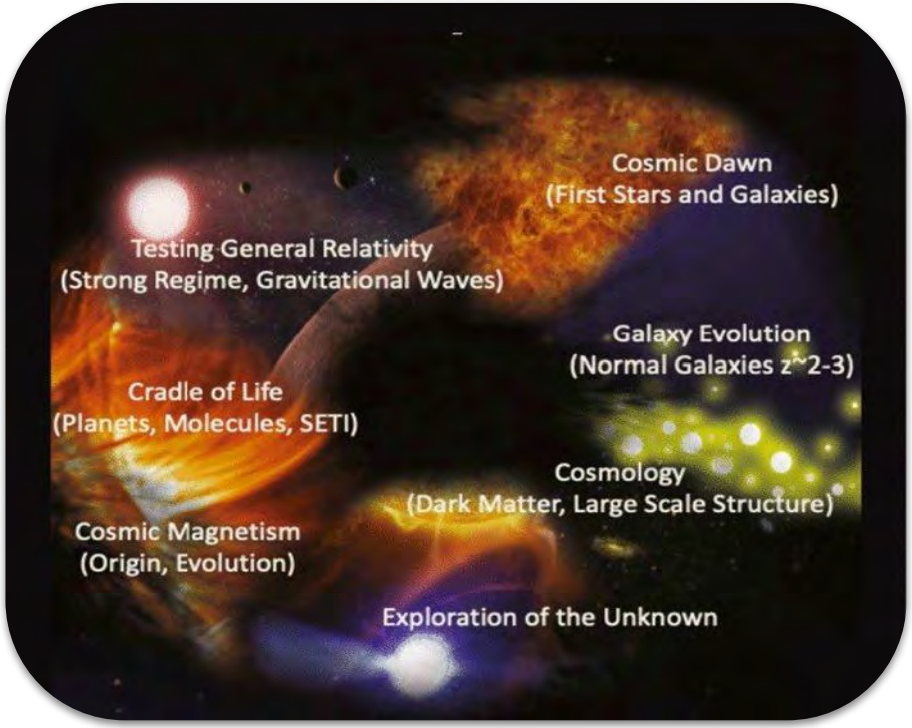
StackHPC

Square Kilometre Array

StackHPC

The Square Kilometre Array (**SKA**) Observatory (**SKAO**) is a next-gen radio astronomy facility covering the frequency range **50 MHz to 15 GHz**.

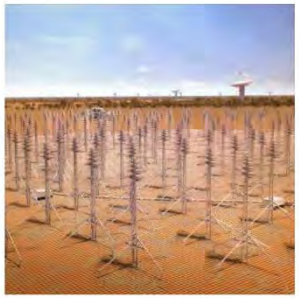
Credit: SKA Observatory



Main Science Drivers. Credit: SKA Observatory

SKA Regional Centre Network

StackHPC



SKAO

130PF Supercomputer

SKA1-Low (Aperture Arrays) → SKA Central Signal Processor ↔ SKA Science Data Processor → 350 PB/year

SKA1-Mid (Dishes) → SKA Central Signal Processor ↔ SKA Science Data Processor → 350 PB/year

Observatory Data Products

SRC Science Analysis Platform Vision Document de Boer, et al (2023)

SRC | Net

SKAO Regional Centre Network

SKA Regional Centre

SKA Regional Centre

SKA Regional Centre

SKA Regional Centre

...

SRC Science Gateway

User

User

User

User

What is 750 PB?

StackHPC

- 100GB = 2hr long 4k movie
- 1PB = 11,000 4k movies

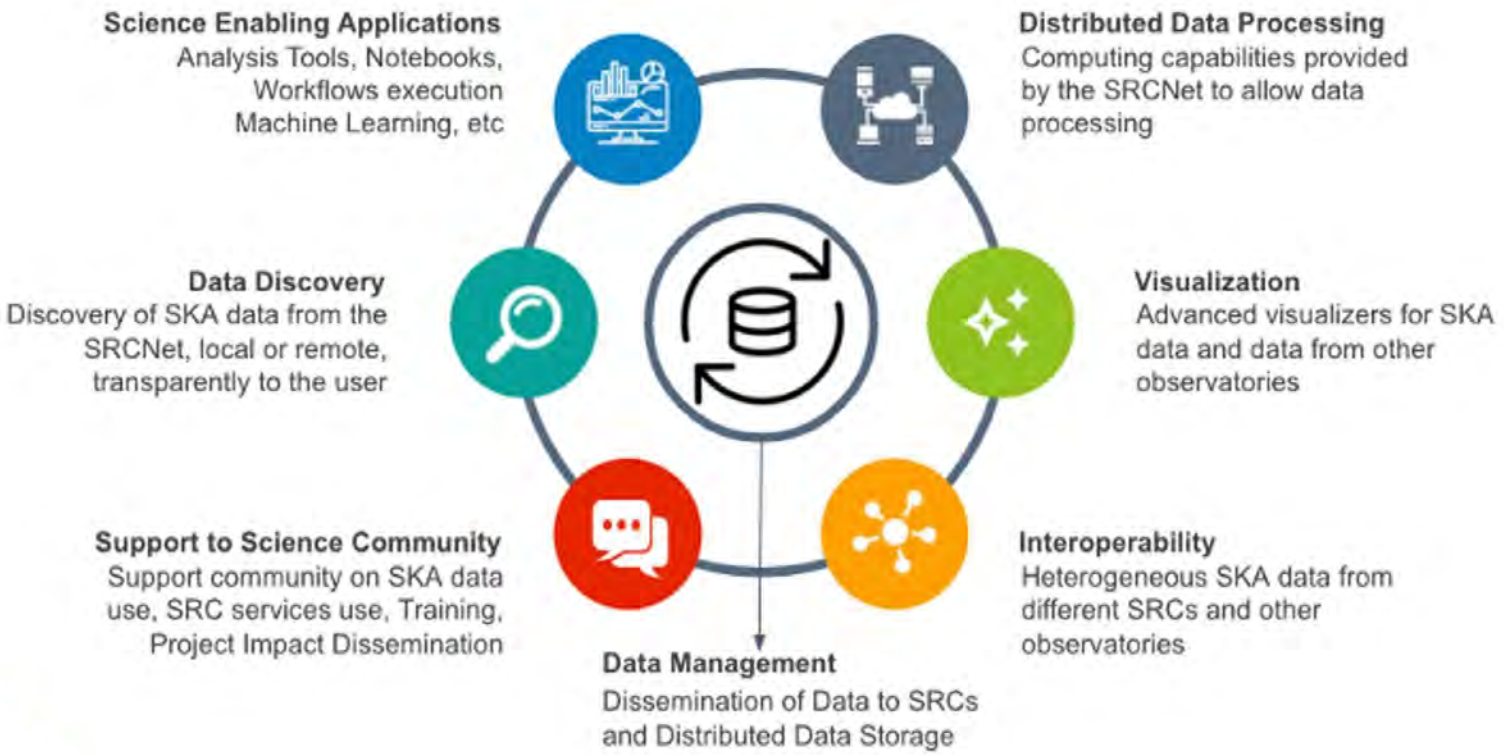
- 200TB = 6 months nonstop 4k movies
- 1PB = 2.5 years nonstop 4k movies
- 750PB = 1750 years nonstop 4k movies

- 1750 years ago Roman Emperor Aurelian died

Credit: G.Dallorto
via Wikipedia

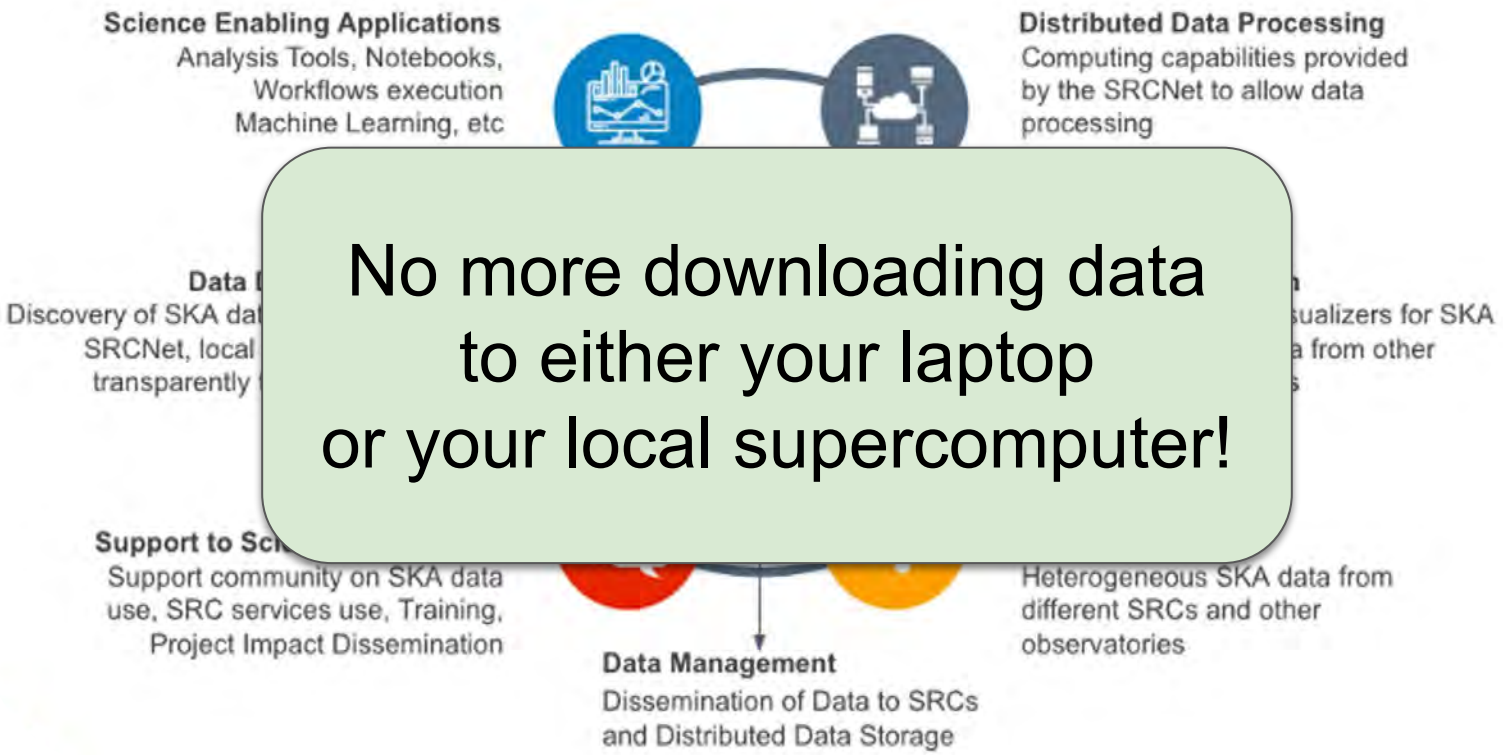
SRCNet and UKSRC

StackHPC



SRCNet and UKSRC

StackHPC




uk
SRC
SKAO Regional Centre United Kingdom


zimuth


UNIVERSITY OF
CAMBRIDGE
Research Computing Services

StackHPC

Azimuth, SKA and StackHPC

StackHPC

SKA SCIENCE DATA PROCESSOR

AlaSKA

2016-2022

Azimuth

2021-

2018-

IRIS Federated e-Infrastructure

iris DiRAC

2023-

UKSRC and SRCNet

uk SRC

SKAO Regional Centre United Kingdom

UKRI Science and Technology Facilities Council

Dawn: UK's Fastest AI Supercomputer

StackHPC

StackHPC

StackHPC Company Overview

StackHPC

- Formed 2016, based in Bristol, UK
  - Based in Bristol with presence in Oxford, Cambridge, France and Poland
  - Currently around 30 people
- Founded on HPC expertise
  - Software Defined Networking
  - Systems Integration
  - OpenStack Development and Operations
- Motivation to transfer this expertise into cloud-native HPC & HPDA (AI)
- “Open” Modus Operandi
  - Upstream development of OpenStack capability
  - Consultancy/Support to end-user organizations in managing HPC service transition
  - Scientific-WG engagement for the Open Infrastructure Foundation
- Hybrid Cloud Enablement

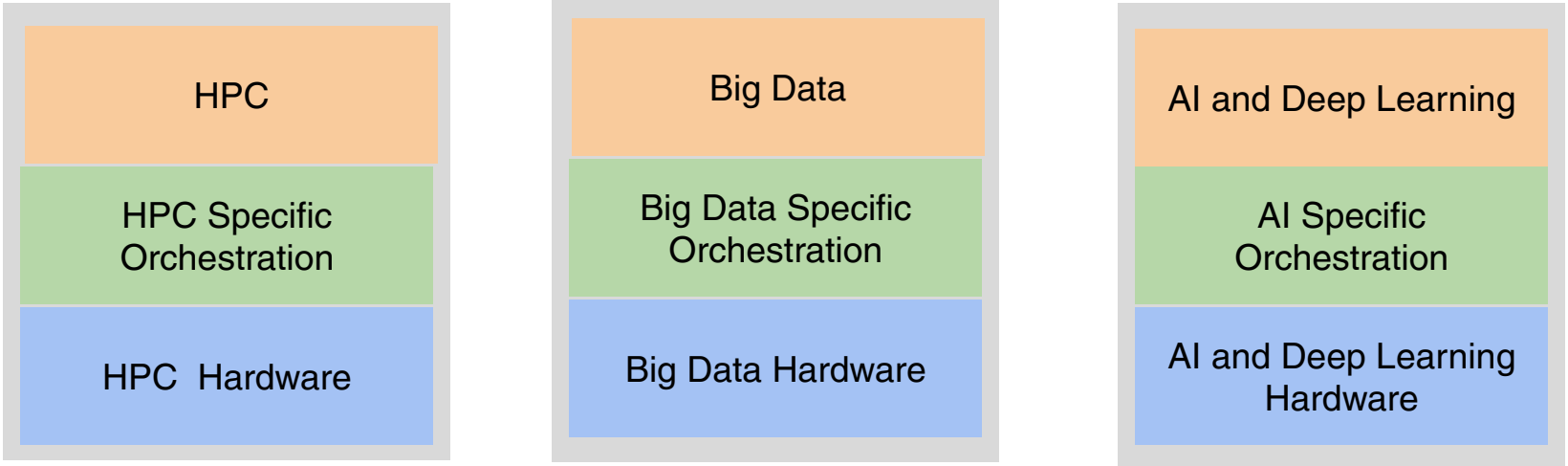
StackHPC

Open Infrastructure FOUNDATION

SILVER MEMBER

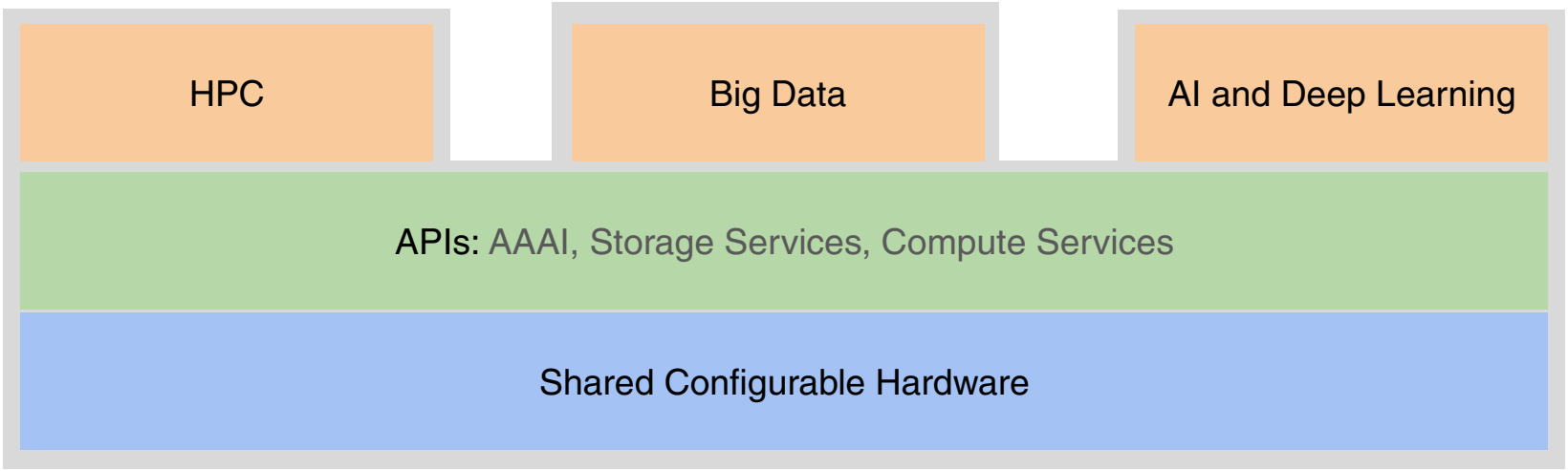
Traditional HPC and AI

StackHPC



Convergence of HPC, AI and Cloud

StackHPC



Dev**O**ps **R**esearch and **A**ssessment

StackHPC

Four key measures of Software Delivery Performance:

- **Lead Time:**
  from customer request to being satisfied
- **Mean Time to Restore (MTTR):**
  failure will happen, get good recovery
- **Change Fail Percentage:**
  a proxy for quality throughout the process
- **Deployment Frequency:**
  a proxy for small batch size

Get Maximum Value from your Investments

https://unsplash.com/photos/2FaCKyEEtis

Get Maximum Value from your Investments

Self-Service HPC and AI

StackHPC

Reconfigurable and Isolated Infrastructure

Performance to extract maximum value

Azimuth Self-Service Reproducible Apps

LOKI: The OpenInfra Standard

StackHPC

More than 85% of OpenStack deployments include Kubernetes
(OpenStack 2022 User Survey)

- **Linux**
  Open Operating System standard
- **OpenStack**
  Open Cloud standard
- **Kubernetes**
  Open Container Orchestration standard
- **Infrastructure**

Azimuth Co-Development

StackHPC

JASMIN

uk | SRC

SKAO Regional Centre United Kingdom

GRAPHCORE

DiRAC

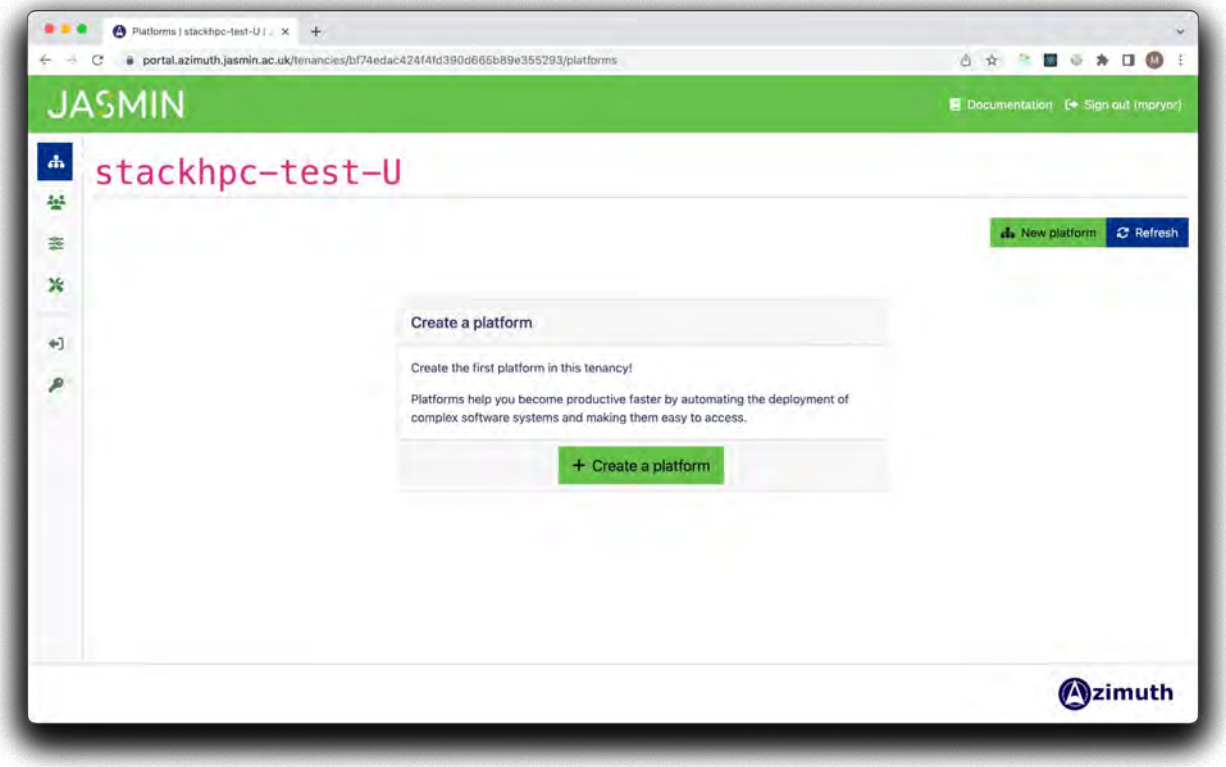
Azimuth Self-Service Applications

StackHPC

Add your own apps here!

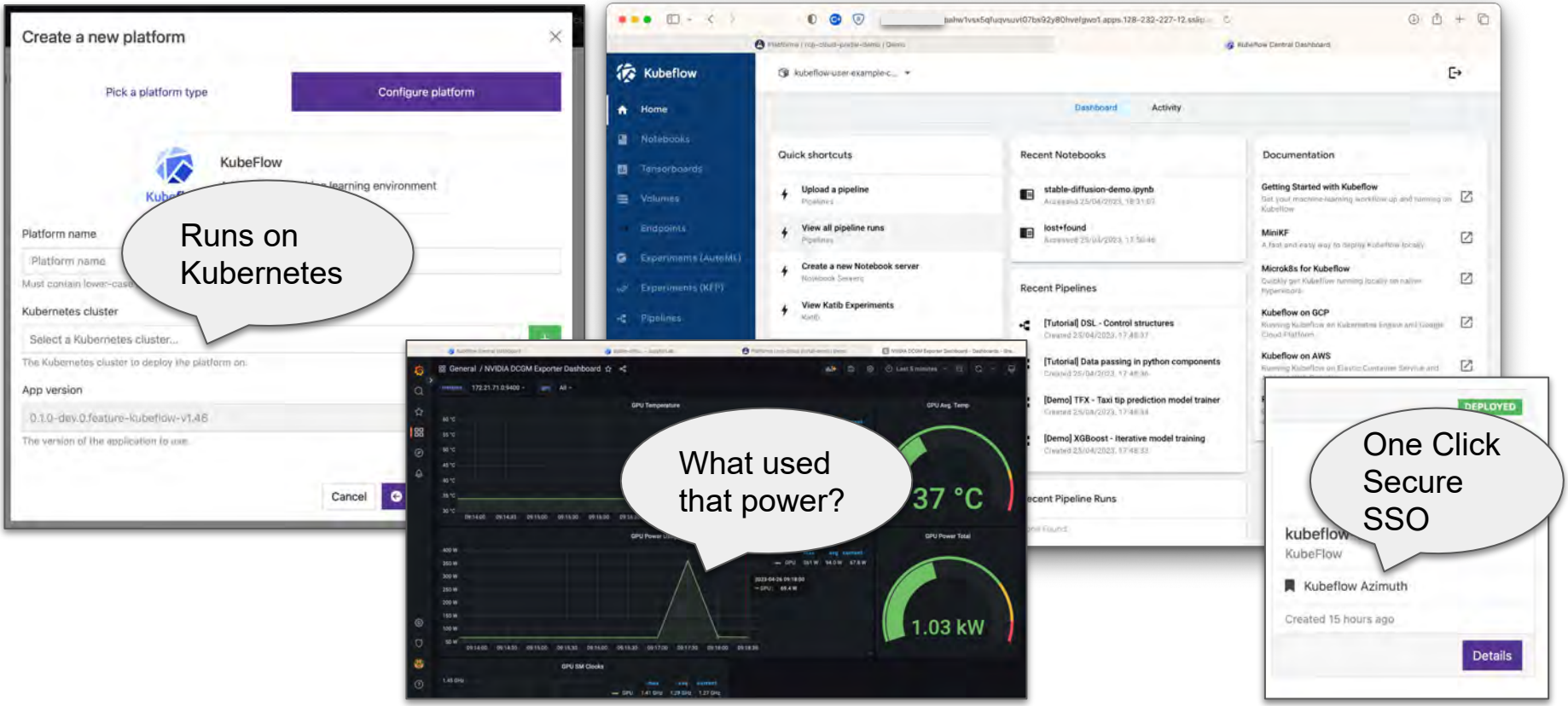
Bigger Laptop via Guacamole

StackHPC



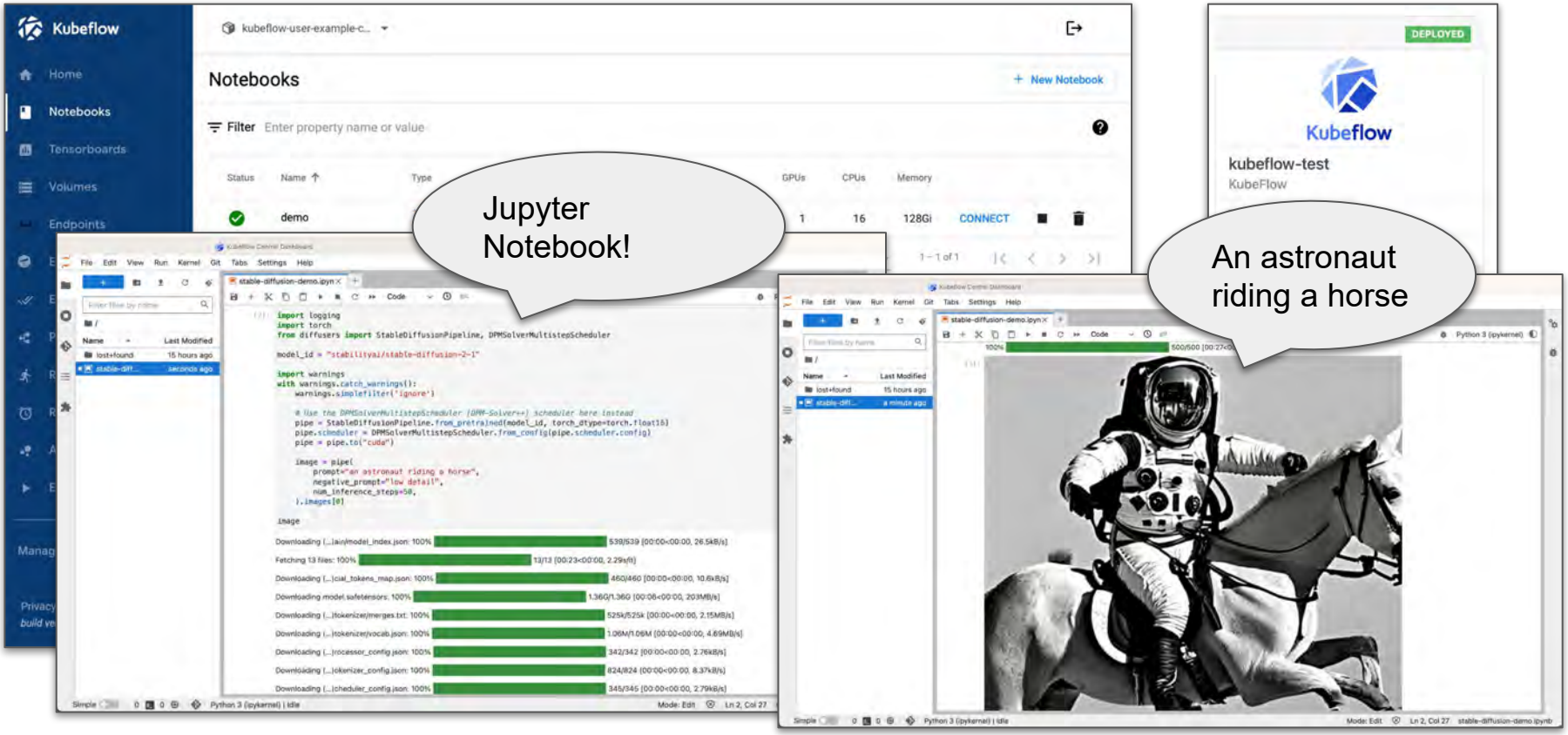
GPU enabled Kubeflow

StackHPC



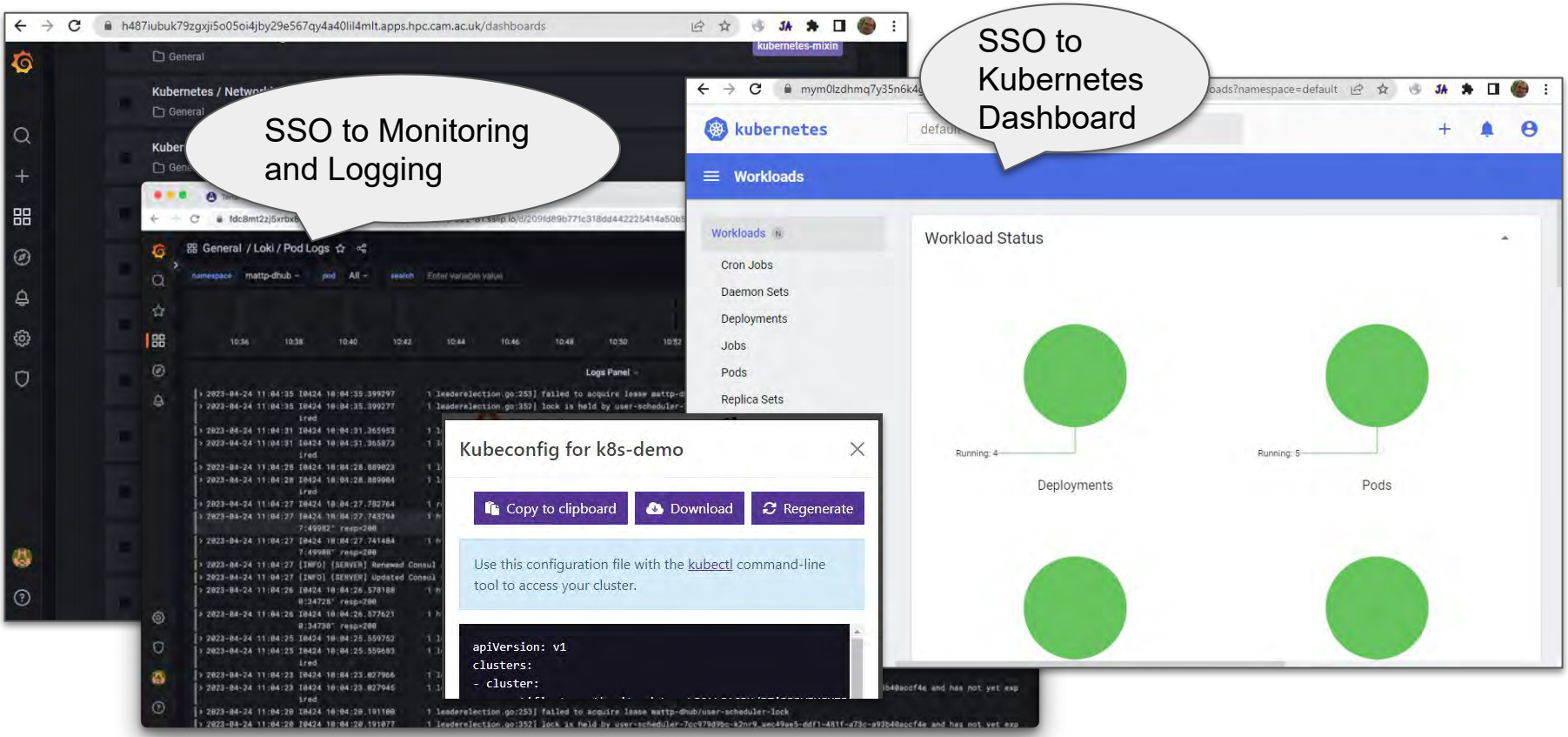
GPU enabled Kubeflow

StackHPC



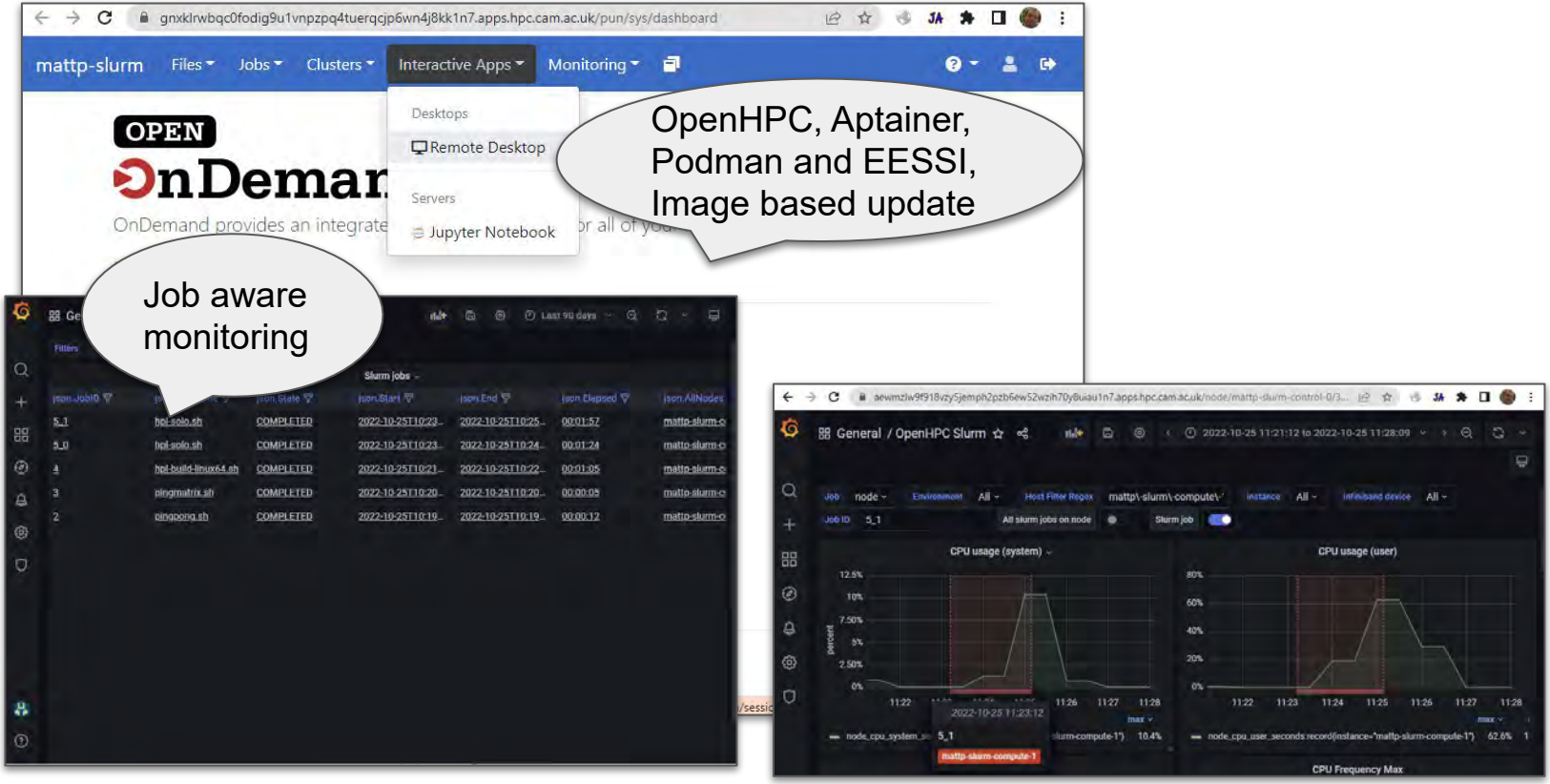
Kubernetes using Cluster API

StackHPC



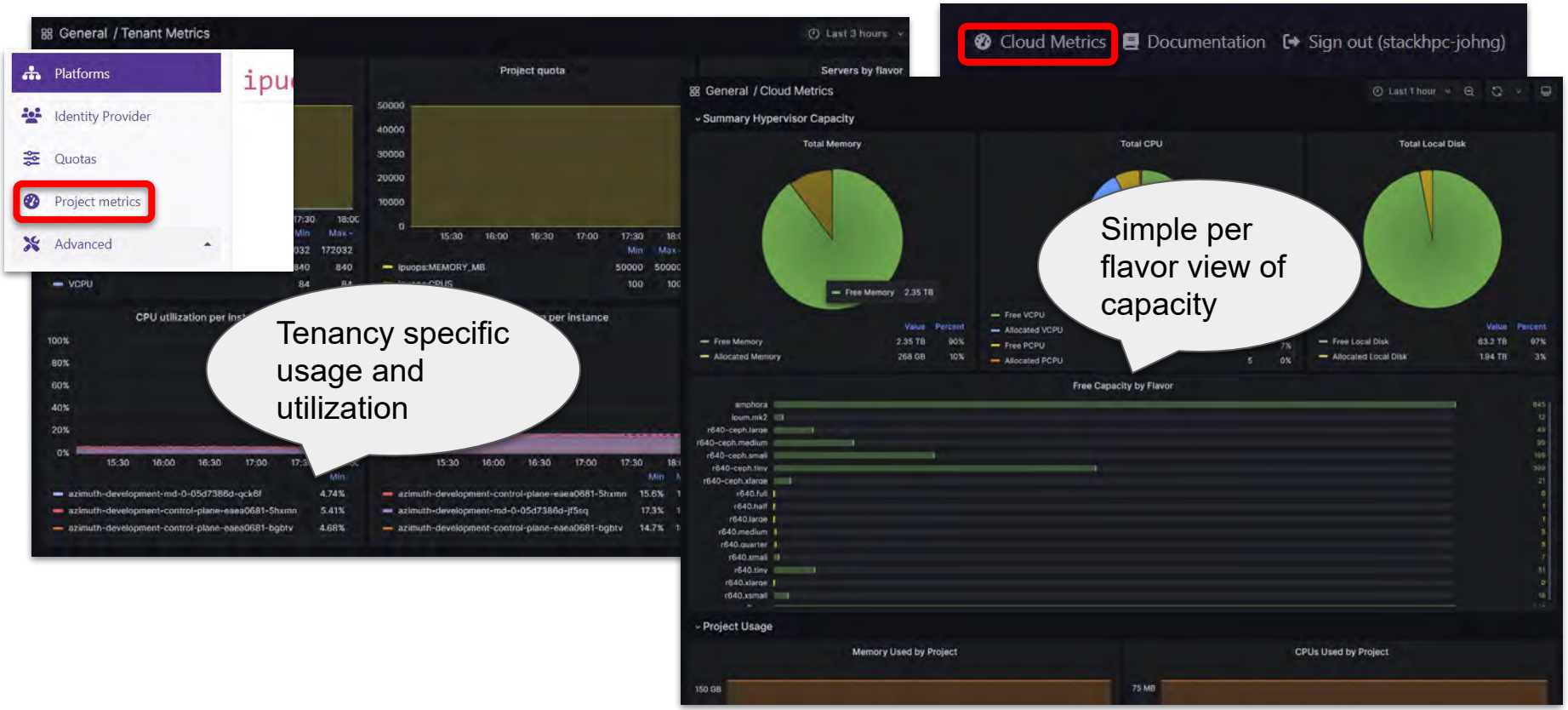
Slurm with Open OnDemand

StackHPC



Cloud Capacity and Utilization

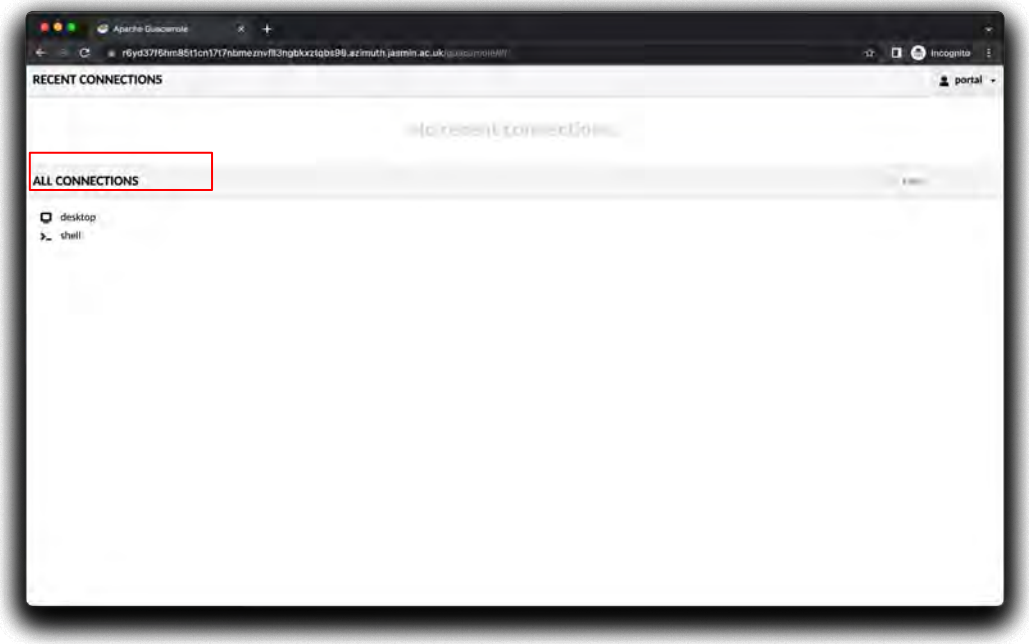
StackHPC



External Application Users

StackHPC

- Azimuth user is Application Admin with access controlled by cloud operator, typically via AAI proxy
- **Application admins** can grant access to **application users** that do not have an Azimuth account
- Each cloud tenancy gets an independent identity provider
  - Manage users and group memberships
  - Application users can be local or federated from another IdP
  - Uses Keycloak realms




uk
SRC
SKAO Regional Centre United Kingdom


zimuth

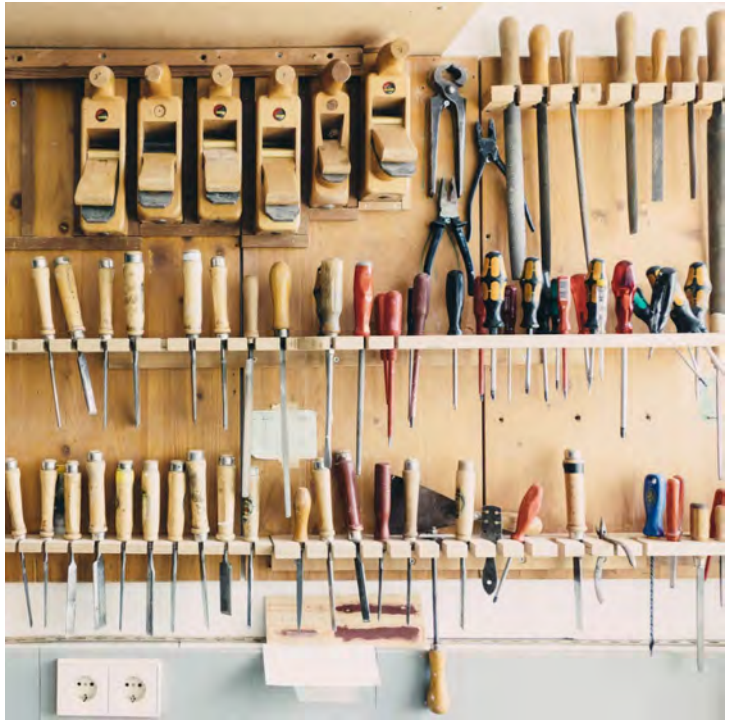

UNIVERSITY OF
CAMBRIDGE
Research Computing Services

StackHPC

How can I do this?

StackHPC

- Fully Open Source solution
- StackHPC support available
  - For Azimuth, OpenStack and Ceph
- Azimuth workshops available
  - Builds on OpenStack and Ceph workshops
  - Session 1: Users meet Azimuth
  - Session 2: Operators meet Azimuth
  - Session 3: Design your azimuth-config
- Co-development opportunities available
  - Improved storage integrations, Federated batch
  - Resource reservations, Hybrid cloud



https://stackhpc.github.io/azimuth-config/try/

Thank you!

https://www.stackhpc.com
https://www.uksrc.org
https://www.hpc.cam.ac.uk

StackHPC

CIUK 2023 Presentations

Dr Owain Kenway (University College London - Centre for Advanced Research Computing)

Bridging the Professional Services/Academic gap

Abstract: UCL ARC (Centre for Advanced Research Computing) has been building on the success of its predecessor, "Research IT Services" in having one of the first RSE teams in the UK by both building out – building on the concept of the RSE as a “research IT staff scientist” to include roles for Research Infrastructure, Data Science, Data Stewards and Research Managers – and by building up – harmonising conditions between professional services and academic strands to build a “best of both” for staff working there. ARC staff have standardised job descriptions at all grades, a promotion round every year (like academia), have permanent positions and run services (like professional services) and so on. In this talk as Head of Profession Dr Kenway will present these structures and the lessons learned from implementing them and running them for almost two years, along with a call to action for other institutions to follow suit.

Bio: Dr Kenway is the Head of Research Computing in the Centre for Advanced Research Computing at UCL. In this role he runs the group responsible for all UCL’s centrally provided HPC and Research Data Storage infrastructure as well as services for private Cloud and AI. He has worked at UCL designing and supporting these services for well over a decade and helped design the change from being a professional services department to a unique (at least at UCL) hybrid of academic and professional services and is “Head of Profession” for the Research Infrastructure Developers at 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

About ARC

- ARC is new hybrid academic/professional services department spun out of ISD (central IT) at UCL – a little over 100 staff total and growing.
- Dual reporting at Director level (Professional Services and Academic).
- One PS director and three academic associate directors.
- Responsible for “Research IT” services provided centrally at UCL.
- Research “staff scientists”.
- Researchers (inc. PI on grants).
- Teaching.

The challenges

- UCL previously had no concept of a department in PS that does research:
  - Do we give or take overheads?
  - How do we bend finance/project process/HR/procurement to our will so that we can be effective researchers?
  - How do we give staff meaningful progression so that we retain staff if they are trapped in the PS “you are hired into a role and never progress” HR process?
  - Can we stop people being turned into managers to go up grades?
- How do we balance service provision with research/teaching work?
  - Still challenging but we’re working on it!
- Will we ever have a part of a building which is ARC?

The opportunities

- We can design how Research IT professions work!
- Lots more options for satisfying work:
  - Teaching
  - Research
- Direct experience in teaching and research can inform the design of systems.
- Central IT/Finance's new rules mean we can hire enough people to staff the department!

The core problem

Junior Research IT Professional:

- Handles support board
- Does some basic software installs
- Helps install hardware
- Learning on the job
- Primary editor: Notepad++, Nano, VS Code(?)
- Writes code: Python

This job is great, much better than doing a postdoc!

The core problem

Research IT Professional:

- 2nd/3rd Line
- Expert at getting difficult software to work.
- Designs hardware deployments.
- Researches complex problems and invents solutions
- Primary editor: Vim or Emacs
- Writes code: Fortran/C++/Cuda…

Oh god please don't promote me – I want to keep doing technical work

The core problem

Senior Research IT Professional:

- Attends governance meetings
- “Negotiates” with vendors and other “senior stakeholders”
- Line management
- Has vast wealth of experience that is rusting through lack of use
- Primary editor: **Microsoft Word**
- ~~Writes code~~: **Microsoft Excel**

My life is pain, and I don’t know how I got here.

The core problem

- The vast, overwhelming majority of people who want to do “Research IT” want to do (and are good at) technical work and don’t want to become “mediocre managers” to progress. *(I may be projecting here a bit)*

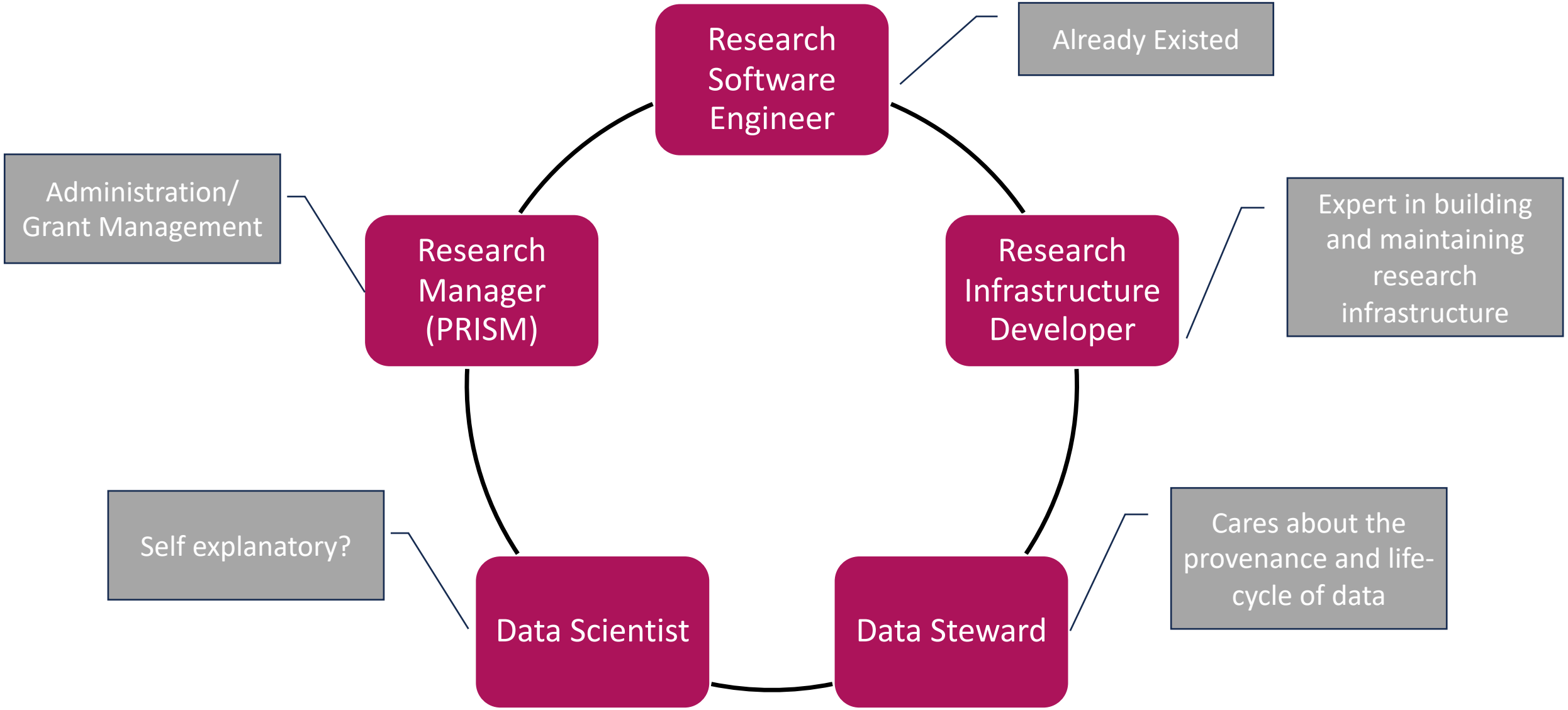
BUT

- Professional services structures **ONLY** reward becoming a manager.

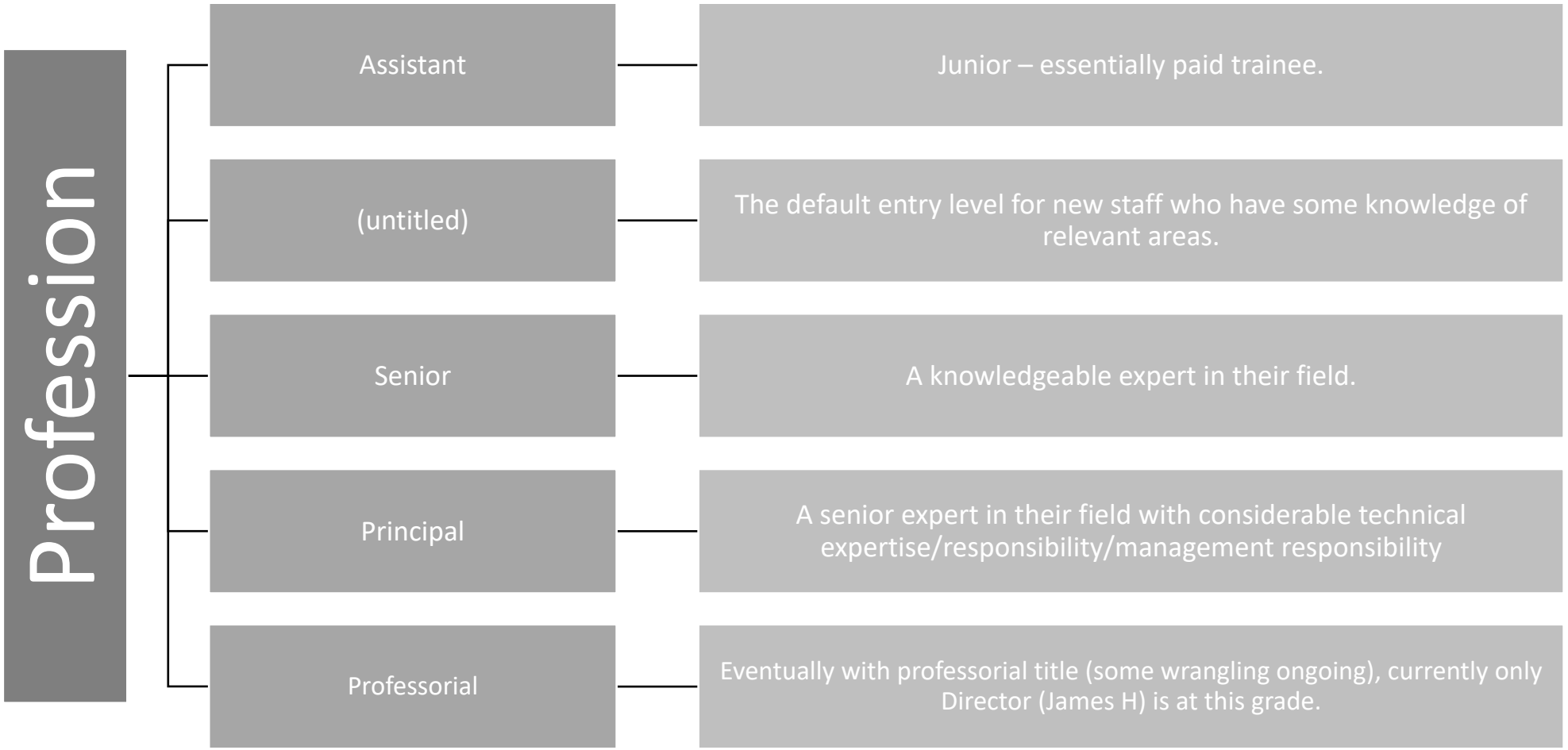
Goal

We need to design a professional structure that gives staff seniority due to technical expertise/leadership.

The Professions



Grade Structure (6 → 10)



"I'm a Principal Research Infrastructure Developer. I am Head of Research Computing and Head of Profession for the Research Infrastructure Developers."

Duties by grade

(note for space reasons these are very summarised as each area is worth a slide in itself!)

- Senior:
  - Technical leadership – design, or service ownership, or delivering technical projects.
  - Teaching/training leadership
  - Line management/MSc supervisor/PhD second supervisor.
- Principal:
  - Strategic leadership – responsible for long term planning for a department, or for a major initiative/external funded project/University Strategy/PI etc.
  - Head of profession.
  - Technical expertise and leadership – e.g. owning a key technical architectural area.
  - “Product” Ownership
- Professorial:
  - National or international leadership.
  - Several or more of the Principal roles.
  - Senior leadership at UCL.

Promotion

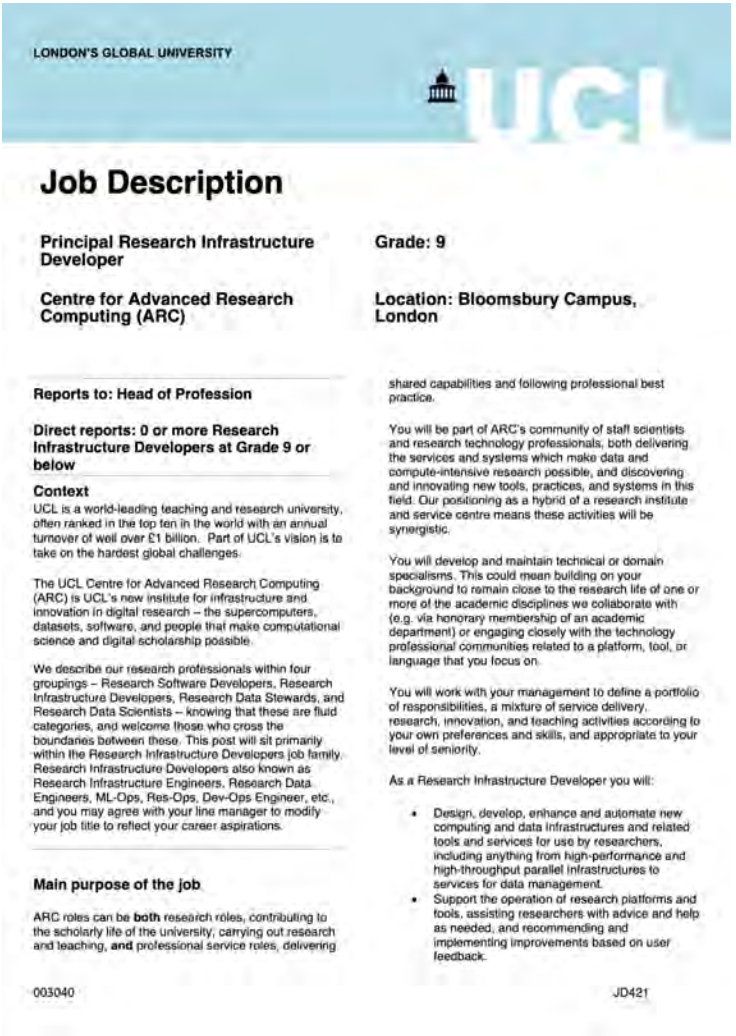
- ~~Academics get promoted! PS staff don't!~~
- Not anymore!
  - Yearly promotions round in the summer:
    - Standardised case format with internal/external references.
    - With or without support of line manager.
    - Much lighter weight than a job application.
    - Promoted "in post" i.e. your post goes up a grade with you in it!
    - Can fall back to a spinal point increase.

- Not just ARC! Central IT are doing this too!
- In 2022 we did our first promotion round and promoted 8 people.
- In 2023 we promoted 7 people.

Rolling recruitment

- Rather than having specific posts people are recruited into, we have rolling advertisements for all professions and grades and ”good people are appointed at the appropriate grade and developed”.
  - Don’t miss people because jobs are over/under-graded.
  - Don’t miss people because they apply to the wrong profession.
  - Direct application through LinkedIn is possible (no more ROME).

Actually a bit of a fudge underneath to keep HR + finance + the law happy.



LONDON'S GLOBAL UNIVERSITY

UCL

Job Description

Principal Research Infrastructure Developer

Grade: 9

Centre for Advanced Research Computing (ARC)

Location: Bloomsbury Campus, London

Reports to: Head of Profession

Direct reports: 0 or more Research Infrastructure Developers at Grade 9 or below

Context

UCL is a world-leading teaching and research university, often ranked in the top ten in the world with an annual turnover of well over £1 billion. Part of UCL's vision is to take on the hardest global challenges.

The UCL Centre for Advanced Research Computing (ARC) is UCL's new institute for infrastructure and innovation in digital research – the supercomputers, datasets, software, and people that make computational science and digital scholarship possible.

We describe our research professionals within four groupings – Research Software Developers, Research Infrastructure Developers, Research Data Stewards, and Research Data Scientists – knowing that these are fluid categories, and welcome those who cross the boundaries between these. This post will sit primarily within the Research Infrastructure Developers job family. Research Infrastructure Developers also known as Research Infrastructure Engineers, Research Data Engineers, ML-Ops, Res-Ops, Dev-Ops Engineer, etc., and you may agree with your line manager to modify your job title to reflect your career aspirations.

Main purpose of the job

ARC roles can be **both** research roles, contributing to the scholarly life of the university, carrying out research and teaching, **and** professional service roles, delivering shared capabilities and following professional best practice.

You will be part of ARC's community of staff scientists and research technology professionals, both delivering the services and systems which make data and compute-intensive research possible, and discovering and innovating new tools, practices, and systems in this field. Our positioning as a hybrid of a research institute and service centre means these activities will be synergistic.

You will develop and maintain technical or domain specialisms. This could mean building on your background to remain close to the research life of one or more of the academic disciplines we collaborate with (e.g. via honorary membership of an academic department) or engaging closely with the technology professional communities related to a platform, tool, or language that you focus on.

You will work with your management to define a portfolio of responsibilities, a mixture of service delivery, research, innovation, and teaching activities according to your own preferences and skills, and appropriate to your level of seniority.

As a Research Infrastructure Developer you will:

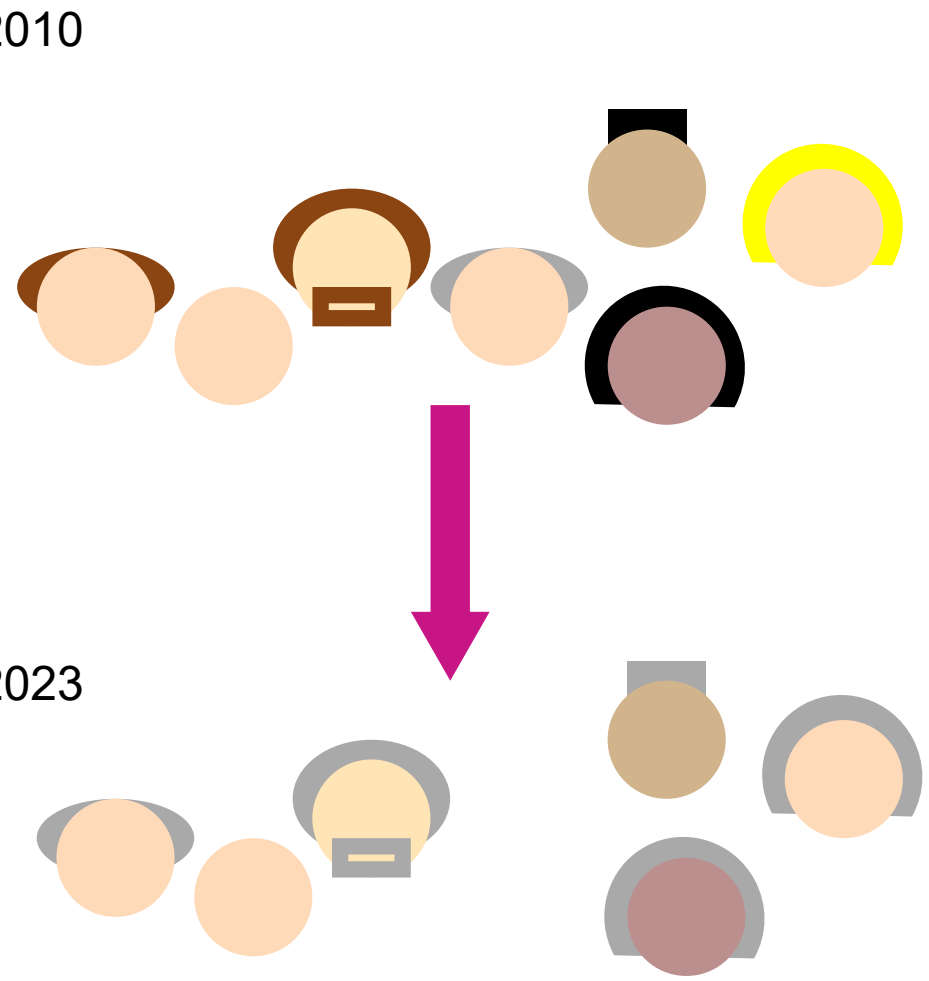
- Design, develop, enhance and automate new computing and data infrastructures and related tools and services for use by researchers, including anything from high-performance and high-throughput parallel infrastructures to services for data management.
- Support the operation of research platforms and tools, assisting researchers with advice and help as needed, and recommending and implementing improvements based on user feedback.

003040

JD421

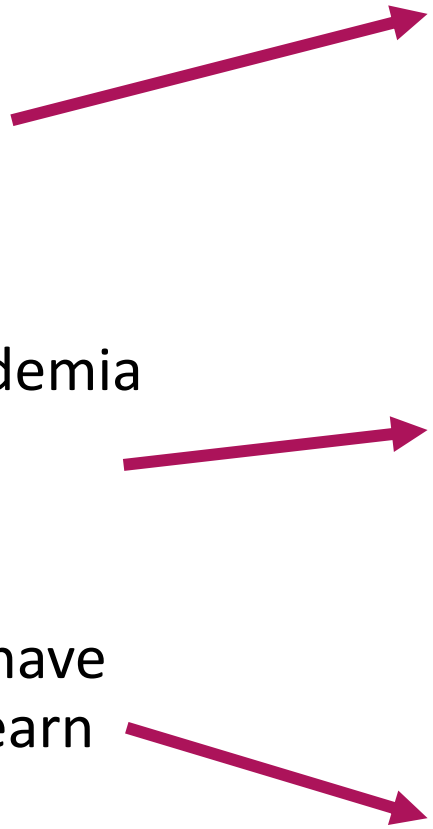
Creating future employees

- HPC community is small and is not growing.
  - Same people flit about between universities and vendors.
  - Cloud companies make things worse.
  - All getting older and greyer as we approach retirement.
  - No new blood.



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



Teaching course in “Data Engineering” with CS

Two apprentices started in September 😀😀

Grade 6 “entry” level posts with learning on the job

Lessons learned

- Changing the behaviour of the wider organisation is hard:
  - “Unorganised passive resistance” – other PS depts agree to things in principal but do not actually change behaviour:
    - Some things are working well after hard battles (HR/Recruitment, Finance, Procurement) – others are not (Estates).
    - Requires endless persistence to follow up and make sure changes are made.
  - Achieving change requires more admin support than you can possibly imagine:
    - Several full-time staff handle the organisational interface layer between ARC and the rest of UCL and they are over-worked.

Lessons learned

- Staff abused by multiple previous re-orgs view all change with suspicion:
  - Initially there was a gap between new and existing staff in willingness to seize opportunities from the change – this is improving.
  - Related to this: existing staff who were hired to be professional IT staff may not actually want to be researchers!
    - Need to make sure promotions + grading scheme understands + rewards excellence in this area too!

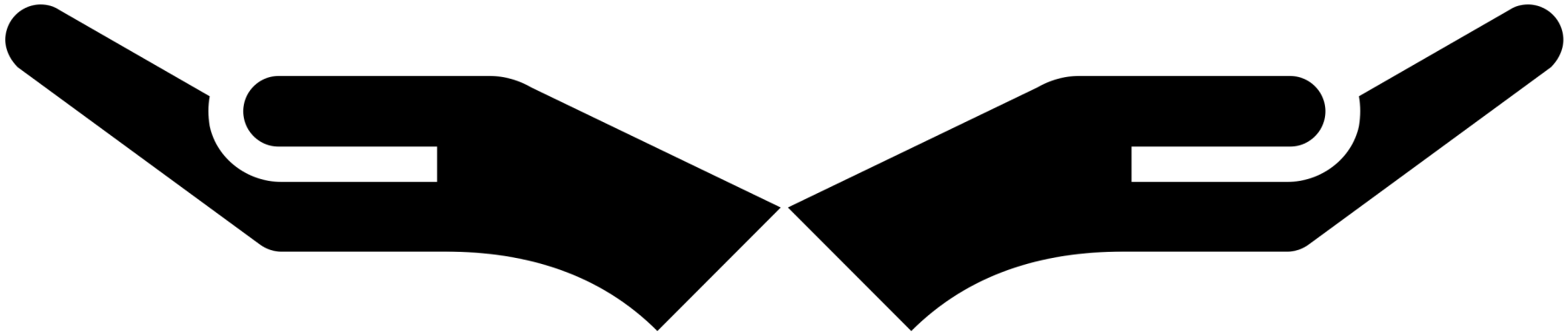
Lessons learned

- Recruiting at scale is hard:
  - At least three admin staff helping from HR + ARC.
  - All “Product Owners” involved in shortlisting.
  - Interview panel requirements.

- “General” JDs are very wide and can turn off candidates.

Offer of support

- We think this is a great model for how to run the “career management” part of running a research IT department.
- We haven’t got all the rough edges sanded off yet (e.g. hiring fudges).
- We want to spread this practice to other institutions and are happy to come and discuss what we did with your senior managers/finance/HR/whoever to help make the case.



CIUK 2023 Presentations

Matt Penn (King's College London)

King's e-Research portal: making things easier, securely

Abstract: Since King's central e-Research (research computing) department was formed in 2021 we have been developing a bespoke web portal for users to configure and access our computational services (HPC, Private Cloud, TRE). Engineered with security and ease of access as primary goals the portal enables (among other things): SSH multi-factor authentication management, group membership management, web proxy management to private cloud, virtual desktop access. In this talk we will detail the implementation of some of these features and describe how they are making access to advanced computational resources easier for researchers across King's. Along with the technical elements we will describe the skills required to embark on such a project, the collaboration involved with our central IT department and road ahead as we start to think about software sustainability.

Bio: Matt leads the infrastructure and operations teams in the e-Research department at King's College London. In the 3 years since the department was launched the team have rolled out HPC, private cloud and TRE services as part of King's Computational Research, Engineering and Technology Environment (CREATE). Prior to this Matt managed the computing team in the Faculty of Natural, Mathematical and Engineering Sciences, before that he spent 10 years in the hedge fund industry in various systems roles.

King's e-Research Portal

Making things easier, securely

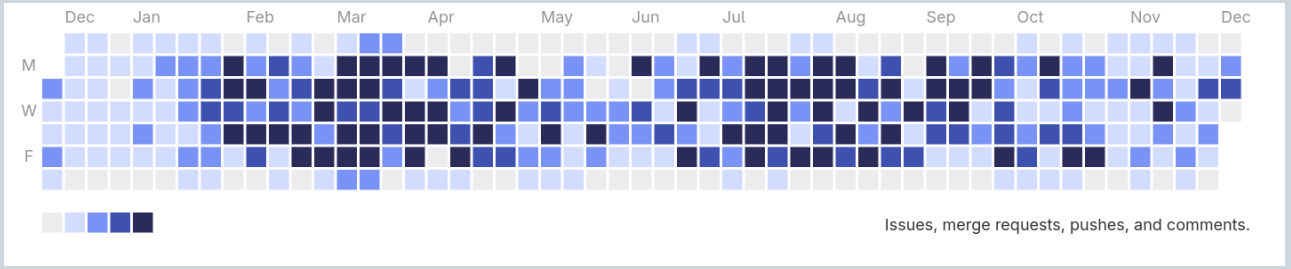
Matt Penn 7/12/2023

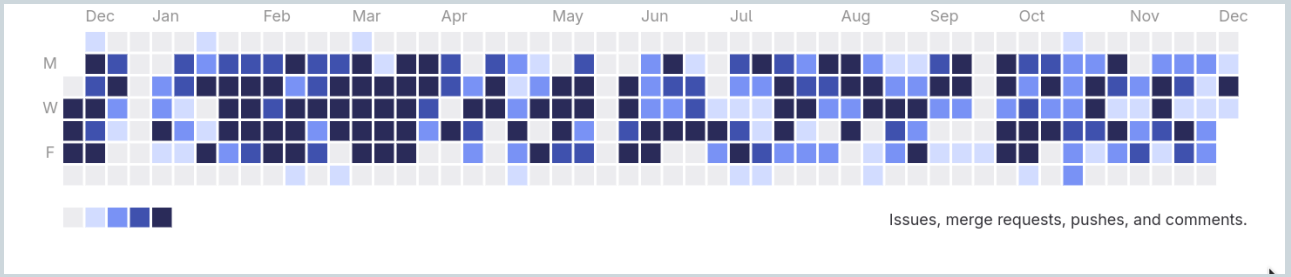
King's e-Research

- In late 2021 funding was secured to hire and build a central research computing function at King's
- Prior to this there was no institution wide research computing facilities
- 16.5 FTE
  - Director
  - Infra, Data, Information Governance and Trusted Research Environment Leads
  - 2 x Research Software Engineers (4 posts being appointed)
  - 2.5 x Senior Infrastructure Engineers
  - 6 x Research Operations Engineers
  - Admin Manager

Senior Infrastructure Engineers

- Skylar Kelty
- Xand Meaden
- Deep infrastructure knowledge
  - Linux administration at scale
  - Software development
  - Storage
  - Hardware
- Lots of python, Puppet, PHP, ...





King's CREATE

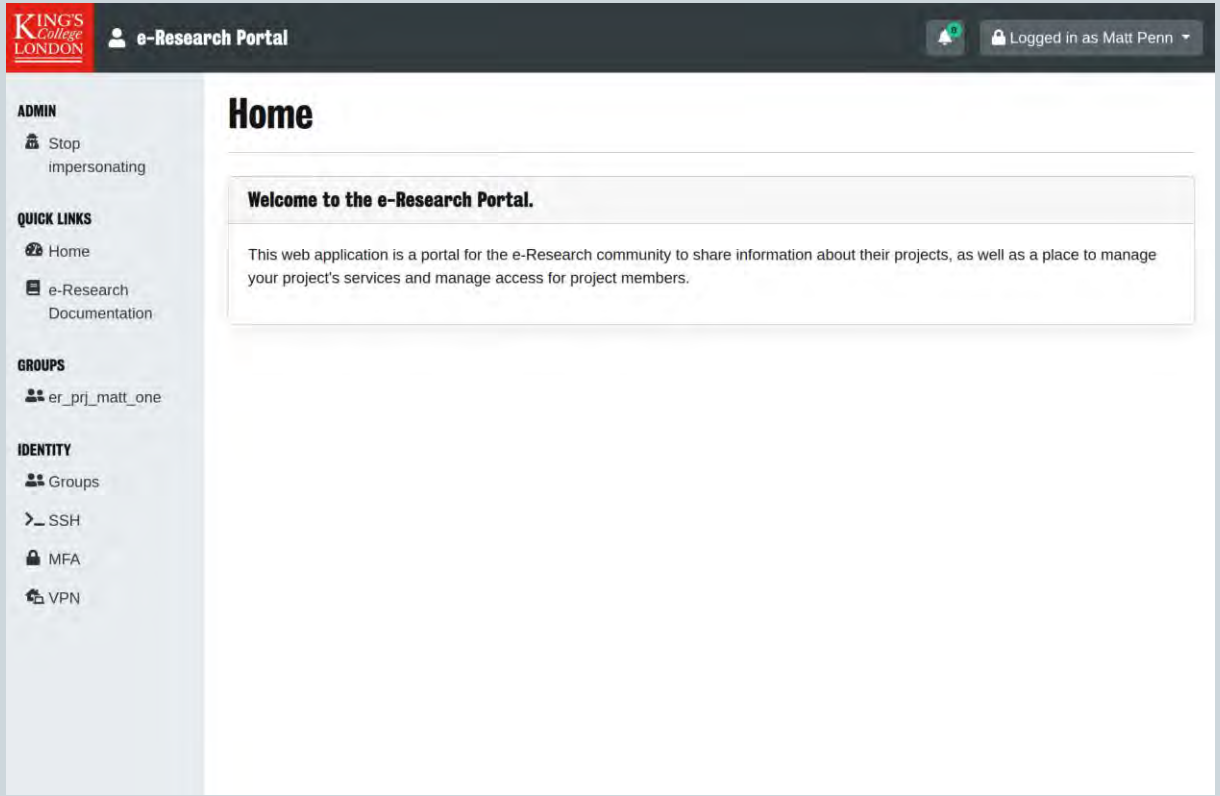
- HPC, private cloud, TRE, web farm
- Went live in April 2022
- Security is a priority
- Ease of access
- All deployment and maintenance in-house
- StackHPC for OpenStack support
- We use Ceph and Ubuntu for everything

e-Research Portal

- In this talk
  - SSH MFA
  - Group management
  - Web proxy self-service
- Not in this talk
  - Guacamole based VDI (to Trusted Research Environments and Private Cloud)
  - Service account provisioning (for application access to shared storage)
  - Read-only account provisioning (for sharing of datasets via SFTP)
  - OpenVPN certificate generation
  - Static website, WordPress and PHP-FPM provisioning
  - OpenStack project provisioning

e-Research Portal

- PHP/Laravel app
- Organically developed since 2021
- Very KCL specific
- Integrated with King's ID via SAML 2
  - SimpleSAMLphp
  - Portal -> idp.er -> MS "Entra" IdP
- Entra can be configured for tighter MFA timeouts on specific "apps", e.g.:
  - 7 days in general
  - 6 hours for TRE functionality
- Entra can implement "terms of use" acceptance flow



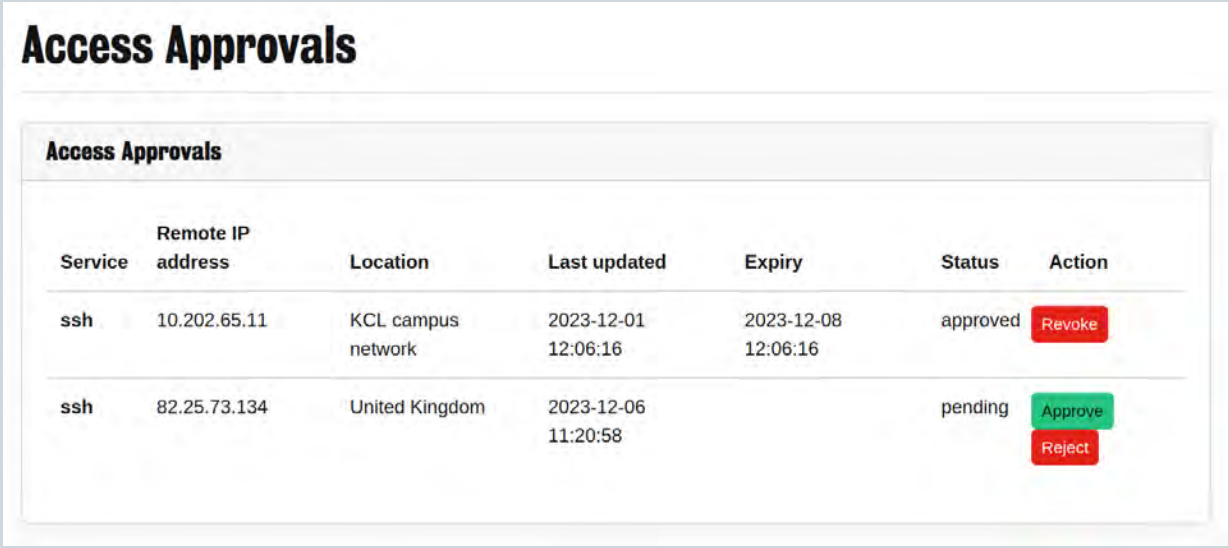
Securing SSH access with UX in mind

- Started exploring SSH MFA in 2020
- Wanted to avoid
  - Managing OTP tokens: confusing for users, more to implement for us
  - Microsoft RADIUS: PAM plugin didn't support prompt for codes
- Suggestion from MSc student with industry experience: move the MFA to a web portal
- If we implement SSH MFA mechanism in e-Research Portal, we:
  - Utilise King's ID and token management
  - Support generic OTP clients in addition to MS proprietary
  - SSH client doesn't need to MFA on every connection

sshd_config

AuthorizedKeysCommand

Banner



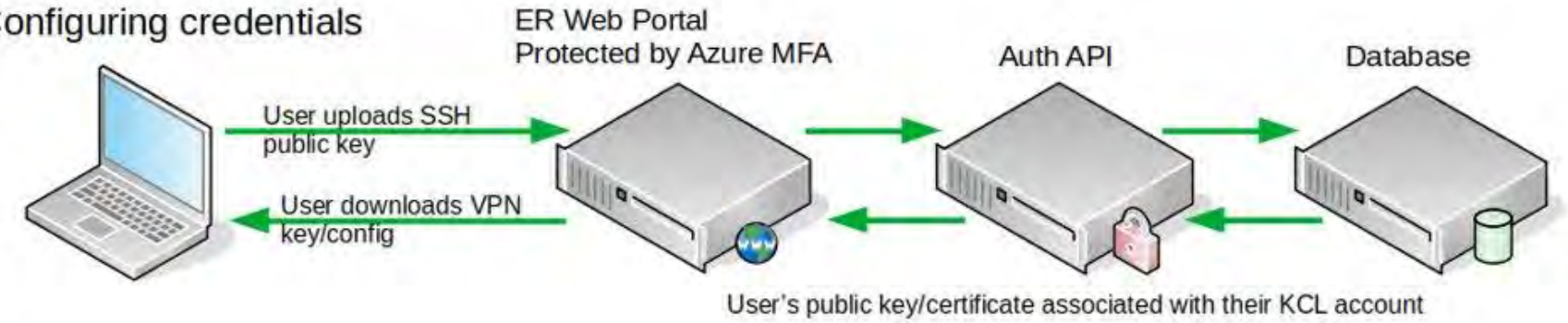
auth_api

- python flask API w/ mysql database
- Tracks connection approve/deny actions made in Portal
- Store account specific loosenings, e.g.:
  - Removal of MFA (for limited access data ingress/egress use cases)
  - Restriction to specific commands (rsync, SFTP)
- Auxiliary scripts, e.g.:
  - bastion/logon nodes: determines source address of sshd connection and confirms MFA state in
  - Slurm nodes: check if user has a running job

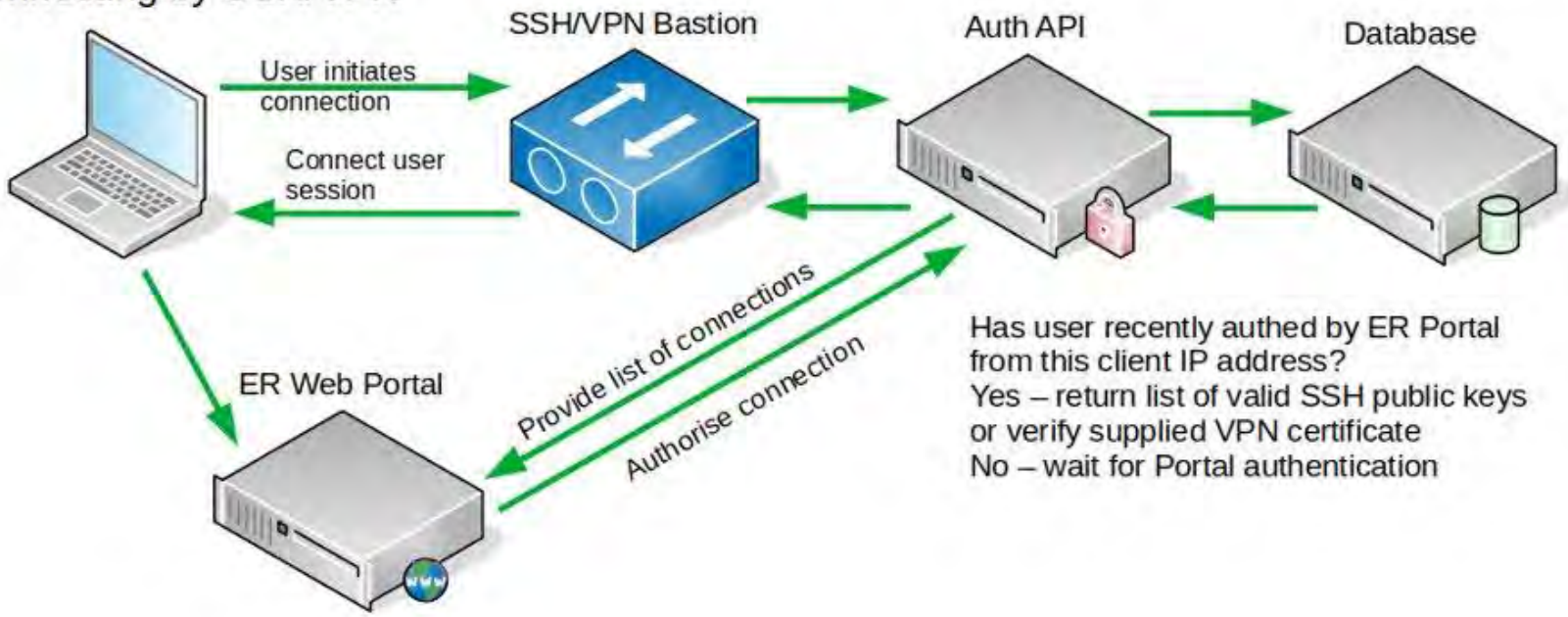
https://github.com/kcl-eresearch/auth_api

e-Research SSH/VPN Bastion Authentication

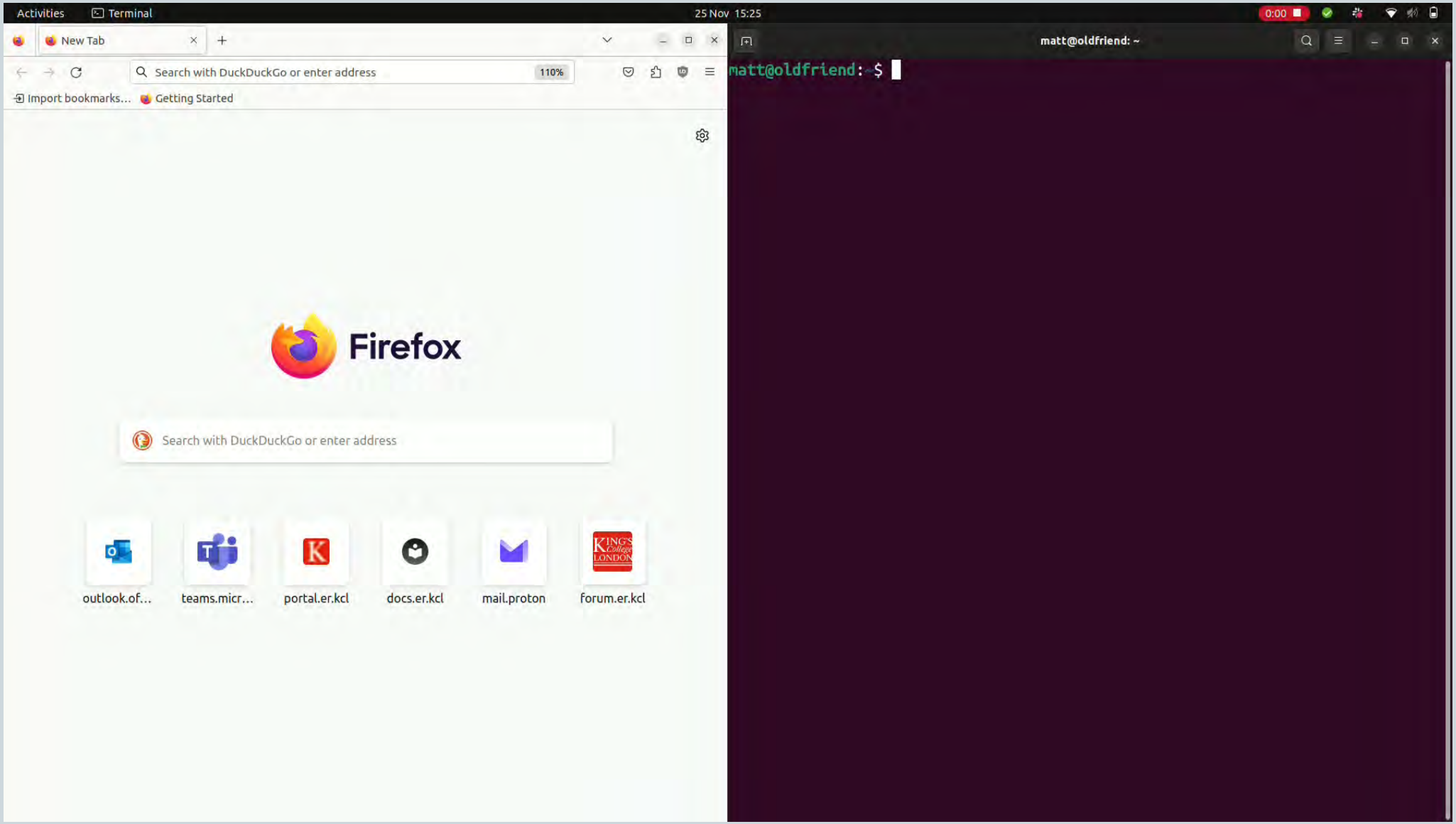
A. Configuring credentials



B. Connecting by SSH/VPN



SSH MFA access to CREATE HPC https://youtu.be/HEVx4CeIv-8



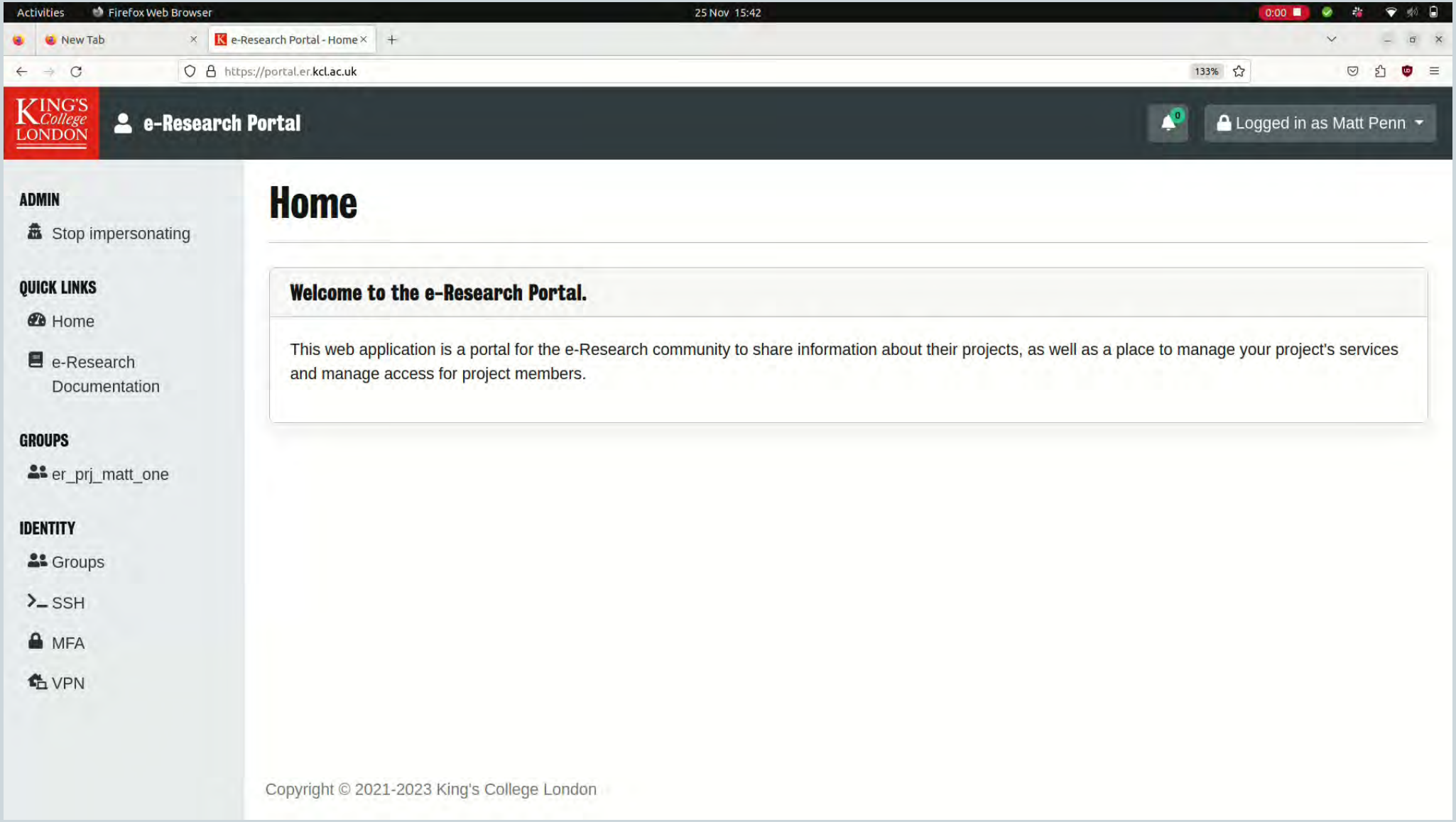
Usage to date

- 1000+ users since CREATE HPC launched in April 2022
- 458 unique HPC logons in October
- Also used for SSH bastion (with -J / ProxyJump) to:
  - OpenStack VMs
  - On-campus SSH servers

Group Management

- Most obvious thing to self-service well for efficiency
- Again, lean on institutional AD
- TRE increased motivations for:
  - Audit log
  - Increase context for approvals (department, user type, photo)
- Sync mechanics
  - Per action pushes to from Portal to AD (e.g. add/remove this user to group)
  - Per group sync job from AD to Portal runs every 6 hours

Group member addition https://youtu.be/-r7AwiWifW0



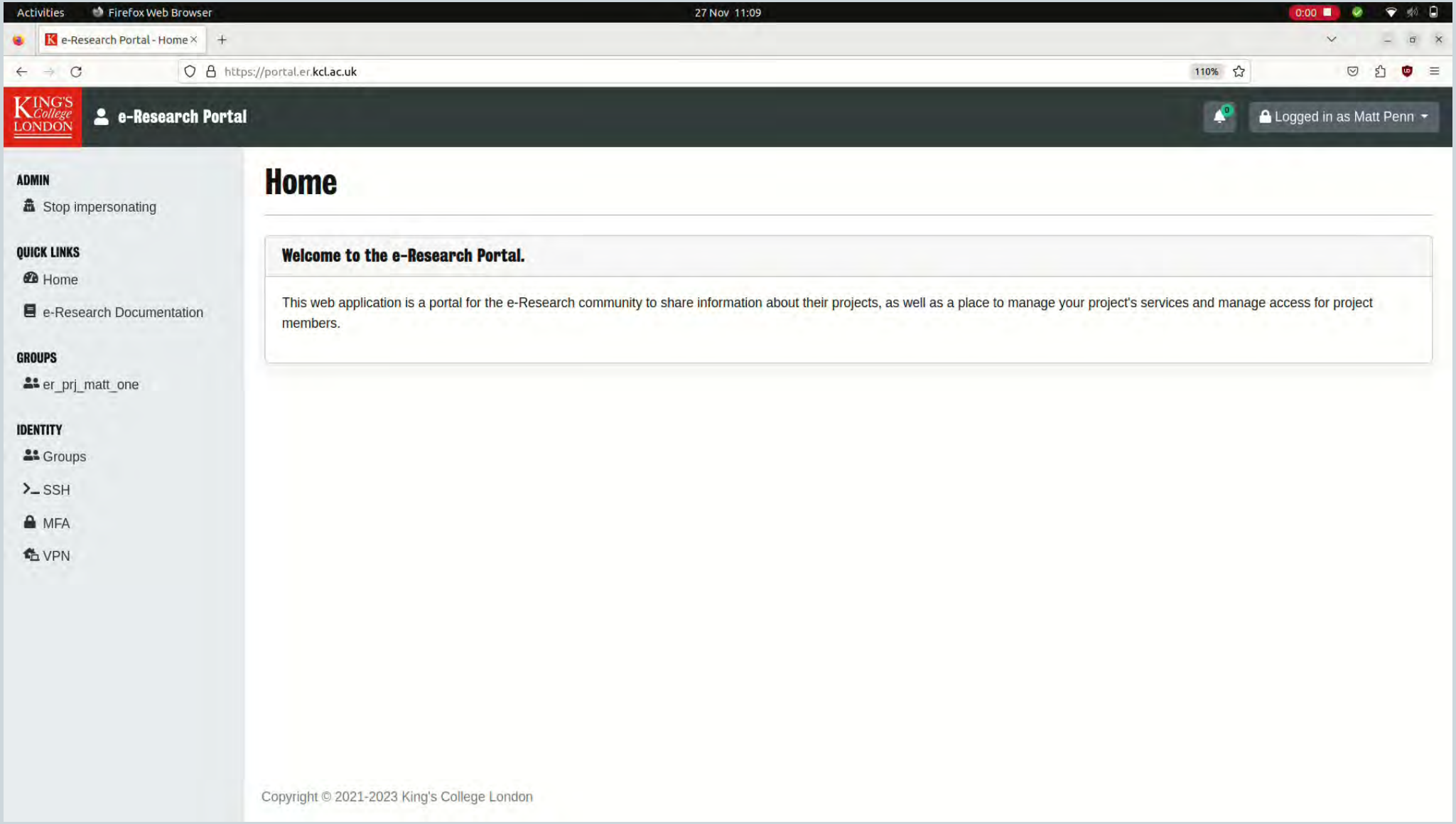
Web proxy self-service

- Want users of our private cloud to be able to expose web applications to the outside world at ease
- NGINX / Apache plumbing
- Take care of certificate signing (Let's Encrypt)
- Built-in vulnerability scanning (OWASP-ZAP)
- Web application firewall support (ModSecurity)
- Allow authn/authz from mod_mellon upstream of app
  - Allow restricted access to internal applications to save users time on implementing themselves
  - Allow "crusty" things to be run behind appropriate security controls

https://github.com/kcl-eresearch/webfarmd

Proxying to OpenStack website and enabling institutional auth

https://youtu.be/ja-DbF13myI



Use case example 1 – crusty old stuff :)

- We've all been there right...
- … vendor application tied to scientific instrument
- CentOS 6
- HTTP
- Basic auth
- We can layer on top:
  - HTTPS
  - MFA
  - Group based authorisation

Relevance alarm!!

Use case 2 – web app for inference

- Recently deployed A30/A40/A100 hypervisors in OpenStack
- Many research groups also have GPUs in private hypervisors
- Proxy through to OpenStack instance with GPU
- Want to present trained models trained on the HPC cluster externally

Disclaimer: early case so we set this up for user (using Portal)

but expecting more use as we roll out training and documentation on this pattern

Use case 2 – LLMs as predictive model from patient data

- https://foresight.sites.er.kcl.ac.uk/
- LLM trained on records from 1,000,000 KCH patients,
  20,000 SLaM patients
- Use inputted patient timeline to predict what will happen next
- Funded by NIHR Maudsley BRC
- Preprint https://arxiv.org/abs/2212.08072

You need to work well with your IT department/identity team

- AD OU with delegated access
- Entra/Azure AD SAML registrations and configuration
- GraphAPI (this one also involved information compliance)

Thinking about software sustainability

- Very custom stuff
- Bus factor is a worry
- Sharing knowledge and effort with RSE team
- Possibility of open source projectification
- Hiring junior posts (inc. industry placement)
- Keep calm, assess risk and trust your engineers

Thanks!

- Sky and Xand for the incredible work
- The rest of the e-Research team
- King's IT
- King's researchers
- Our partners
  - StackHPC
  - Lenovo

Questions?

Contact details/for more information

Now

Later

matt.penn@kcl.ac.uk

HPC-SIG Slack

https://docs.er.kcl.ac.uk/

CIUK 2023 Presentations

Dimitrios Bellos (Rosalind Franklin Institute)

Accelerating Structural Biology: From PC to HPC

Abstract: The Rosalind Franklin Institute is a centre dedicated to tackle important health research challenges by developing new technologies. It is an interdisciplinary institute, which provides step-change research through intervention, imaging and interpretation. AI accelerates Structural Biology in these key areas. In structural biology studies, cryo-EMs produce multi-terabyte datasets which all need to be analysed in reasonable time. Earlier this year, users analysed their cryo-EM data on the Franklin's cloud infrastructure, however the time to process these data was increased to a few days as the number of users increased 10-fold. This prompted us, to rely on a higher degree on external computing resources, using our partners at the Baskerville Tier 2 HPC. In our attempt to do so however, we faced a significant challenge. How to make Baskerville more convenient for our users, who never worked on HPCs and prefer UI tools, with the minimal amount of training. This talk will cover how the RELION suite was able to run graphically on Baskerville. How fast data transfer data were facilitated with Globus. How we engaged all related groups to resolve technical issues, perform testing and profiling. And how we offered training material through documentation and open communication channels.

Bio: Dimitrios is a researcher and software developer, in the Artificial Intelligence and Informatics (AI & I) theme in Rosalind Franklin Institute. As a member of the Franklin's AI & I Core team he works on the development and support of Franklin's compute and data infrastructure. Furthermore, he researches into DL/ML approaches for the processing or enhancing of Electron 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 and segmentation, deep/machine learning, computer vision and image processing.

The Rosalind Franklin Institute

Accelerating Structural Biology: From PC to HPC

Dr Dimitrios Bellos

Rosalind Franklin Institute, Baskerville HPC, Collaborative Computational Project for Electron cryo-Microscopy (CCP-EM)

Rosalind Franklin Institute
Artificial Intelligence and Informatics:
- Dr Laura Shemilt
- Gabryel Mason-Williams
- Dr Joss Whittle

Rosalind Franklin Institute
Structural Biology:
- William Bowles

Baskerville HPC:
- Dr Gavin Yearwood
- Dr James Allsopp
- Dr Jenny Wong
- Dr Simon Hartley

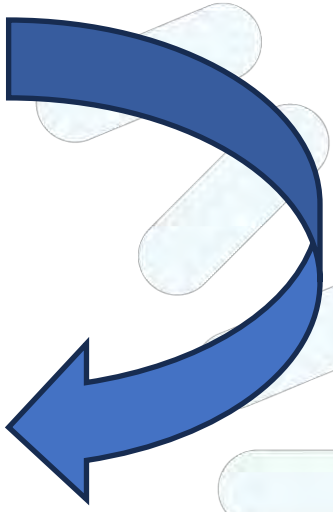
CCP-EM:
- Dr Colin Palmer
- Dr Tom Burnley

The Rosalind Franklin Institute

Software used for Structural Biology

Structural Biology Software :

- It is tested to run on a single machine - PC.
  Computational bottleneck
- Have Graphical User Interface (GUI) that it is preferred
- It is being used to process multi terabytes of data
  Even though, not large local storage or slow connection with external
- A single user can use the PC at a time



Can we bring and use Structural Biology Software on HPCs ?

High Performance Compute cluster Software:

- Ideally should run on multiple machines
  High compute resources
- They are operated via terminal commands
- High speed connection with large storage
- Multiple users can schedule jobs and use it simultaneously

The Rosalind Franklin Institute

REgularised LIkelihood OptimisatioN (RELION)

One of the most commonly used software for Structural Biology is RELION.
It employs empirical Bayesian approaches for electron cryo-microscopy (cryo-EM) structure determination.

- Bringing RELION on an HPC can accelerate science tremendously and increase the number of publications.
- Operations that take 2 weeks in a single machine may now be done in a few days.
- The advantage of RELION is that it has the capacity to be run HPC, especially a GPU cluster.

The Rosalind Franklin Institute

REgularised LIkelihood OptimisatioN (RELION)

An example of highly impactfully publication thanks to RELION

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.

The Rosalind Franklin Institute

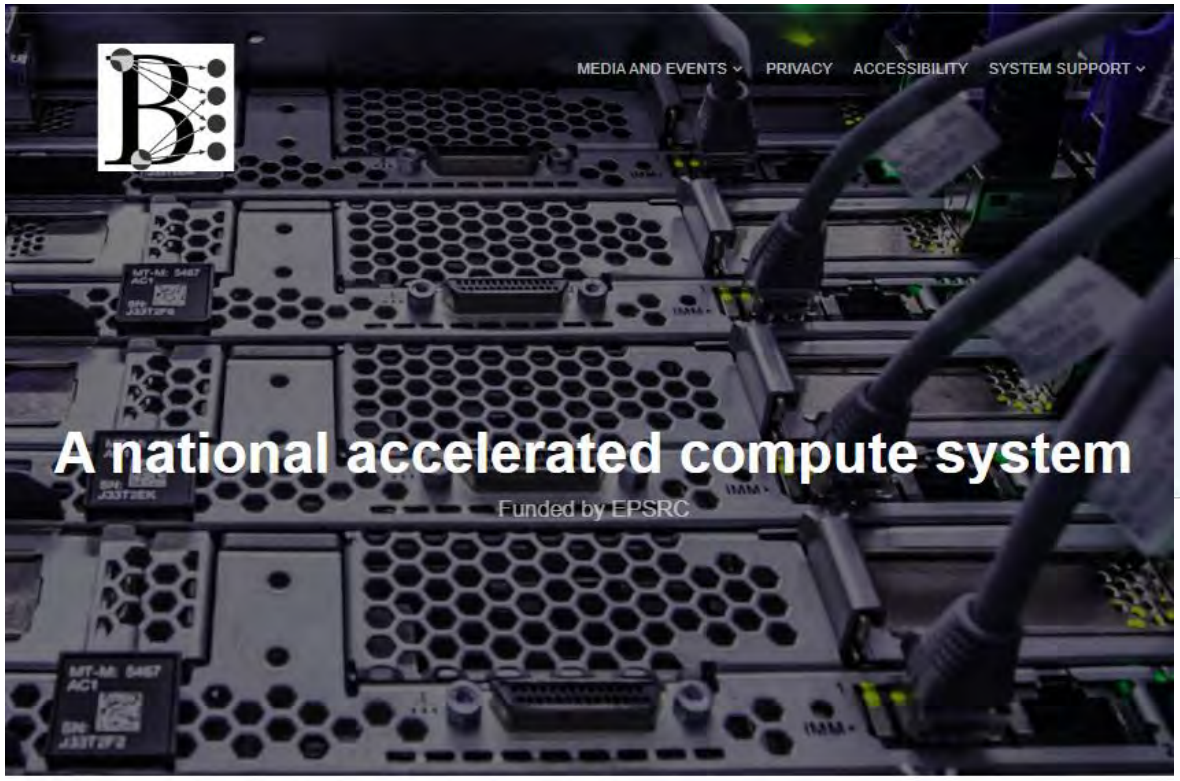
REgularised LIkelihood OptimisatioN (RELION)

However, to bring RELION in an HPC (Baskerville) there were some requirements to make it highly accessible to its users

1. Offer a way to use RELION GUI approach
2. Allow fast data transfers between the cluster and the data storage
3. Resolve potential technical issues, perform testing and profiling when this requires multiple groups to co-ordinated (Franklin biologist and AI core Team, Baskerville HPC, CCP-EM)
4. Educate the users how to use RELION on an HPC instead of PCs and single machines.

The Rosalind Franklin Institute

Baskerville HPC cluster



Baskerville launched to users in July 2021

Read details of our Baskerville launch event.

208
GPUs

52
Compute Nodes

5400
Storage (TB)

Baskerville HPC offers provides a high compute resources and in particular there are 52 nodes and a total sum of 208 high performance Nvidia A100s GPUs.

Offers a large data storage with fast connection

Optimal for running software that can run in a multi-node multi-GPU setting

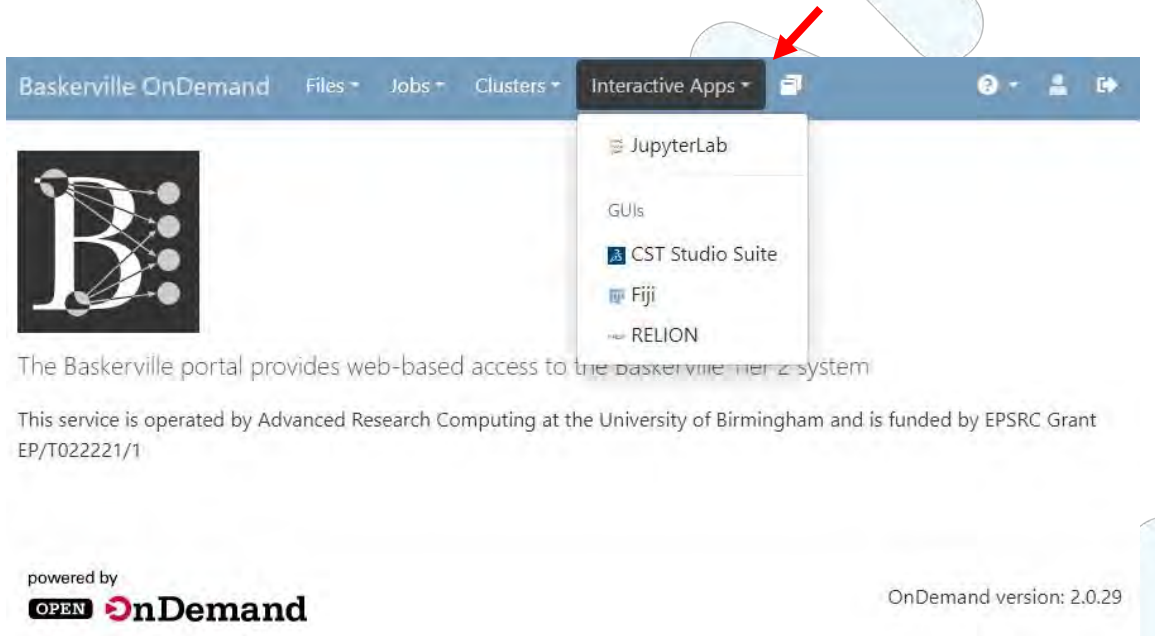
The Rosalind Franklin Institute

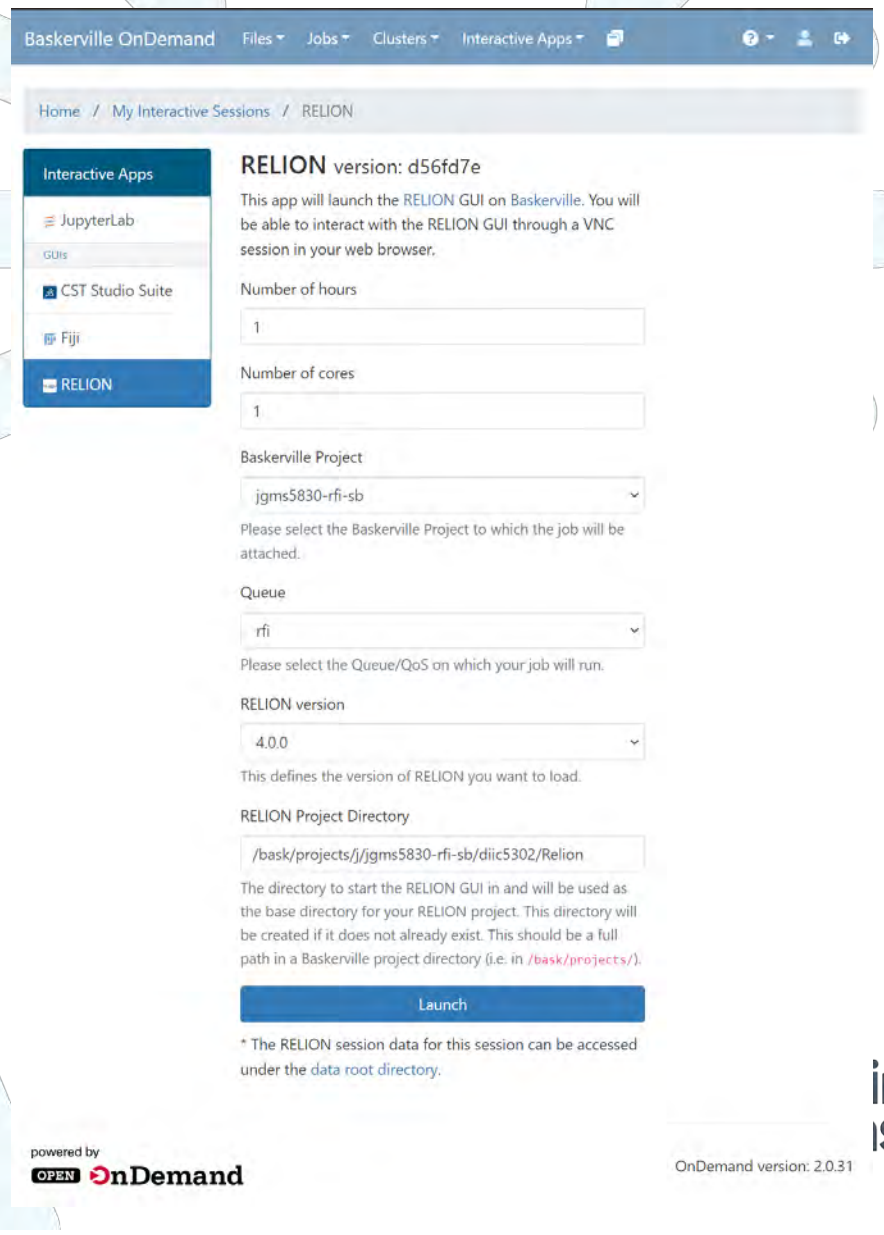
Interactive RELION App on Baskerville HPC

As it is widely known, learning to use a software via a GUI is is easier than with terminal commands

An interactive app has been created on [Baskerville Portal]() that allows the users to launch RELION on Baskerville and to display its GUI

To login to Baskerville Portal 2FA is being used





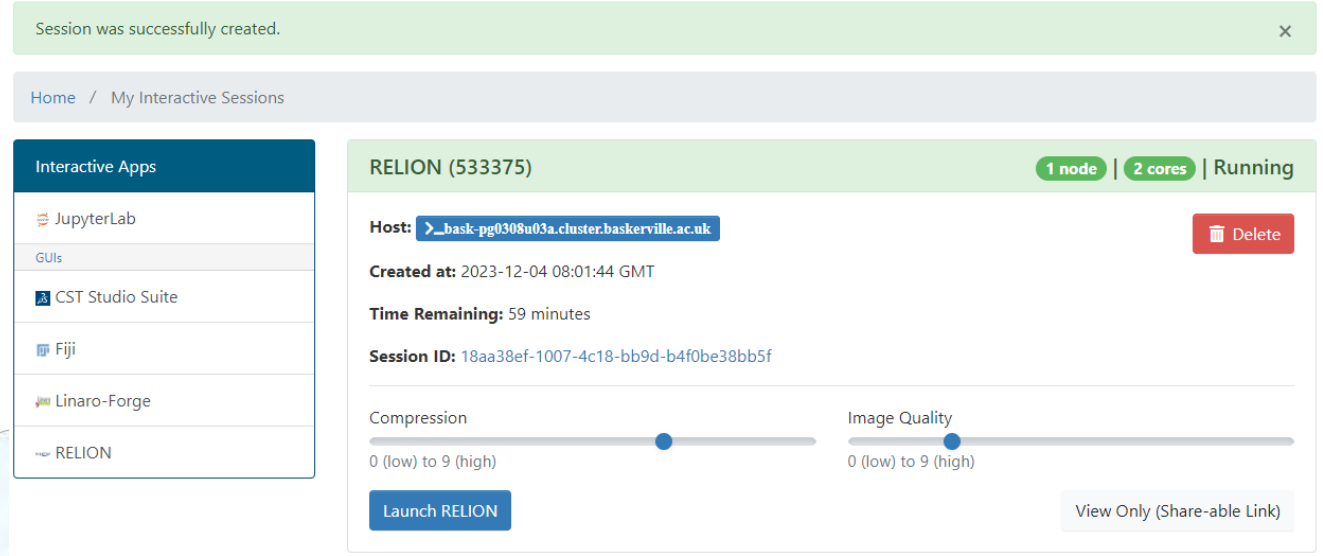
Interactive RELION App on Baskerville HPC

When a new interactive app is being requested a single GPU on a single node has to be reserved

This is needed for the graphics part of the Graphical User Interface

After the user clicks the button ‘Launch RELION’ a new browser tab opens with RELION GUI.

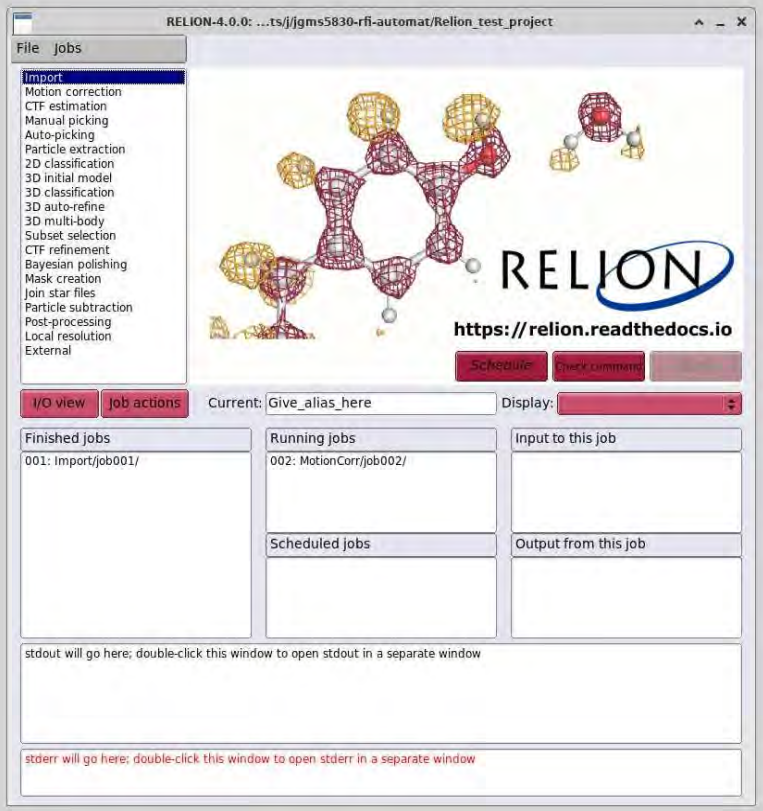
When the GUI is no longer need the user is advised to delete the session, thus releasing the allocated GPU

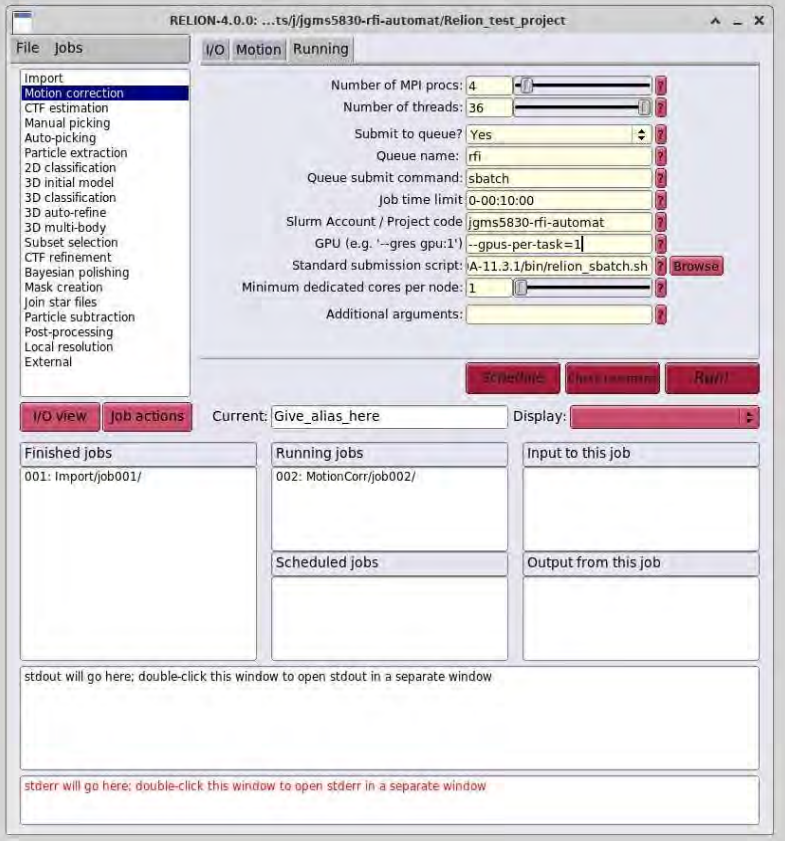


Interactive RELION App on Baskerville HPCs

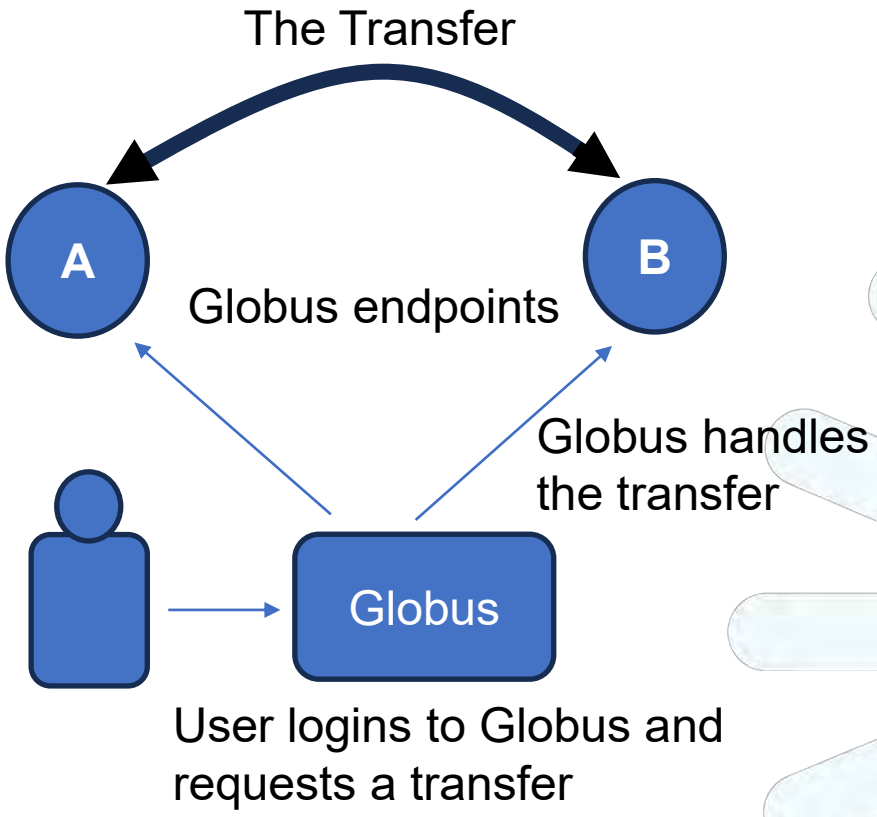
For almost every process on the list there is a GPU acceleration option. In the final ‘Running’ tab the user can specify slurm scheduler options.

After clicking ‘Run!’ a new independent slurm job will be submitted that will not be killed if the Interactive session is closed





Fast data transfers to Baskerville using Globus



Globus is a service that allows fast data transfers between machines where a Globus endpoint has been setup

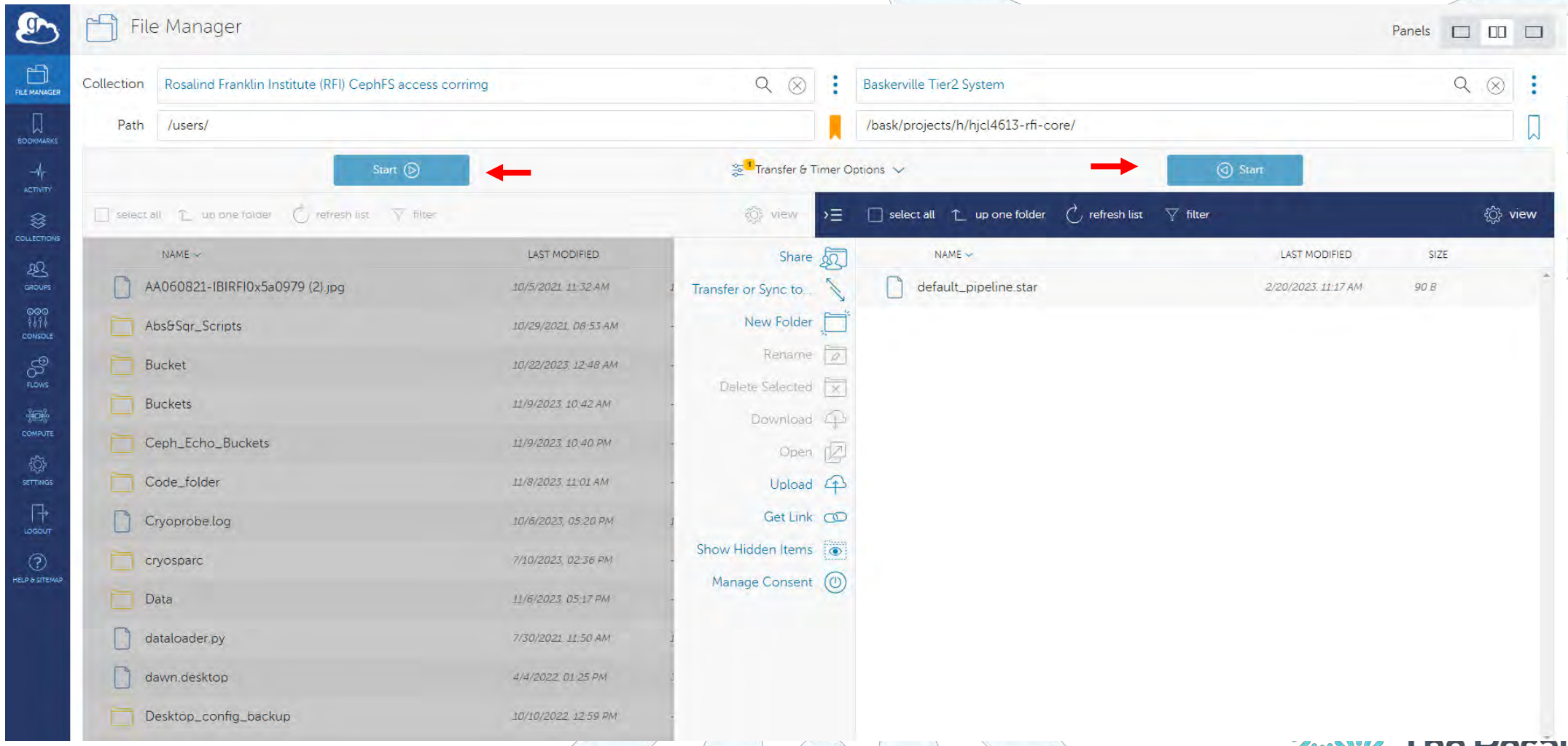
Globus allows encryption and different levels of the visibility for its endpoints and permissions

Transfers initiated via Globus website

Very intuitive website design

The Rosalind Franklin Institute

Fast data transfers to Baskerville using Globus



The Rosalind Franklin Institute

Fast data transfers to Baskerville using Globus

To put things in perspective:
From 29/09 - 25/10 were able to transfer 724TB to Baskerville

Logins to Baskerville's Globus endpoint it requires to authenticate with 2FA

The authentication last only for 30 days

Transfers to and from Baskerville are forced to be encrypted and this option cannot be disabled by the users

Baskerville has a large storage and every Baskerville project had different storage quota set by the project investigator (PI). It is the users' responsibility to do periodic clean ups.

The Rosalind Franklin Institute

Engaging all related groups

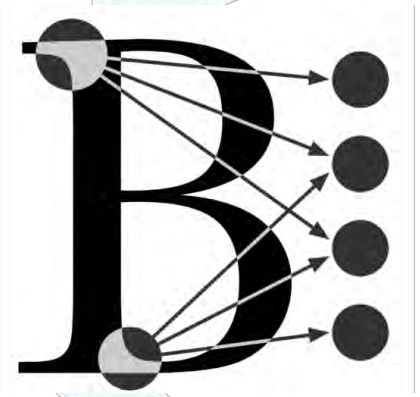
We organised meetings with different groups separately (CCP-EM, Baskerville, Franklin biologists), but also joined meetings

During these meetings we arranged to reserve Baskerville resources to be able to solve issues with live testing

Using these meeting we were able to resolve:

- Allow submitted RELION to not be killed after the interactive app session is deleted
- Find additional dependencies that were missing
- Find argument limits that we are working on removing
- Provide some initial compute argument recommendations to the users

The Rosalind Franklin Institute



CCP-EM

The Rosalind Franklin Institute

Training materials and documentation for users

- A week ago, on 27th and 28th of November the annual Baskerville training took place in Harwell campus.
- We have a website to our users where they can read recommended RELION compute settings for a single machine
- We are working on a guide for recommended compute settings on Baskerville. To do so we are working with the users helping them test and profile their Baskerville jobs.
- There is a training and documentation material offered to our users on how to use Baskerville
- Baskerville has a very comprehensive documentation page (https://docs.baskerville.ac.uk/)

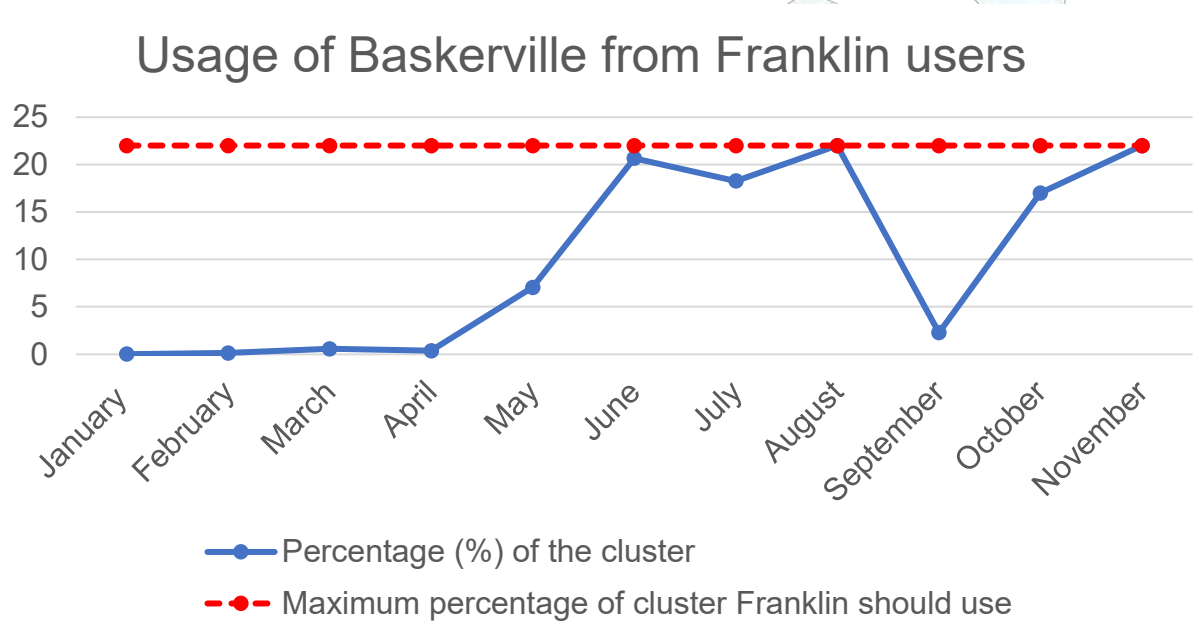
The Rosalind Franklin Institute

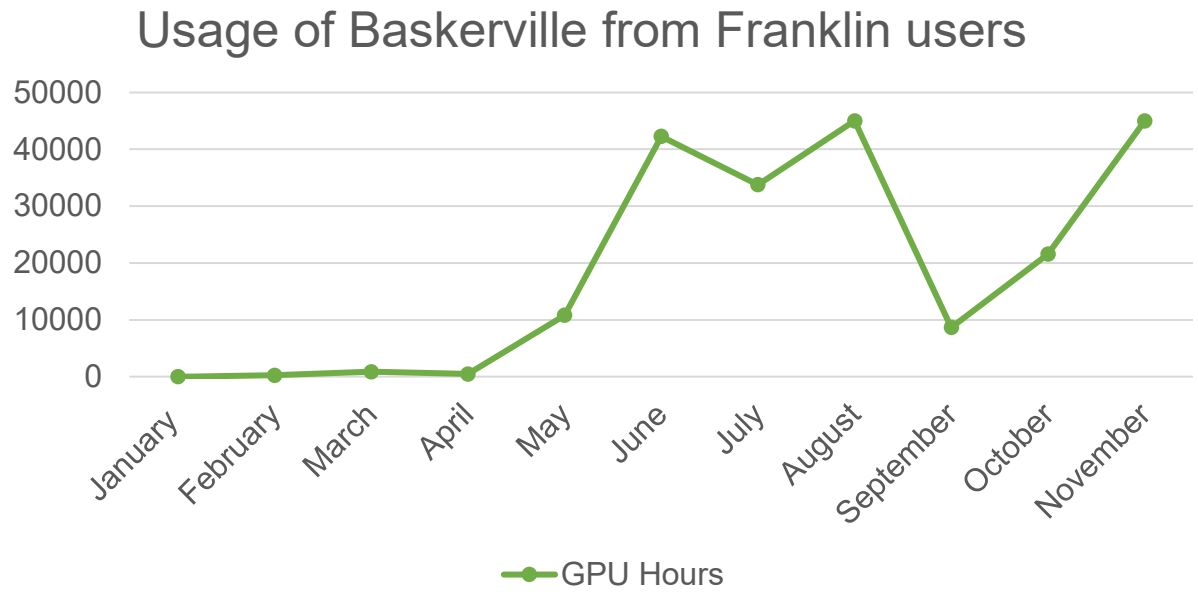
Open communication channels and support

There is a ticketing system in Franklin where users can email issues they experience when using RELION on Baskerville

The is also a baskerville-rse Slack channel which our users can use to directly contact the Baskerville team if there is an issue that needs to be addressed to them

Usage of Baskerville from Franklin Users since the beginning of the year (2023)





Conclusion and Future Plans

Things that helped:

- RELION can launch with a GUI made it more accessible to our users
- Use of Globus has facilitated fast transfers
- Joint meetings along all related groups, with dedicated resources for live testing accelerated the fixing of issues
- Offering training, documentation material and continuous support to our users

Things that do not help

- Difficult to write documentation on recommended compute options (for efficiency)
- Many manual steps required from the users
- Difficulty for users to interpret error outputs

Future Plans:

- Offer an automated system to run jobs on Baskerville (Transfer data & Compute & Return data)
- Create guides on recommended compute values

The Rosalind Franklin Institute

Thank you for your attention

The Rosalind Franklin Institute

CIUK 2023 Presentations

Maria Fando (CCP4, STFC)

Delivering HPC Power for Structural Biologists with CCP4 Cloud

Abstract: The increasing complexity of tasks, extensive reliance on databases, and the automation of CCP4 Software for protein structure determination from X-ray diffraction images have created a demand for computing resources that often surpass what is available within typical structural biology research groups. This has led to the necessity for distributed computing in crystallography, utilizing High-Performance Computing (HPC) facilities. One significant challenge is to provide users with an interface that seamlessly connects personal computing with access to HPC resources and cloud-based storage. Here we present details of CCP4 Cloud, a cloud-based project management interface enabling researchers to access these resources.

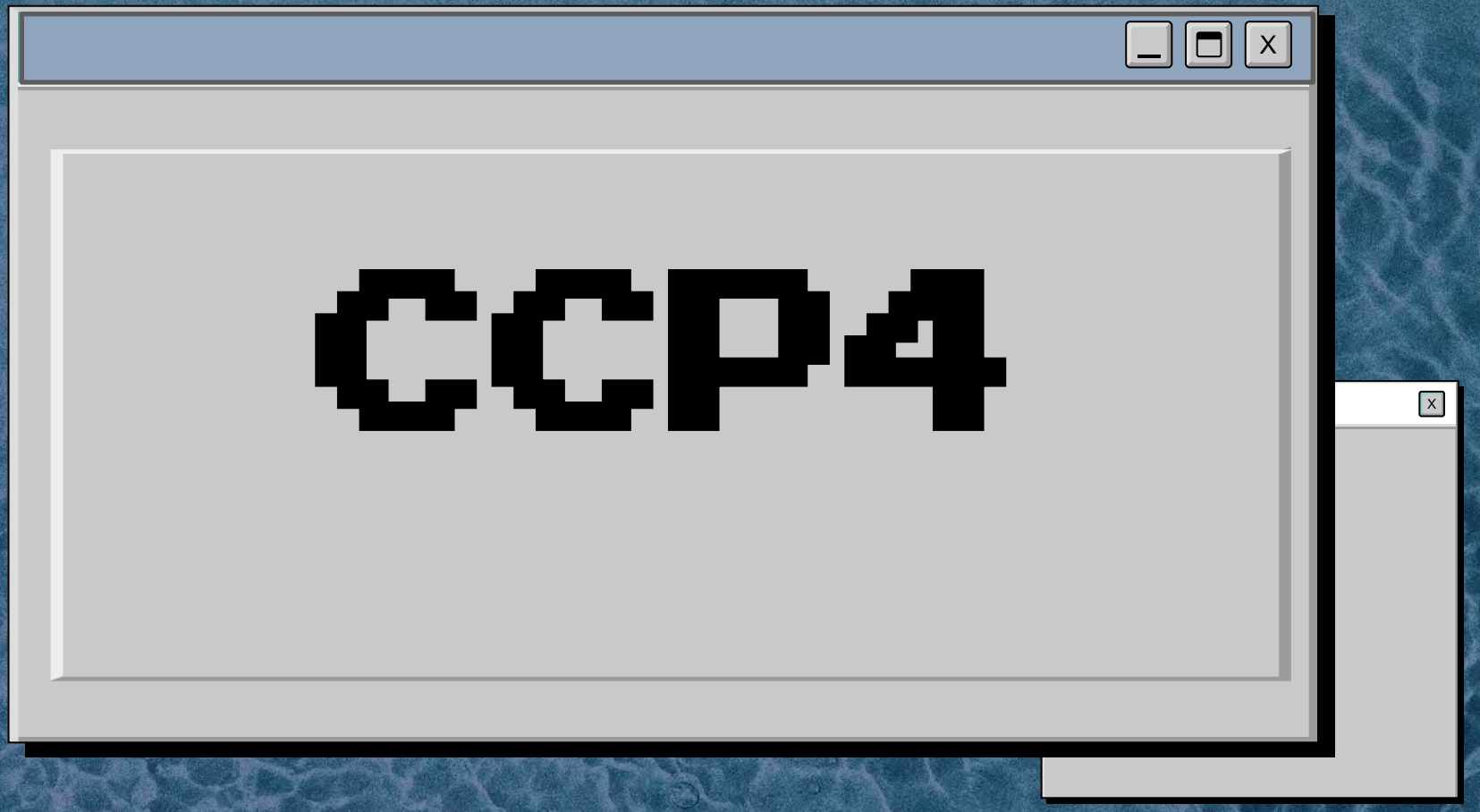
Bio: Since 2013 Maria has been working on various projects in molecular biology. In 2020 she completed her PhD in molecular biology at the Institute of Protein Research RAS. In her current post within STFC, she channelled her expertise and passion in protein crystallography into a new realm by contributing to the development of crystallography software within CCP4.

Maria Fando

Delivering HPC Power for Structural Biologists with CCP4 Cloud

Computing Insight UK 2023

07/12/2023


CCP4

"Collaborative Computational Project No. 4 in Protein Crystallography"

CCP4

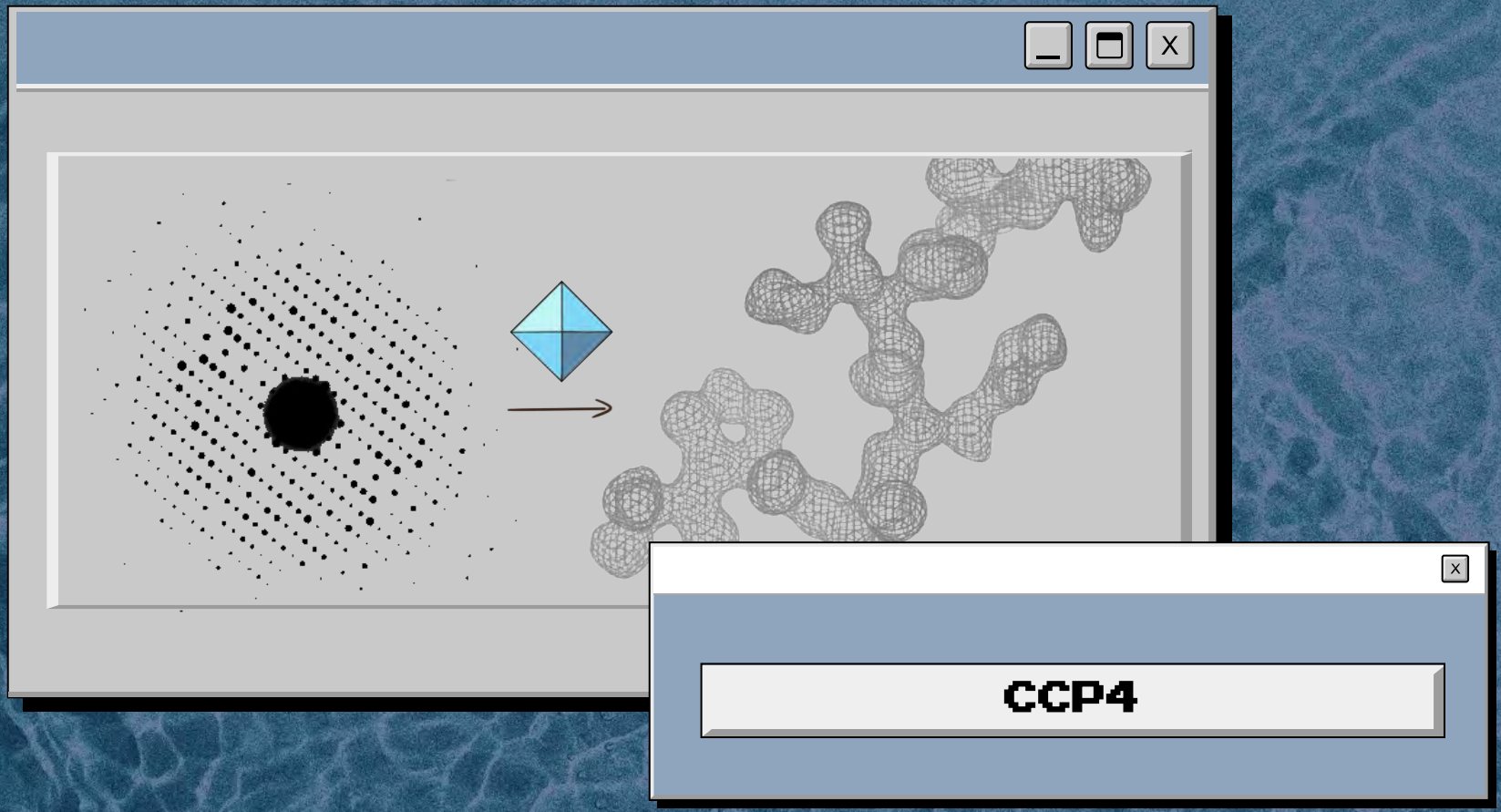
CCP4 was set up in 1979 to support collaboration between researchers working on MX software in the UK, and to assemble a software suite to satisfy the computational requirements of the relevant UK groups

CCP4

The CCP4 Suite is a comprehensive suite of software for macromolecular crystallography. It contains about 700 program components

CCP4

CCP4 user's map


CCP4

diamond

Active Site

Drug inhibitor in crystal

Data in Crystallography

Progress in the determination of three-dimensional macromolecular structures from diffraction images is achieved partly at the cost of increased computational time and data volumes

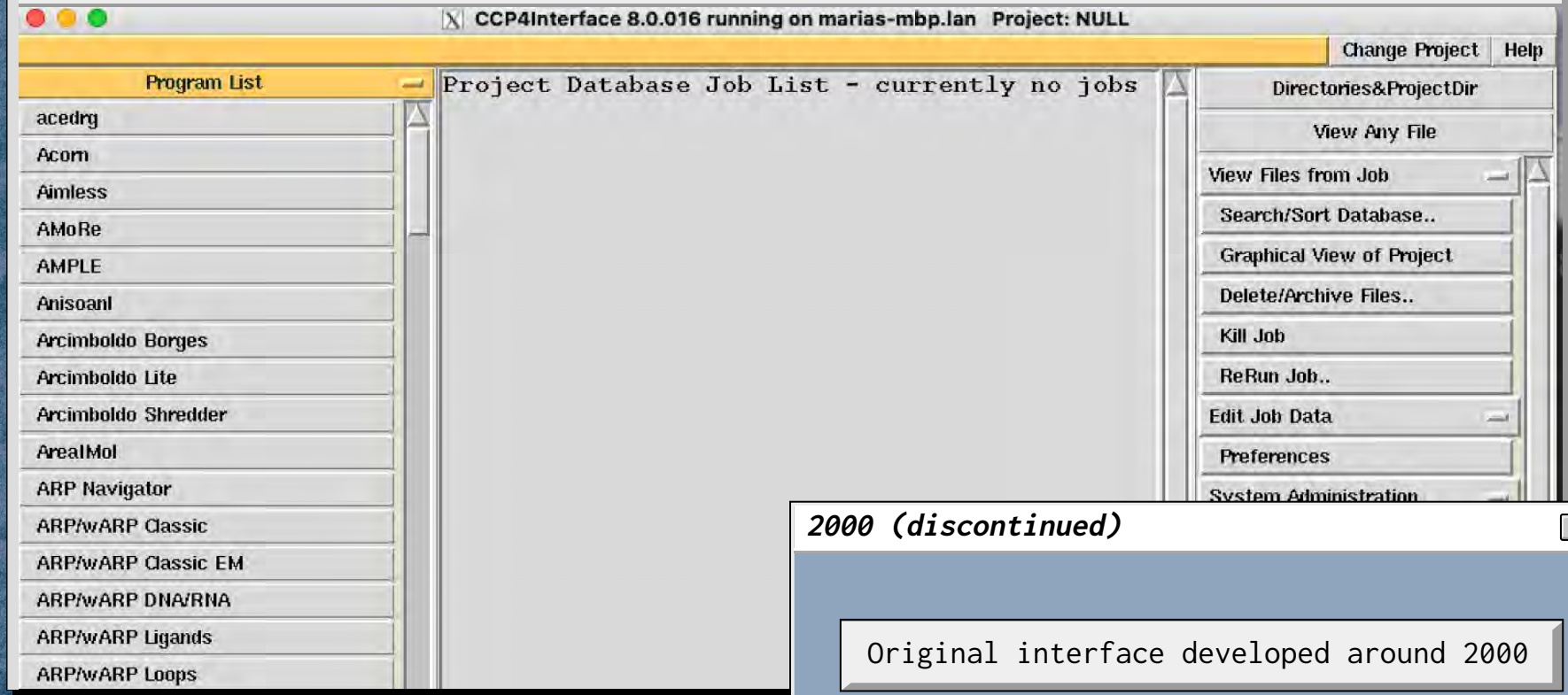
CCP4 shifted heavily to an automatic structure solution at the cost of higher CPU demand

Combination of data, AI and computing power starts giving a boost for automation

Recent example: Structure determination using predicted models

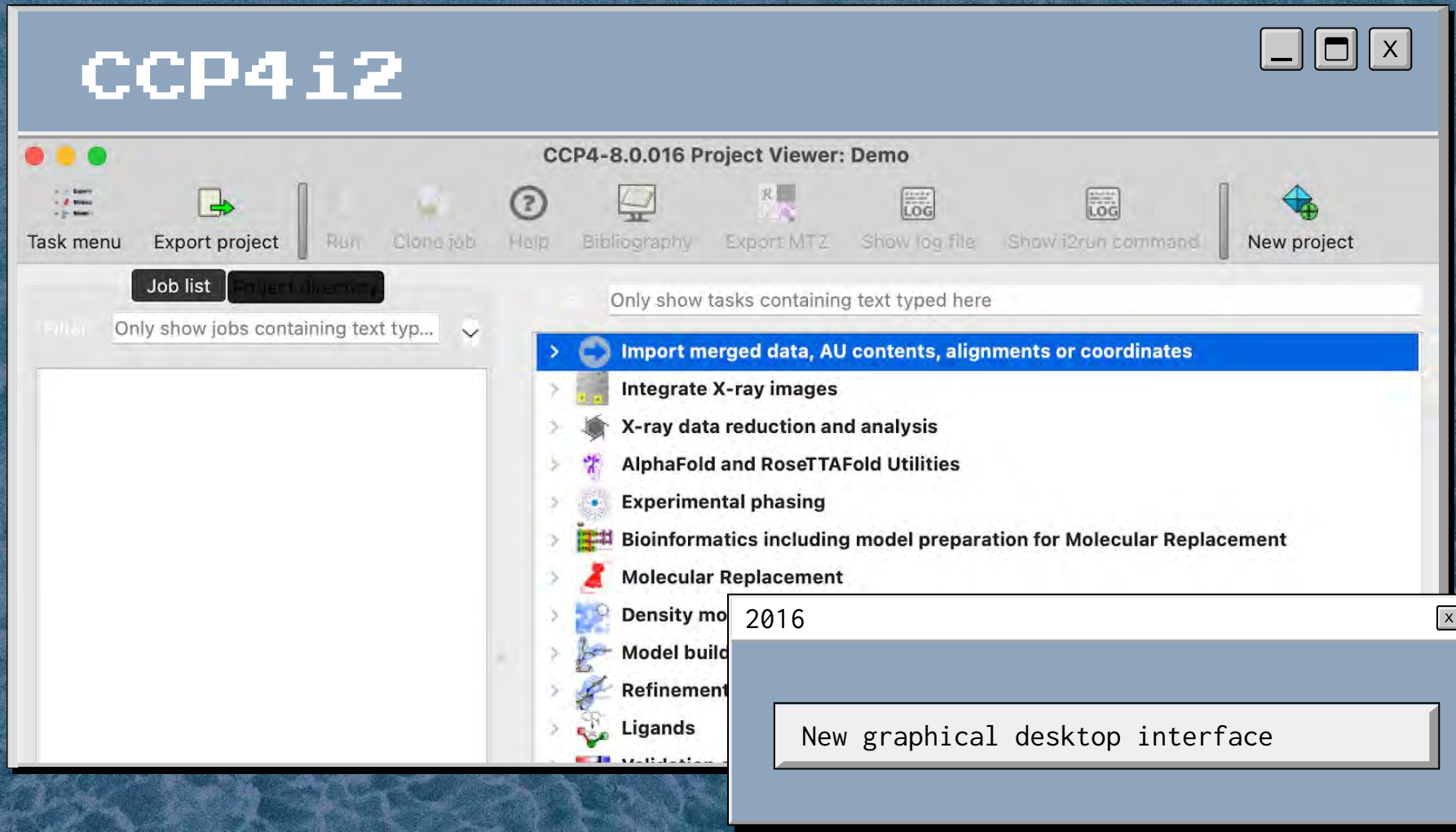
- unprecedented accuracy in predicting protein structures in 3d
- expanding proteomics to genomics scales (from 180K known to 200M structures in AlphaFold (by DeepMind) and 772M in ESM (by Meta) datatbases)

CCP4i



2000 (discontinued)

Original interface developed around 2000


CCP4i2
CCP4-8.0.016 Project Viewer: Demo
Task menu
Export project
Run
Clone job
Help
Bibliography
Export MTZ
Show log file
Show i2run command
New project
Job list
Only show jobs containing text typ...
Only show tasks containing text typed here
Import merged data, AU contents, alignments or coordinates
Integrate X-ray images
X-ray data reduction and analysis
AlphaFold and RoseTTAFold Utilities
Experimental phasing
Bioinformatics including model preparation for Molecular Replacement
Molecular Replacement
Density mo
Model build
Refinement
Ligands
2016
New graphical desktop interface

CCP4 Cloud

cloud.ccp4.ac.uk

My Projects/Demo

54% 0%:1% maria

Open | Add | Rename | Clone | Move | Delete | Export | Import | Join | Tutorials | Help

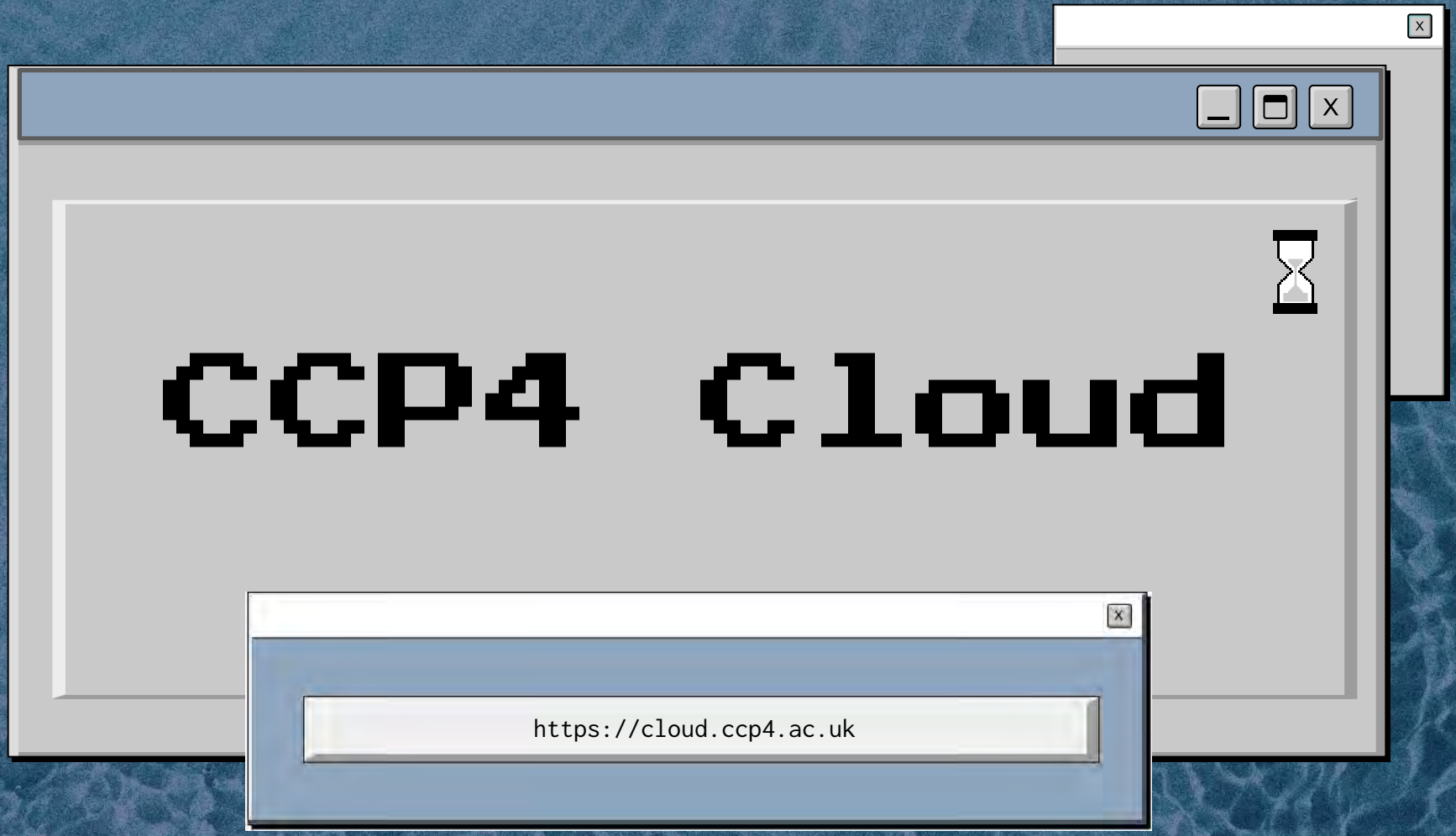
ID	Name	R_{free}	Disk (MBytes)	CPU (hours)	Date Created	Last Opened

There are no projects in folder "My Projects/Demo".

Use "Add" button to create a new Project;
"Import" button for importing a p
"Join" button for joining project
or "Tutorials" button for loa
or click on page title or folder

2018

browser based


CCP4 Cloud
https://cloud.ccp4.ac.uk

CCP4 Cloud Initiative

Conceived in 2016

- Funded by BBSRC UK and CCP4

Response to demands and trends rapidly emerging in the field

- CPU power (due to increased automation)
- Centralised database support (due to expansion of methods based on data templates)
- Software as a service (due to increased size and complexity of software setups)
- Supporting distributed projects for team work
- Cloud model for geographically-agnostic access and project data safety
- Supporting personal mobile platforms (tablets and smartphones)
- Communication with data facilities (synchrotrons, PDB, AFDB, etc)

Availability

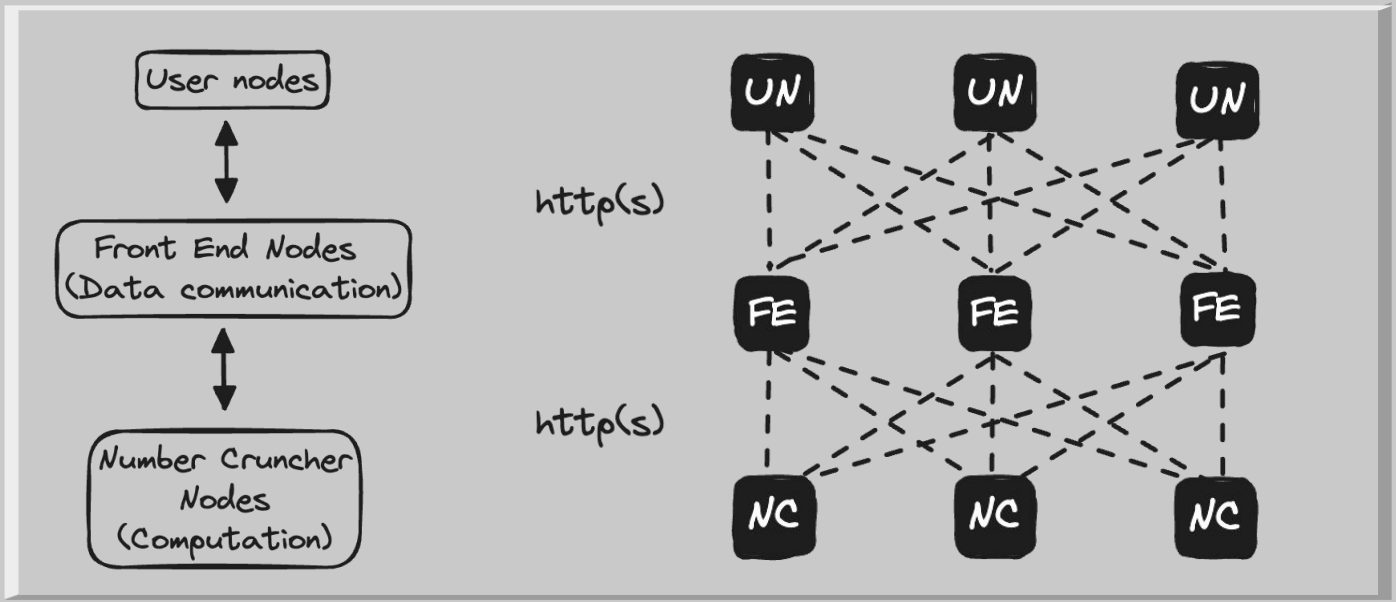
https://cloud.ccp4.ac.uk

https://www.ccp4.ac.uk/download

Part of standard CCP4 distribution package starting from 7.1 release series (since June 2020)

CCP4 Cloud Architecture

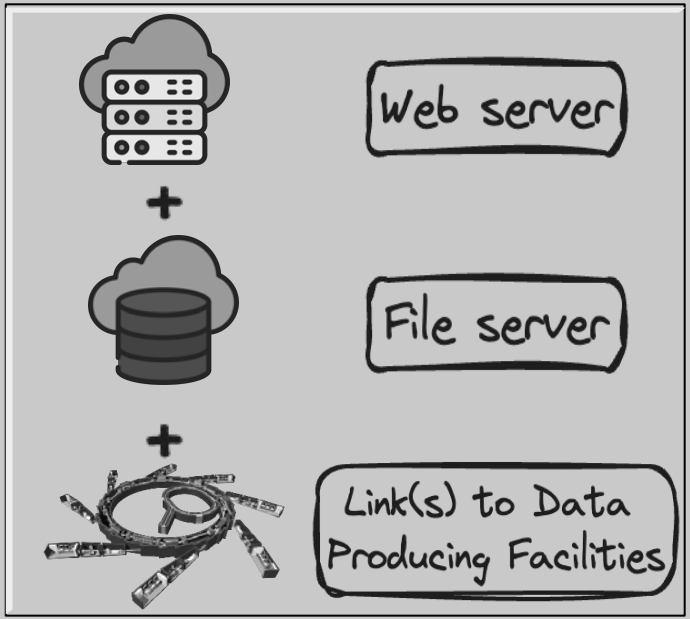
CCP4 Cloud Architecture



CCP4 Cloud Architecture

Front End Nodes:

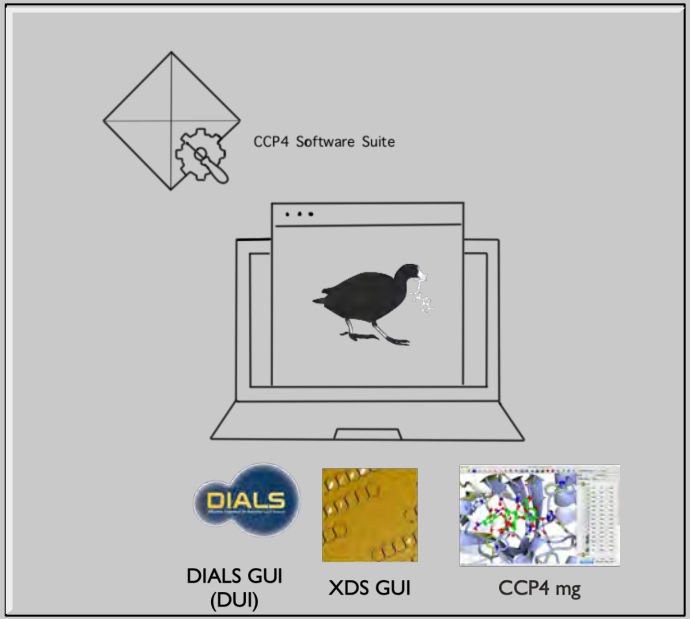
- provide all data logistics in the Cloud
- represent web-servers with storage for user data and projects
- may obtain data from Data Producing Facilities (experimental at the moment)
- do not run any calculations
- despatch jobs to Number Cruncher Nodes



CCP4 Cloud Architecture

CCP4 Cloud Client:

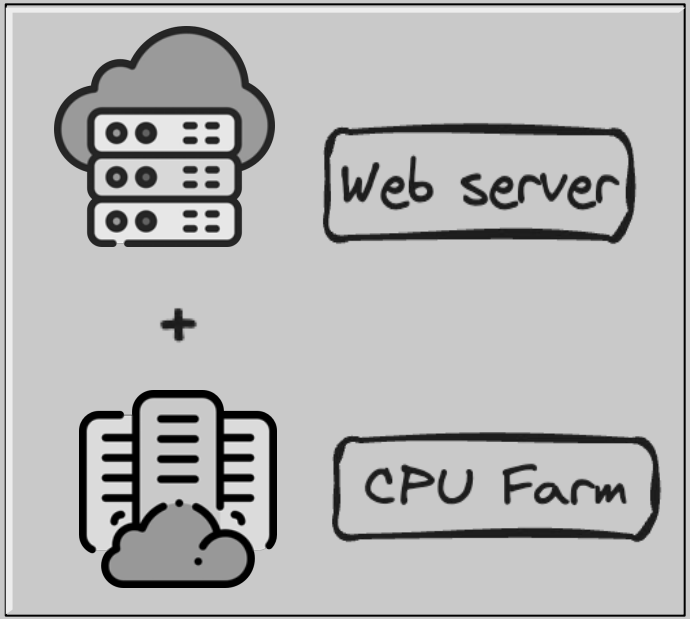
- local server which effectively makes user's device a part of CCP4 Cloud, proprietary to that user
- used to run interactive desktop applications, and also image processing where image data cannot be placed in the Cloud
- installs out-of-box as a CCP4 package



CCP4 Cloud Architecture

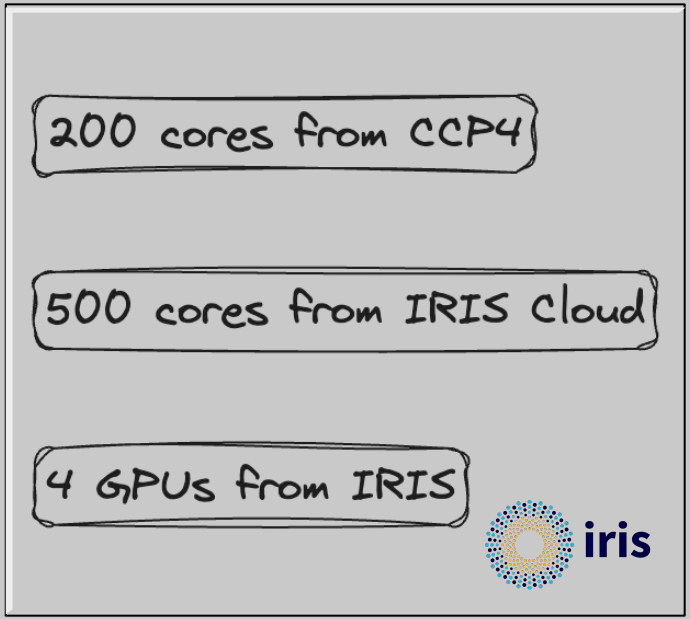
Number Cruncher Nodes:

- only run calculations
- receive jobs from Front End Nodes and send results back to the sender FE
- may be placed on a single or multiple hardware hosts

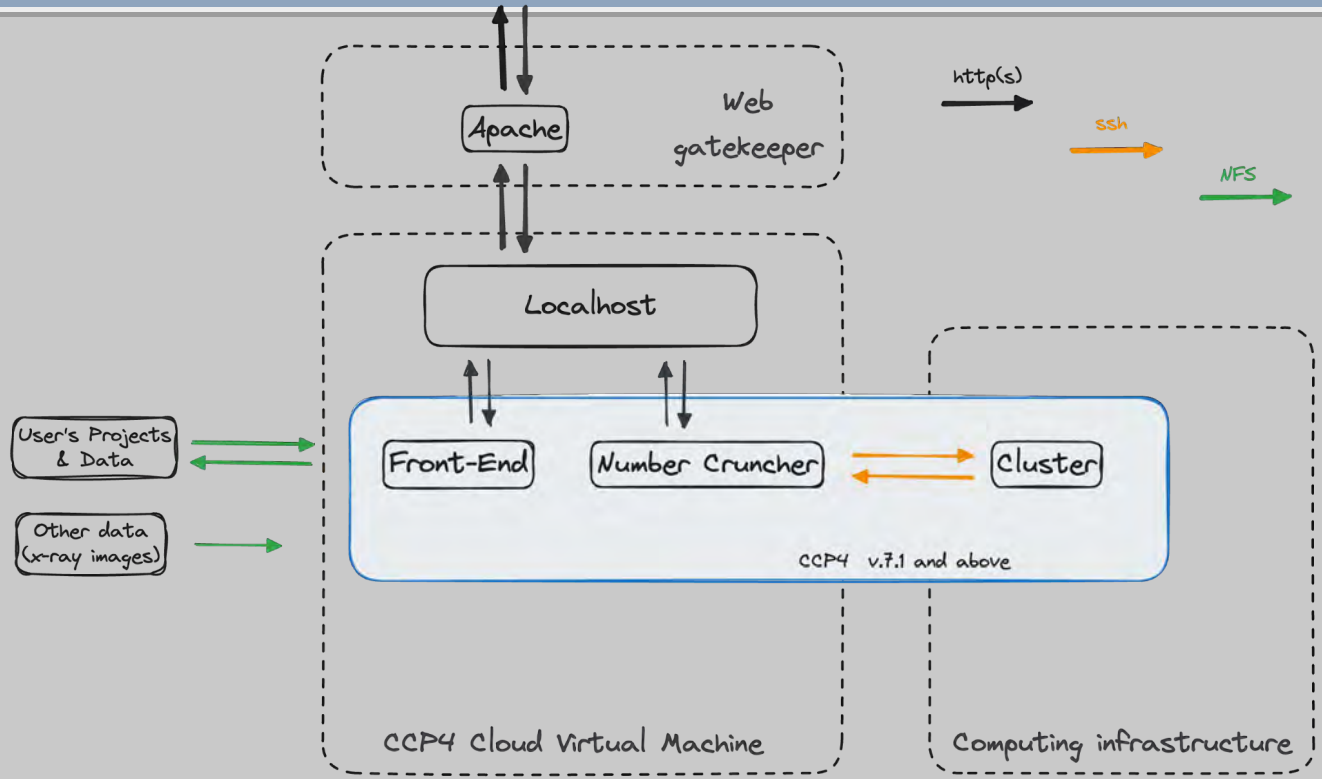


CCP4 Cloud Architecture

- Computing back-end can be a cluster (SLURM, GRID engine, etc) or queue-less system
- NC Web-Server and back-end can be on same or different machine



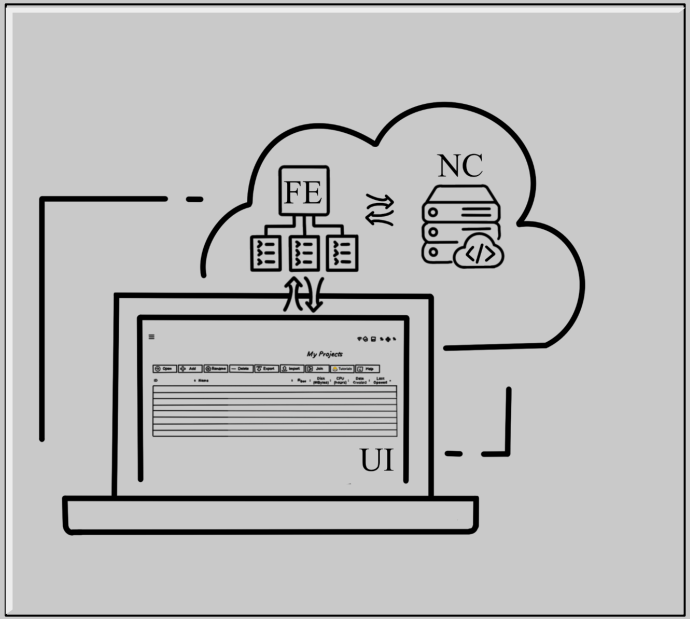
Communication protocols



CCP4 Cloud configurations

CCP4 Cloud configurations

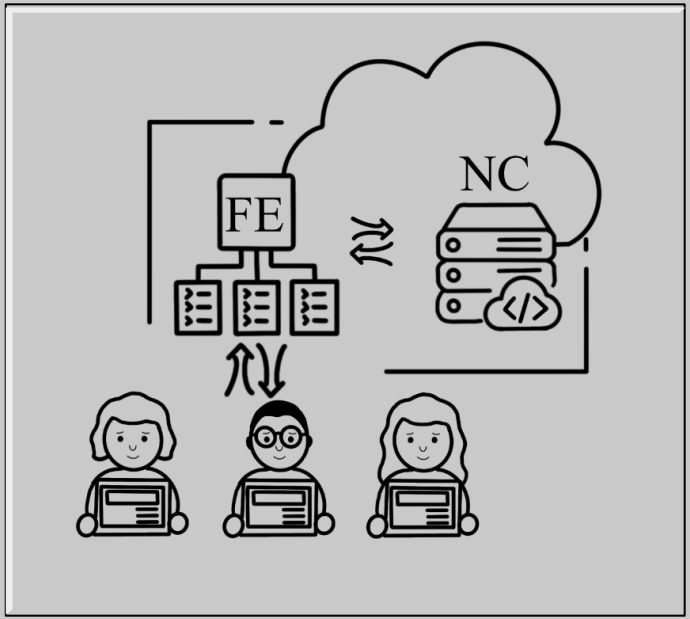
Single-host configuration, suitable for an individual working without need for internet connection



FE - Front End; NC - Number Cruncher; UI - User Interface

CCP4 Cloud configurations

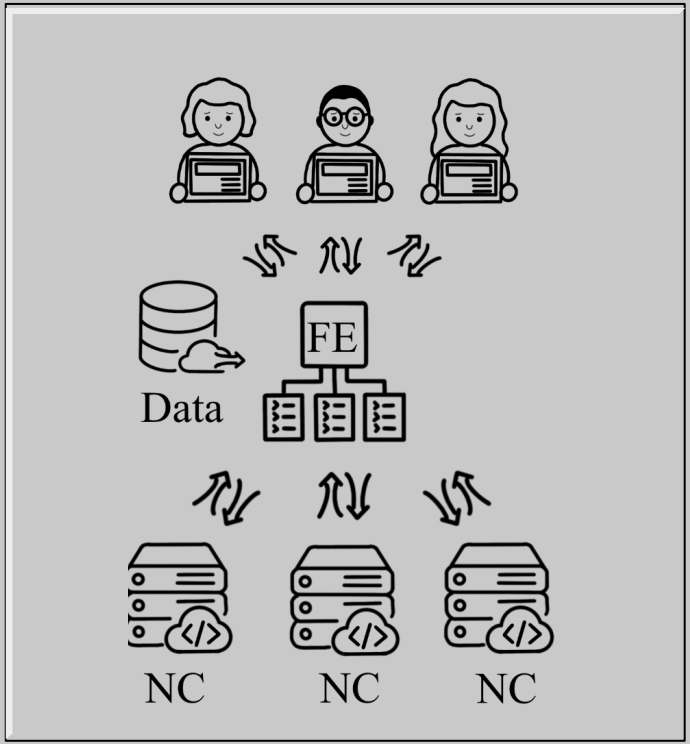
A multi-user setup using a central host machine, suitable for small to medium-sized laboratories.



FE - Front End; NC - Number Cruncher

CCP4 Cloud configurations

Fully distributed, multi-component setup with single point of access, suitable for large facilities and research centres

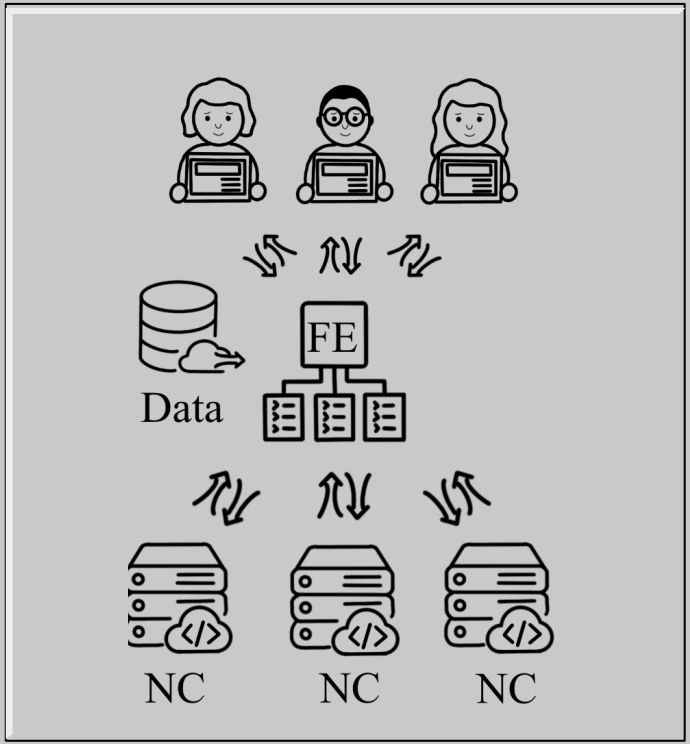


FE - Front End; NC - Number Cruncher

CCP4 Cloud configurations

Fully distributed,
multi-component setup with
single point of access,
suitable for large facilities
and research centres

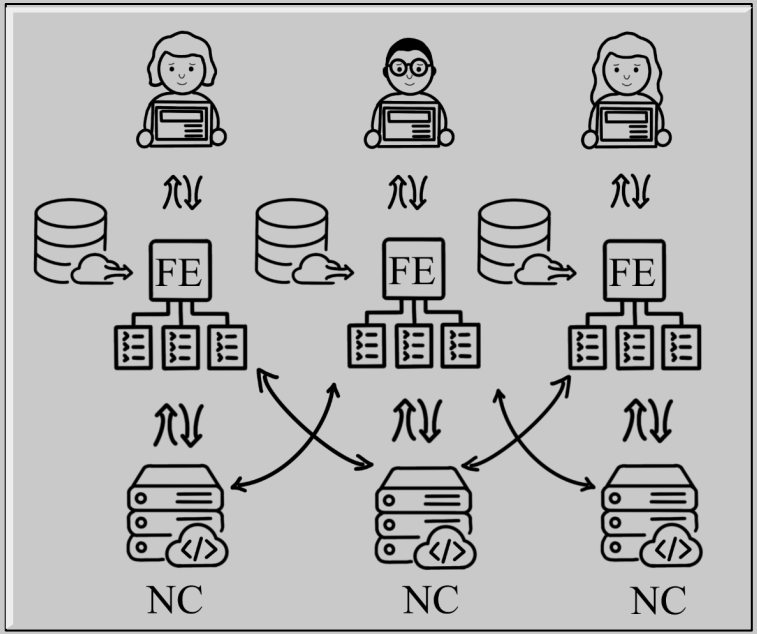
https://cloud.ccp4.ac.uk



FE - Front End; NC - Number Cruncher

CCP4 Cloud configurations

Fully distributed, multi-component setup with multiple points of access, suitable for large facilities and research centres



FE - Front End; NC - Number Cruncher

Implementation details

- All server nodes are based on the Node JS platform
- Browser side: HTML5, WebGL, custom Javascript widget framework based on jQuery, jQuery-UI and React
- Job launching framework: Python
- Job report framework: RVAPI (dynamic web content) from CCP4
- Job workflow framework based on abstract task and data models
- Update mechanism
- Script-assisted installation, auto-setup for CCP4 Cloud Client
- No principal restrictions on the number and location of computational nodes

CCP4 Clouds instances

Main CCP4 Cloud instance at CCP4-Harwell from 2018:

- - Over 4,500 user accounts
- - Over 100,000 jobs/year

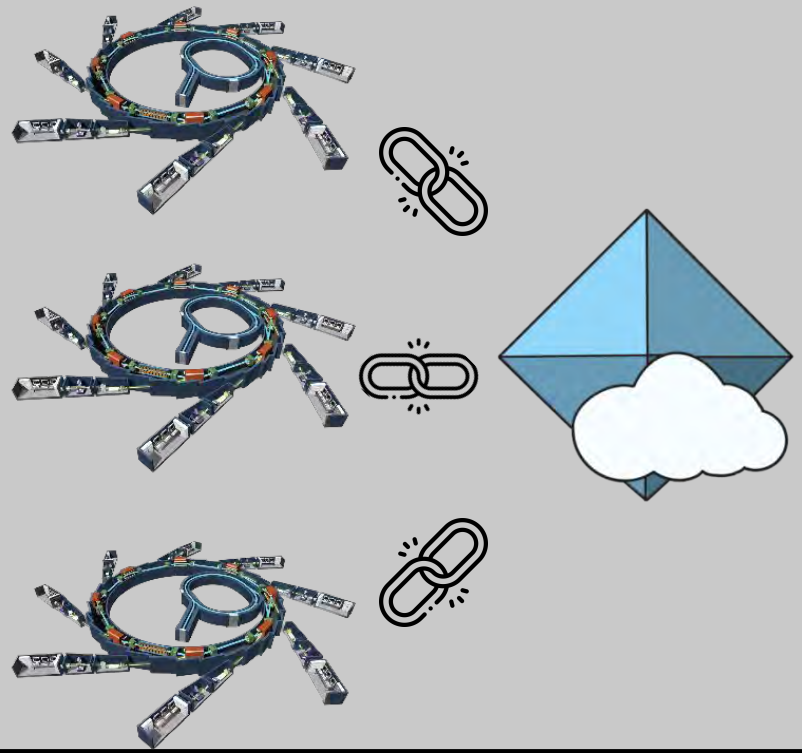
CCP4 Cloud instances at partner sites, including industrial sector:

- EMBL (Hamburg)
- Francis Crick Institute (London)
- Newcastle University
- University of Exeter
- Incyte Inc (Virginia, USA)

Future plans

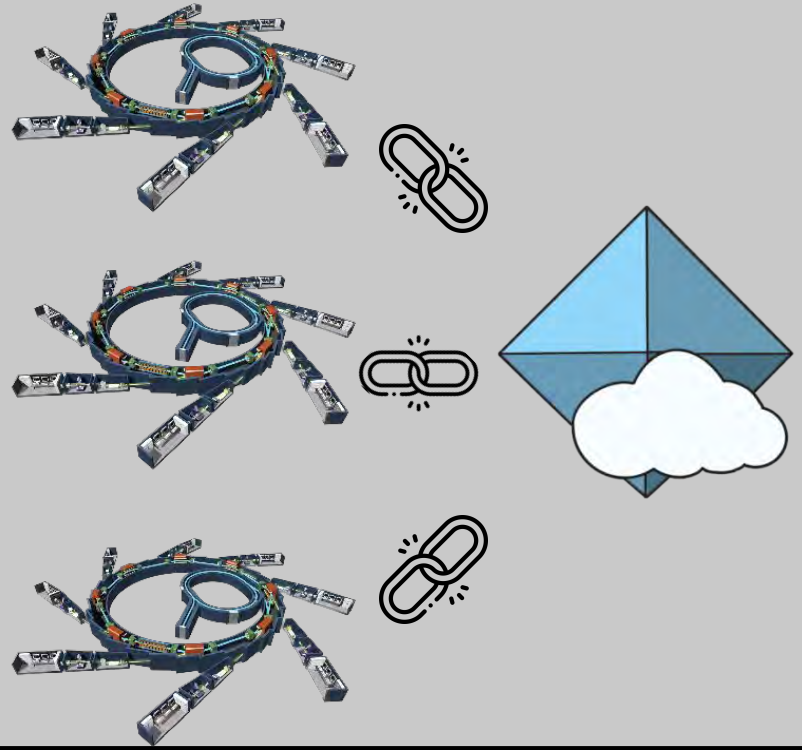
Future plans

Make data links between experimental facilities and in-house X-ray diffractometers and CCP4 Cloud



Future plans

Make data links between experimental facilities and in-house X-ray diffractometers and CCP4 Cloud



ada lovelace centre

Future plans

CCP4 Cloud Archive:

- Develop
- Maintain
- Popularise

Conclusions

CCP4 Cloud

Mitigates software complexity

- ➔ Supporting wide variety of computing platforms is difficult
- ➔ Full installation with 3rd party databases and software is difficult

Meets methods and software demands

- ➔ Modern automatic methods require more CPU and memory than most local setups can afford

Facilitates data logistics and distributed team working

- ➔ Growing volumes of data from modern sources are difficult to handle locally
- ➔ File exchange in distributed collaborations is usually a mess

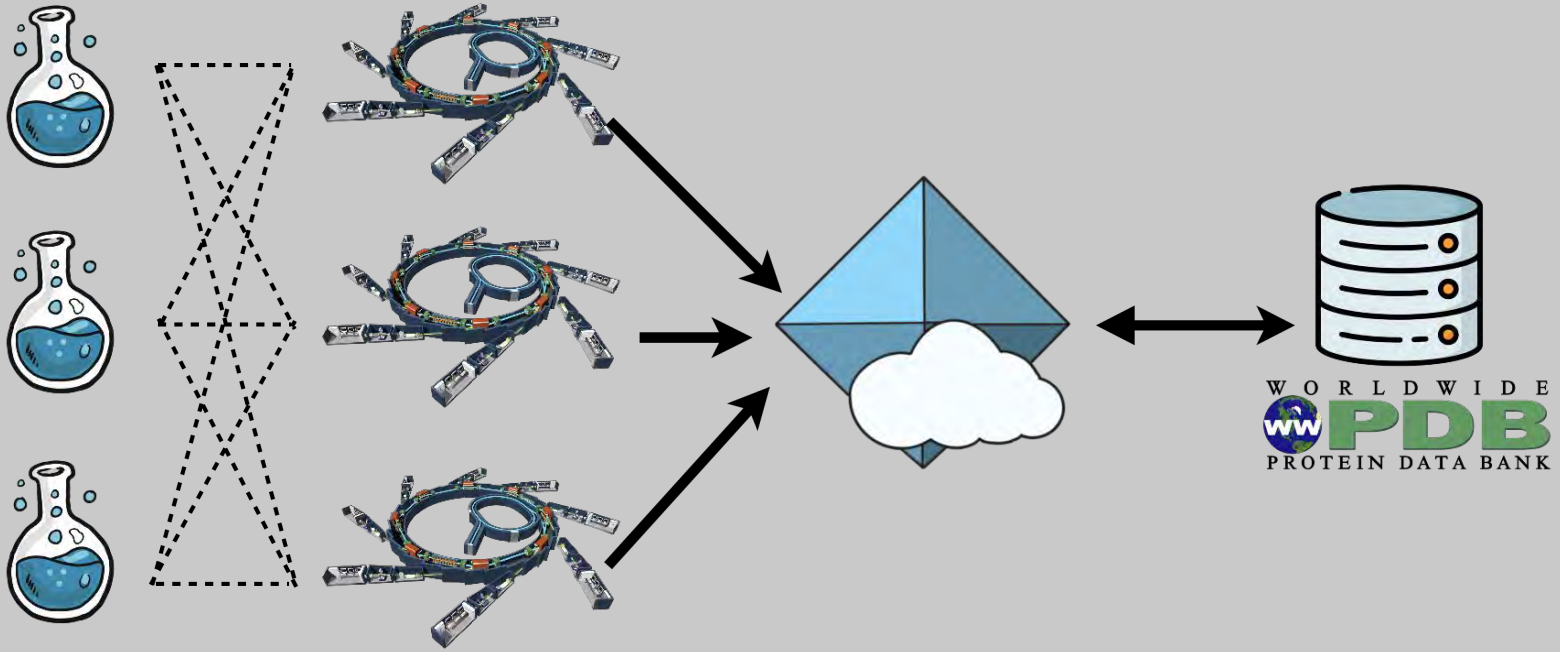
Provides for data security and retention

- ➔ Increasingly more difficult to distribute for modern systems and corporate environments
- ➔ Cloud solutions are safer, getting preferential in industry
- ➔ The lifetime of data stored in the cloud is considerably longer (effectively infinite) than that usually achieved with locally maintained hardware

CCP4 Cloud's key features

- Software, resources and data as a service: go-and-use
- Cross-platform compatibility: can be used on Windows, Linux, Mac OSX, tablets and smartphones
- Rich project development functionality
- All stages of structures solution: from image processing to PDB deposition
- Integrates access to web-resources such as PDB and AFDB
- Facilitates teamwork by sharing projects in real time with various levels of access
- Can be run locally
- Can be installed in a lab, institute or firm
- Highly configurable and adaptable to using mixed distributed computational resources
- Integrated documentation and tutorials

Data Production in Structural Biology



Research Labs → Imaging Facilities → Structure Solution → Data Bank → Utilisation

Acknowledgments

STFC, CCP4, Harwell, UK:

Eugene Krissinel, Andrey Lebedev, Oleg Kovalevskyi, Ronan Keegan, Charles Ballard, Ville Uski, Jools Wills

Newcastle Uni UK:

Martin Noble

Uni Southampton UK:

Ivo Tews

MRC/LMB, Cambridge, UK:

Robert Nicholls

EMBL-EBI, Hinxton, UK:

John Berrisford

Uni Leiden, The Netherlands:

Navraj Pannu, Pavol Skubak

Global Phasing Ltd, Cambridge, UK:

Marcin Wojdyr, Clemens Vonrhein

Uni York, UK:

Stuart McNicholas, Filomeno Sanchez Rodriguez, Paul Bond

Uni Liverpool, UK:

Adam Simpkin, Jens Thomas

Uni Birmingham, UK:

Christopher Oliver

CCP4, STFC & RCaH

CCP4 Collaboration, CCP4 developers

CCP4 Cloud users
Worldwide

CCP4 School hosts

Ed Lowe
Oxford University

Andy Purkiss
Francis Crick Institute, London

Grzegorz Chojnowski
EMBL-Hamburg

Arnaud Basle
Newcastle University

Michael Isupov
University of Exeter

Biotechnology and Biological Sciences Research Council *(BBSRC) UK*

UKRI Research Complex at Harwell

iris

CCP4

UKRI Science and Technology Facilities Council

Newcastle University

THE FRANCIS CRICK INSTITUTE

EMBL European Molecular Biology Laboratory

UNIVERSITY OF EXETER

ada lovelace centre

UKRI Biotechnology and Biological Sciences Research Council

Research grant BB/L0070317/1 (2014-2019)

CIUK 2023 Presentations

Richard Gunn (Programme Director, Digital Research Infrastructure, UK Research and Innovation)

Towards a coherent state-of-the-art national digital research infrastructure

Abstract: This session will provide a progress update on the development of UKRI's Digital Research Infrastructure (DRI) programme over 2023 and future plans.

Bio: Richard Gunn leads on delivering UKRI's strategy to develop a coherent state-of-the-art digital research infrastructure through support for data and computing services, software and skilled professionals. Since joining the Research Councils in 2012, Richard has gained extensive experience in developing strategies, policies and interventions to support research and innovation. He has contributed to a wide-range of initiatives relating to research infrastructures and emerging technologies, including ARCHER2, ExCALIBUR, PRACE, the Henry Royce Institute, and the UK National Quantum Technologies programme. Previously, Richard worked in intellectual property law and gained a PhD from Imperial College London at the chemistry-biology interface.

UK Research and Innovation

Towards a coherent state-of-the-art national digital research infrastructure

Richard Gunn, DRI Programme Director

CIUK, Manchester Central
7 December 2023

UK Research and Innovation

What I'm going to cover

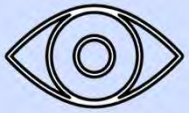
Since the first DRI Congress in March 2023, we have:

- Been working closely with DSIT on delivering the recommendations of the Future of Compute review.
- Developed our programmatic and governance arrangements, including establishing our independent expert advisory group.
- Developed a portfolio of investments for the second ‘ramp up’ phase of the DRI programme.

And there will be much more to come in 2024!

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

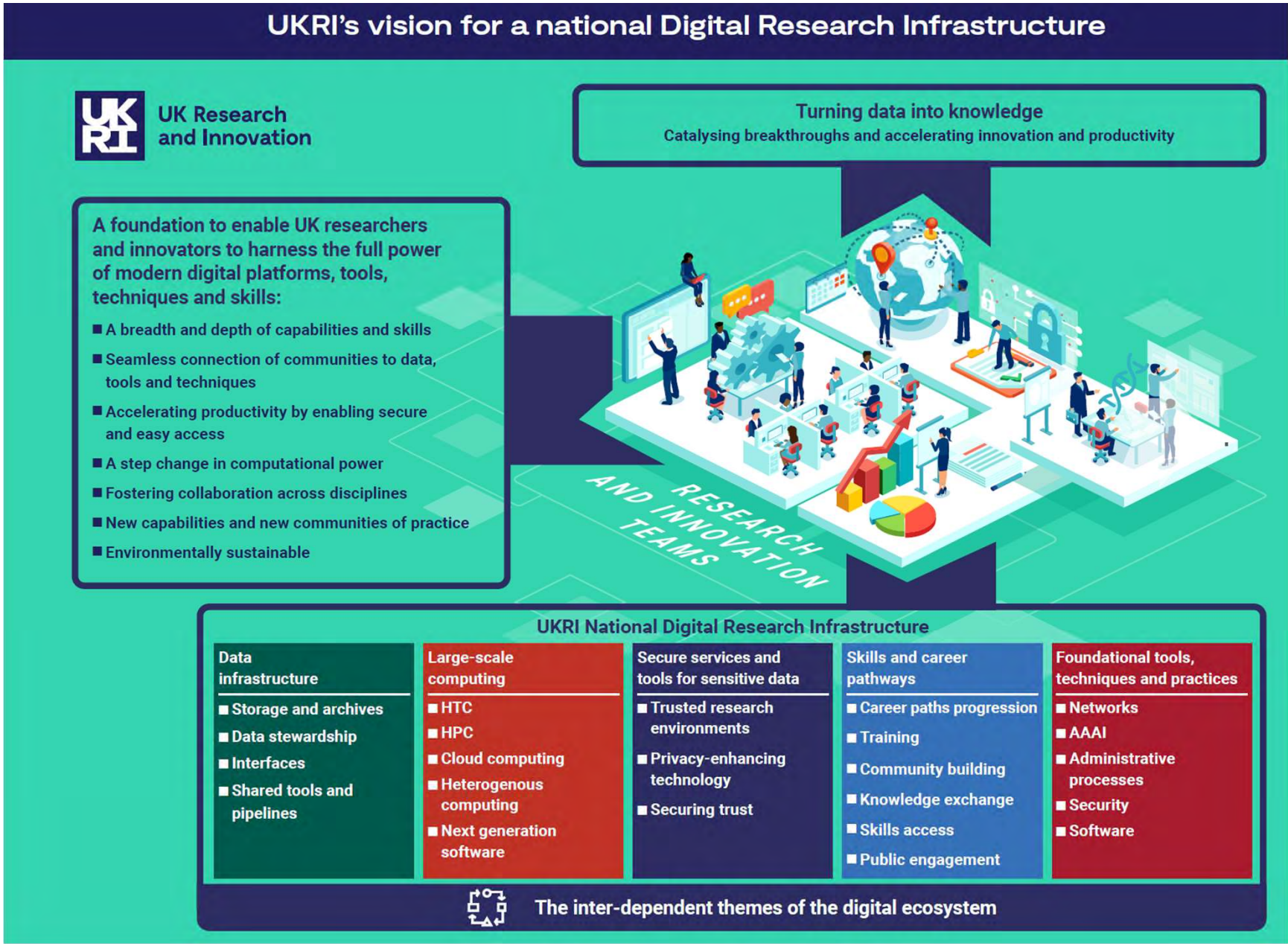
DRI is a system that includes large-scale computing (LSC), data storage, facilities, software, networks, skilled DRI professionals, and other components.
Working with Councils, we will achieve our vision by evolving existing infrastructures to support new communities of practice and, subject to funding,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

UK Research and Innovation

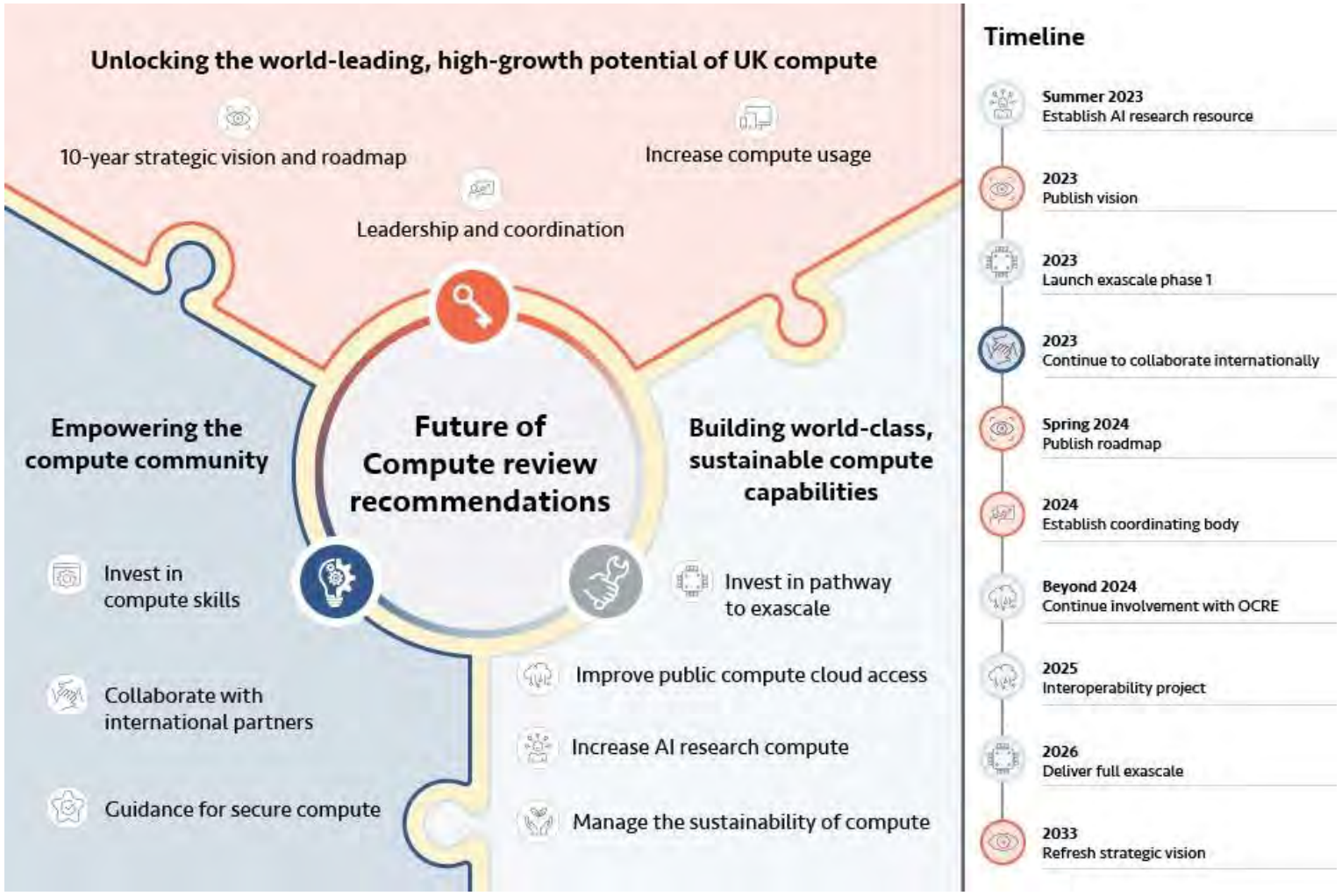
Digital Research Infrastructure

- Long-term planning
- Driven by community requirements
- Environmental sustainability
- Partnership with government and industry
- Five cross-cutting themes

UK Research and Innovation



Future of Compute Review



UK Research and Innovation

Delivering the Future of Compute Review Recommendations

- Our priority is to provide appropriate and ambitious compute capabilities reaching out towards exascale and large AI enabled supercomputing investments for UKRI's diverse research and innovation communities.
- We are working closely with DSIT to support £300m investment in the AIRR programme and deliver the governments ambition for AI.
- We are taking a phased approach towards Exascale.

UK Research and Innovation

SBU0

DRI Projects to Date – First Pilot Phase

Scoping studies to assess the large-scale computing, data infrastructure and software needs across the breadth of UKRI's communities

UKRI **Net Zero Digital Research Infrastructure** Scoping Project

Ada Lovelace Centre, an integrated, cross-disciplinary data intensive science centre.

The launch of the UK trusted and connected **Data and Analytics Research Environments** programme (DARE UK)

Digital Research Infrastructure retreat: a five-day event to help technology specialists to develop additional professional skills

UK Research and Innovation

You can find out more on the UKRI Website

Slide 7

SBU0 [@Richard Gunn - STFC UKRI] edited this slide - original wording in the speaker notes. Can revert back if you'd prefer
Stephanie Bonehill - UKRI, 2023-12-05T16:10:32.610

UKRI's Digital Research Infrastructure Programme

A phased approach…

UKRI secured £129 million for digital infrastructure (profile rising to £70m in year three) of this Spending review (FY22/23 – FY24/25)

UKRI DRI Phase 1 (2021-23) – £34 million invested in:

- A portfolio of interventions to enhance our existing digital infrastructures
- Investments in priority areas including Net Zero and Trusted Research Environments
- Scoping activities to assess data and computing requirements in more detail

UKRI DRI Phase 2 (2023 onwards) :

- Established the Advisory Group for DRI (AGD)
- AGD has made a recommendation for a portfolio of 37 projects.
- Projects will take many forms and will cover all themes of the strategy.

UK Research and Innovation

UKRI Support for Infrastructure

Research Councils
Prioritise, endorse and submit investment bids

Facilities Strategic Advisory Group (FSAG)
Prioritise, endorse and submit UKRI multidisciplinary facilities investment bids

Infrastructure Working Group
Cross-council forum, ideas sharing, collaboration, co-design processes, discuss strategic topics

IAC
Assess bids, recommend portfolio (s)

ExCo
Decide portfolio

Board
Approve portfolio

Formal commission for input from Councils & key stakeholders

Input from DRI Programme Advisory Team on the development of its programme. Advisors are ordinarily drawn from existing UKRI advisory structures.

Digital Research Infrastructure Committee (DRIC)
Cross-Council committee that develops the plan to deliver the UKRI DRI strategy and portfolio options.

AGD
Recommend portfolio options & advise on risks & opportunities

Digital Research Infrastructure Forum
Senior UKRI and government (DSIT, GO Science, ONS) forum providing strategic advice and scrutiny of DRI investment plans.

UK Research and Innovation

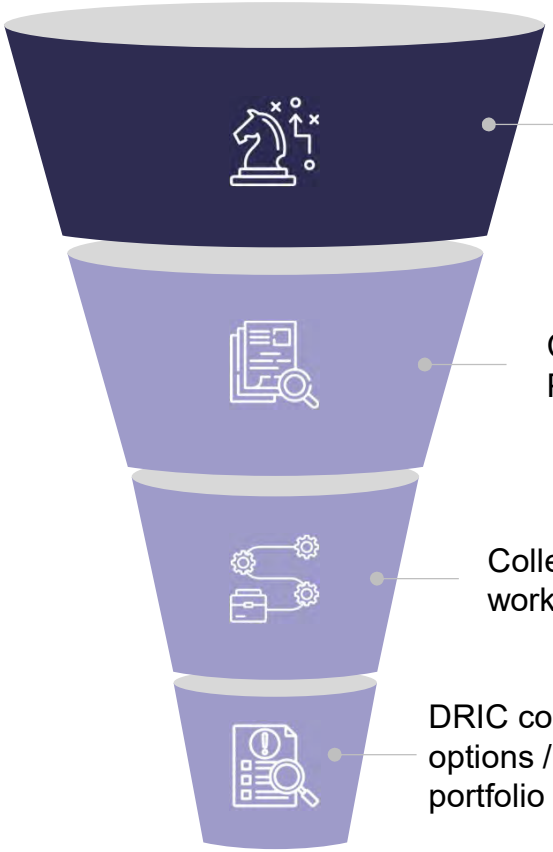
Advisory Group for DRI (AGD): first cohort of members

- David De Roure, University of Oxford
- Christine Orengo, University College London
- Tom Crick, University of Swansea
- Sian John, NCC Group
- James Fleming, The Francis Crick Institute
- Neil Chue Hong, The University of Edinburgh
- Tony Cass, CERN
- Amanda Brock, Open UK

Open recruitment for new members coming soon!

UK Research and Innovation

Phase 2 portfolio Development



- **Council Responses:** Councils have outlined current and planned investments, highlighting their priorities and opportunities for collective investment.
- **Priorities Mapping**: These priorities have been correlated with the DRI Delivery Plan, ensuring alignment and cohesion in the direction of Phase 2.
- **Incorporating Reviews**: Future of Compute review recommendations have been considered, enhancing the effectiveness and relevance of Phase 2 goals.
- **DRIC's Role**: The Digital Research Infrastructure Committee (DRIC) used this extensive evidence to create a series of Strategic Outline Cases under **5 workstreams**
  - DRIC has made progress in developing preferred options and addressing delivery risks, while also assessing and mitigating associated risks.
  - The Exascale, AIRR, and Research Cloud Pilot projects are being implemented under the DRI governance structure with the involvement of DSIT, running concurrently with the aforementioned DRIC activities.

UK Research and Innovation

Priorities for collective investment

Theme	1 Data Infrastructure	2 LSC	3 Sensitive data	4 Skills	5 Foundational tools
Councils' highest priorities for collective investment in Phase 2 (2023-2025) Ranked by Council vote, >3 votes included Priorities don't always map neatly on to strategic themes!	1. Coordination/ overarching network 2. Federation 3. Data standards/ management/ curation 4. Data policy and governance 5. Mapping data infrastructures	1. Access to LSC (e.g., for new communities) 2. Move towards a national system/ scoping future requirements 3. Maximising current investments	1. Priorities for this theme tended to be included in 'Data Infrastructure' more broadly	1. Training 2. Knowledge exchange/ community building 3. Technical 'pools'/ 'ambassadors' to work with researchers 4. Career pathways Note: General agreement with skills working group here	1. Software activities (software as infrastructure, and development of future software)

UK Research and Innovation

Cross-cutting: Net zero, interdisciplinary working

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

Case Studies

Federated Data Services

Expanded support for 6 data infrastructures, including:

iDAH – a national infrastructure for digital innovation and curation for arts and humanities - £3.3M

BioFAIR: A BioCommons infrastructure for UK life science researchers - £10.8M

National Computational Research Services

Transitioning UKRI's Council-focused compute services to 'UKRI National Computational Research Services' to unify hardware, operations, and service provision to serve all of UKRI's communities of practice - £21M.

Enabling Software, Networks, Security and net zero

Establishing authentication, authorization and accounting Infrastructure (AAAI) - £5.9M

Building a cybersecurity community - £0.9M

UK Research and Innovation

Case Studies Continued

Calls for funding & uplifts

Support for DRI Professional workstream – in 2024 there will be an open call for networking activities of digital Research Technical Professionals - £10M over 4 years
Enabling Software, Networks, Security and net zero workstream – in 2024 there will be an open call for a Net Zero DRI Coordinator, who will build connectivity across disciplines and coordinate the development of an application to a future funding opportunity supporting a net zero DRI network - £3M over 4 years.

Funding for software

Software for large-scale compute: building on ExCALIBUR, funding for knowledge exchange, RTP hubs, pathfinder projects, and code porting - £22M.
Enabling Software, Networks, Security and net zero workstream: Piloting approaches for funding software development and maintenance £5.7M.

Federated Data Services

Four cross-cutting projects to reduce silos and explore opportunities for improved data access, data discovery, and federation - £16.1M

UK Research and Innovation

More to come in 2024!

- We will engage across a diverse range of current and potential user communities across UKRI's communities of practice on future compute requirements.
- We will explore the development of a 'statement of requirements' process to contribute ideas for the development of our DRI and recruit a further cohort of AGD members.
- We will develop new activities to expand our support for software and skills.
- We will increase the visibility of UK involvement in relevant international initiatives and organisations.

UK Research and Innovation

UK Research and Innovation

Thank you

@UKRI_news

UK Research and Innovation

UK Research and Innovation

How do we foster a productive compute ecosystem for the benefit of all users, existing and potential?

Sadaf Alam, University of Bristol
Chris Coates, Logicalis
Victoria Moody, Jisc
Tobias Weinzierl, Durham University

UK Research and Innovation

CIUK 2023 Presentations

Martyn Guest (ARCCA, Cardiff University)

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

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

The variety of systems considered focus on both the Intel Sapphire Rapids and AMD EPYC Genoa family of processors. Using the Intel Skylake Gold 6148 and AMD EPYC Rome 7502 as the baselines, an assessment is made across a variety of Sapphire Rapids (8480, 8490) and Ice Lake (8358, 8352Y, 8368Q, 8360Y and 8380) SKUs, with system interconnects from both NVIDIA Networks and Cornelis Networks. Attention is also focused on systems featuring the AMD Genoa and Milan EPYC processors. The former feature the Genoa 32-core 9354 and 48-core 9454 SKUs, the latter the 64-core Milan 7713, 7763 & 7773X and 32-core Milan 7543 & 7573X.

The benefits of the Intel® oneAPI Toolkit and the SPACK Package Manager for HPC 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.

Martyn’s research interests cover the development and application of computational chemistry methods. He is lead author of the GAMESS-UK electronic structure program and has written or contributed to more than 260 journal articles.

COMPUTING INSIGHT UK 2023

"Productive Supercomputing"

7 - 8 DECEMBER 2023

Manchester Central, UK

www.ukri.org/CIUK

UKRI Science and Technology Facilities Council

Performance of Community Codes on Multi-core Processors

An Analysis of Computational Chemistry and Ocean Modelling Applications

Martyn Guest, Jose Munoz Criollo & Thomas Green

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 feature at Daresbury's MEW and successor CIUK conferences.
- Focus on systems featuring **processors from Intel** (Sapphire Rapids & Ice Lake SKUs) and **AMD** (EPYC Genoa & Milan 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.
  - **Two** Intel Sapphire Rapids clusters – the 56-core Platinum 8480 (2.0 GHz) and Platinum HBM 9480 (1.9 GHz).
  - **Five** Intel Xeon Ice Lake clusters, the 32-core Platinum **8358** (2.6 GHz) and **8352Y** (2.2 GHz), the 40-core **8380** (2.3 GHz), 38-core **8368Q** (2.6 GHz), 36-core **8360Y** (2.4GHz) plus other Cascade Lake & Cascade Lake-AP systems.

Introduction and Overview

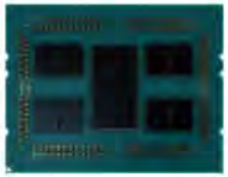
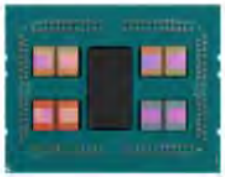
- **Four** AMD EPYC **Milan clusters** featuring the 64-core **7713** (2.0 GHz) and **7773X** (2.2 GHz) and the 32-core **7543** (2.8 GHz) and **7573X** (2.8 GHz).
- **Two** AMD **Genoa clusters** featuring the 32-core **9354** (3.25 GHz) and 48-core **9454** (2.85 GHz) SKUs.

- Consider performance of both synthetic and **end-user applications:**
  - Molecular simulation (**DL_POLY, LAMMPS, AMBER & GROMACS MD codes**);
  - Materials modelling (**VASP, CASTEP**) & electronic structure (**GAMESS-UK**);
  - Ocean modelling codes including **NEMO** and **FVCOM**.
- 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 the 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**.
   - Problems encountered on **AMD Milan** systems. Working code with Intel 2019/5 on AMD Rome systems failed on Milan, with codes hanging at arbitrary core counts. **Intel oneapi resolved many of these** issues.
   - **Performance issues remain** compared to earlier variants of Intel Parallel Studio XE. e.g., a major decline in both VASP and CASTEP performance on AMD EPYC when moving from “mpi/intel/2018/2” to “mpi/intel/2020/2”
5. Consistency through use of **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

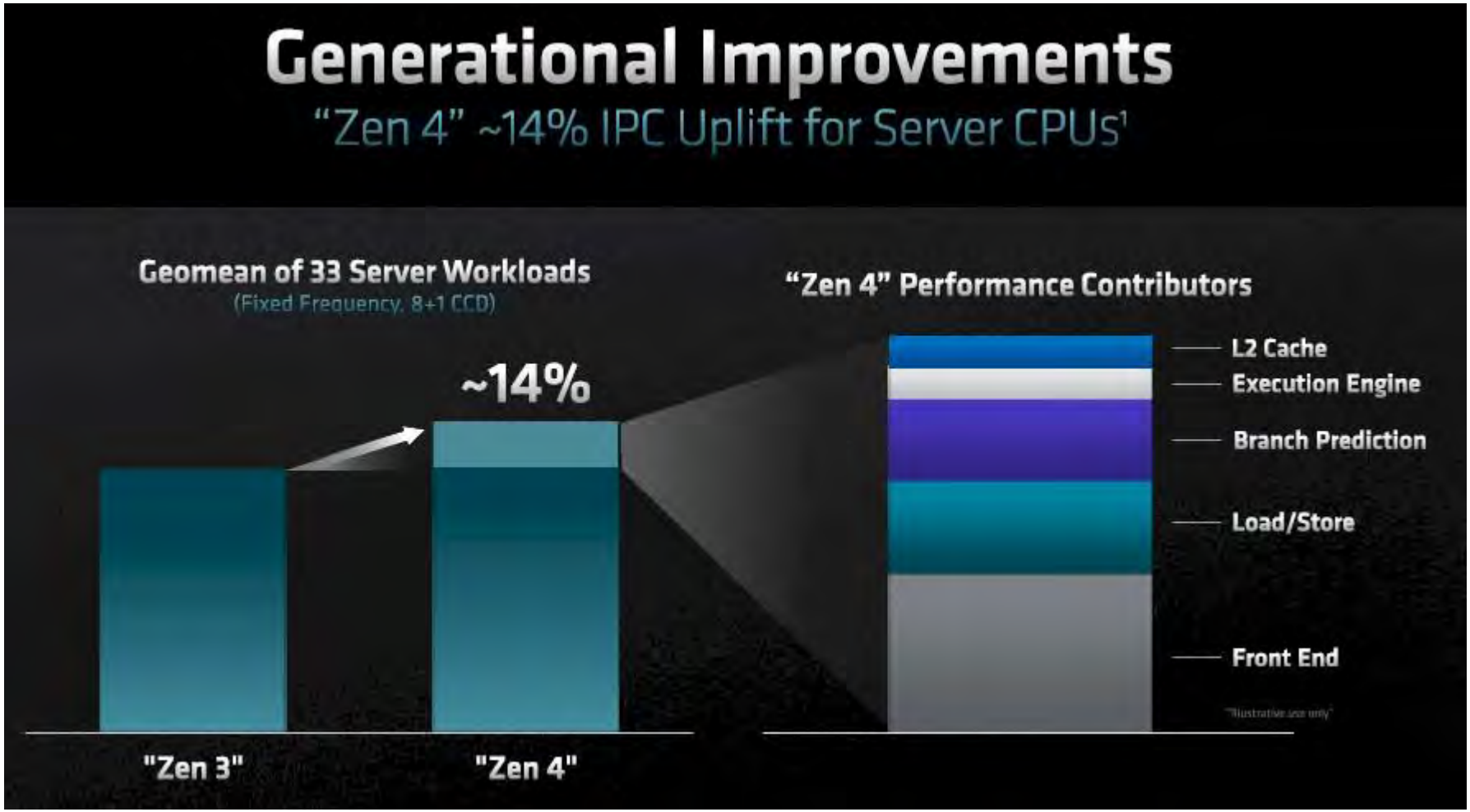
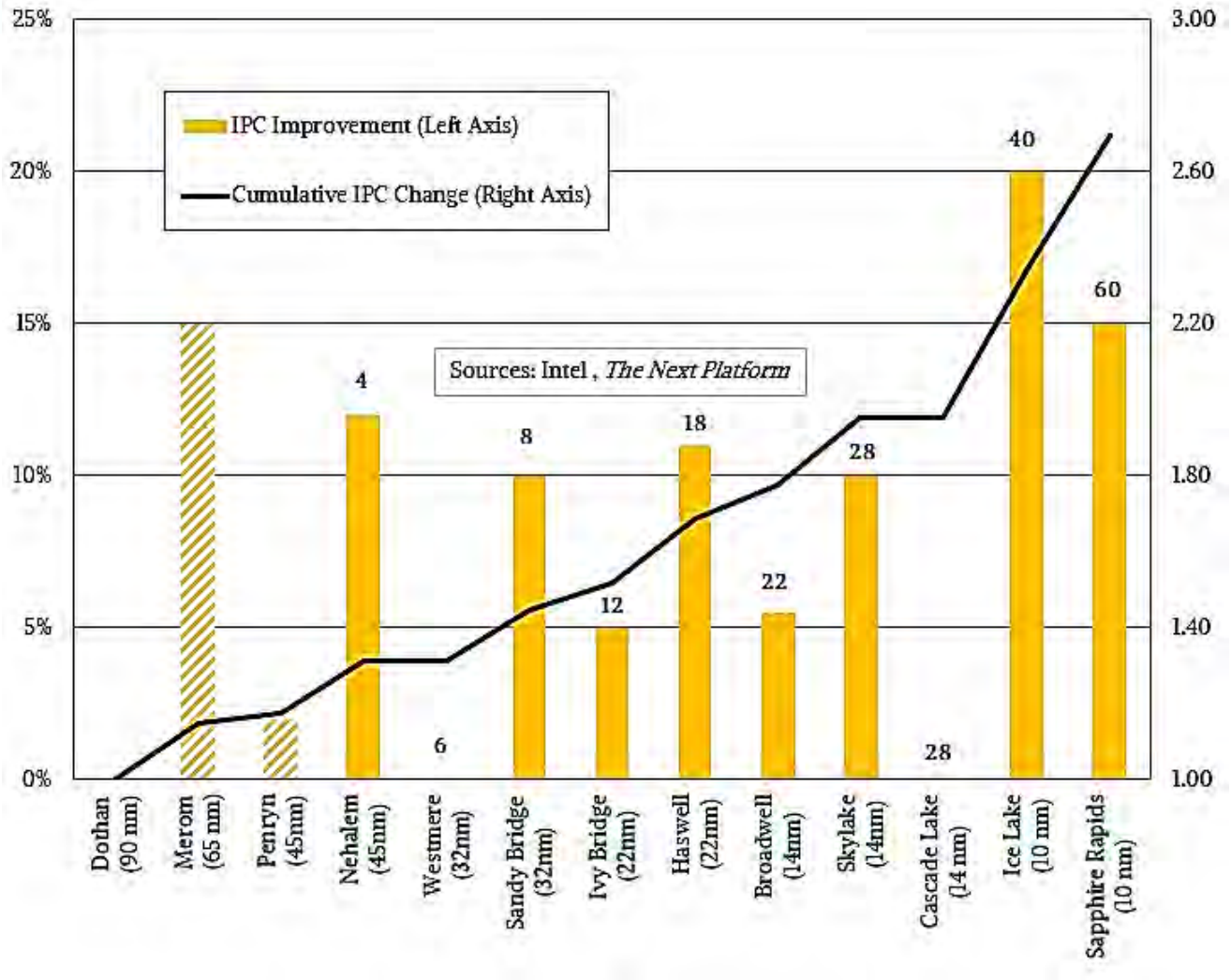


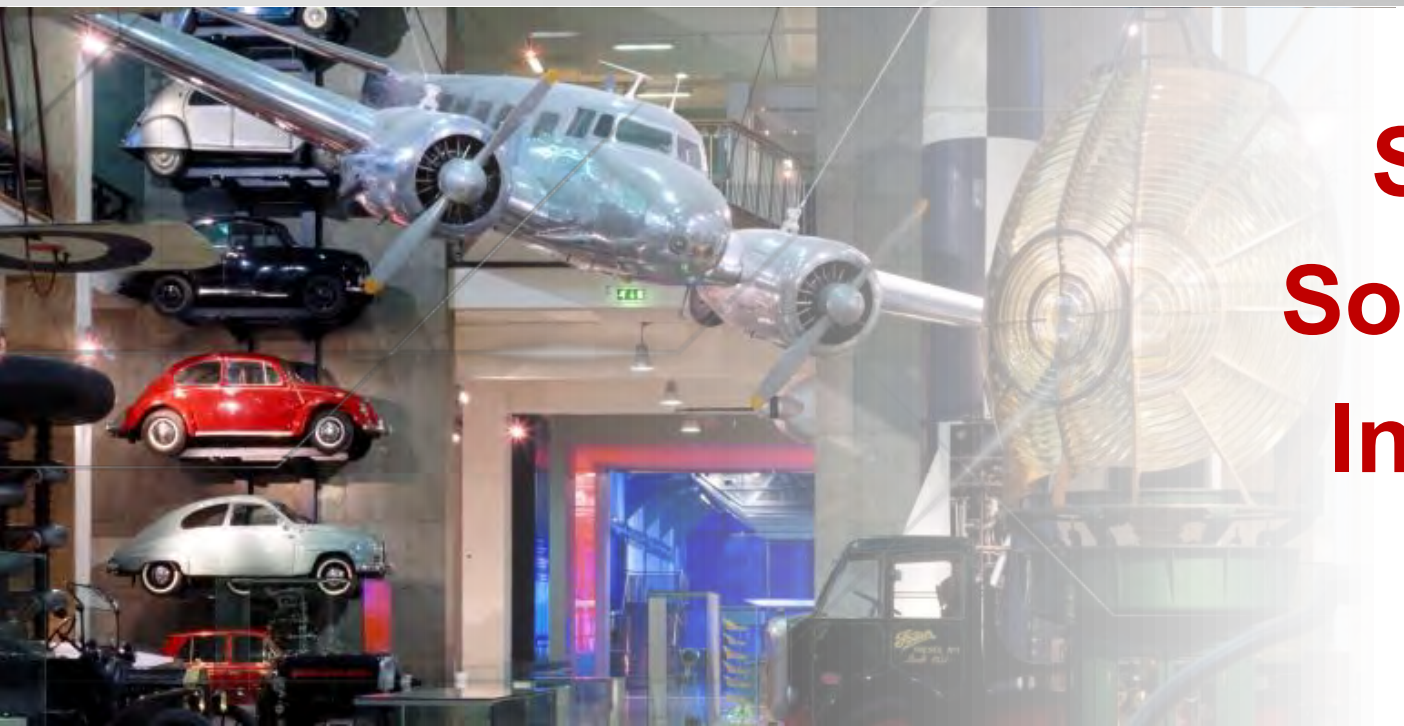
Figure. The move to Genoa is a big leap in performance, starting with the move to the "Zen 4" cores, which are providing a 14 percent increasing in the instructions per clock (IPC) compared to the prior "Zen3" cores used in the Milan Epyc 7003s..

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.

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 (no decommissioned)

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

Intel Xeon Ice Lake Clusters

Cluster / Configuration
Dell Zenith cluster at the Dell Technologies HPC & AI Innovation Lab – Intel Xeon sub-systems with Mellanox HDR interconnect fabric running Slurm
• 50 nodes × Intel **Xeon Platinum 8358 Processor / 2.60 GHz**; # of CPU Cores: **32**; # of Threads: 64; Max Turbo Frequency: 3.40 GHz Base Clock: **2.60 GHz;** Cache 48 MB**;** Default TDP / TDP: 250W; **Mellanox HDR 200Gb/s**
• 70 nodes × Intel **Xeon Platinum 8352Y Processor / 2.20 GHz**; # of CPU Cores: **32**; # of Threads: 64; Max Turbo Frequency: 3.40 GHz Base Clock: **2.20 GHz;** Cache 48 MB**;** Default TDP / TDP: 205W; **Mellanox HDR 200Gb/s**
Ice Lake clusters at Intel's OpenHPC Laboratory with Cornelis OPE fabric running Bright release 8.1 and optane filesystem.
• 4 nodes × Intel **Xeon Platinum 8368Q Processor / 2.60 GHz**; # of CPU Cores: **38**; # of Threads: 76; Max Turbo Frequency: 3.70 GHz Base Clock: **2.60 GHz;** Cache 57 MB**;** Default TDP / TDP: 270W; **Cornelis OPE**
• 4 nodes × Intel **Xeon Platinum 8360Y Processor / 2.40 GHz**; # of CPU Cores: **36**; # of Threads: 72; Max Turbo Frequency: 3.50 GHz Base Clock: **2.40 GHz;** Cache 54 MB**;** Default TDP / TDP: 270W; **Cornelis OPE**
Intel's Endeavour cluster with Cornelis OPE fabric running Slurm
• 8 nodes × Intel **Xeon Platinum 8380 Processor / 2.30 GHz**; # of CPU Cores: **40**; # of Threads: 80;
• 10 nodes × Intel **Xeon Platinum 8360Y Processor / 2.40 GHz**; # of CPU Cores: **36**; # of Threads: 72

Intel Xeon Sapphire 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 8480 Processor / 2.00 GHz**; # of CPU Cores: **56**; # of Threads: 112; Max Turbo Frequency: 3.80 GHz Base Clock: **2.00 GHz;** Cache **105 MB;** Default TDP / TDP: 350W; **DDR5 4800 MT/s**; **Mellanox NDR 400Gb/s** • The 8480 systems are connected to NDR InfiniBand, configured in a fat tree, with each rack of nodes generally using a single edge switch.
Intel's Endeavour cluster with Mellanox HDR and Cornelis OPE interconnect fabrics running Slurm
• 150 nodes × Intel **Xeon Platinum 8480 Processor / 2.00 GHz**; # of CPU Cores: **56**; # of Threads: 112; Max Turbo Frequency: 3.80 GHz Base Clock: **2.00 GHz;** Cache **105 MB;** Default TDP / TDP: 350W; **DDR5 4800 MT/s; Mellanox HDR 200Gb/s; Cornelis OPE**
• 73 nodes × Intel **Xeon Platinum 9480 Processor / 1.90 GHz**; # of CPU Cores: **56**; # of Threads: 112; Max Turbo Frequency: 3.50 GHz Base Clock: **1.90 GHz;** Cache **112.5 MB;** Default TDP / TDP: 350W; **DDR5 4800 MT/s; [Maximum High Bandwidth Memory (HBM): 64 GB]; Mellanox HDR 200Gb/s; Cornelis OPE**

AMD EPYC Milan Clusters

Cluster / Configuration
Dell Minerva cluster at the Dell Technologies HPC & AI Innovation Lab – AMD EPYC Rome and Milan sub-systems with Mellanox HDR interconnect fabric running Slurm
• **4 nodes × AMD EPYC Milan 7543 / 2.80 GHz**; # of CPU Cores: 32; # of Threads: 64; Max Boost Clock: 3.7 GHz Base Clock: **2.80 GHz; L3** Cache 256 MB**;** Default TDP / TDP: 225W; Mellanox HDR-100 **200Gb/s**
• **6 nodes × AMD EPYC Milan 7573X / 2.80 GHz**; # of CPU Cores: 32; # of Threads: 64; Max Boost Clock: 3.6 GHz Base Clock: **2.80 GHz; L3 Cache 768 MB;** Default TDP / TDP: 280W; Mellanox HDR-100 **200Gb/s**
• **170 nodes × AMD EPYC Milan 7713 / 2.00 GHz**; # of CPU Cores: 64; # of Threads: 128; Max Boost Clock: 3.675 GHz Base Clock: **2.00 GHz; L3** Cache 256 MB**;** Default TDP / TDP: 225W; Mellanox HDR-100 **200Gb/s**
• **4 nodes × AMD EPYC Milan 7763 / 2.45 GHz**; # of CPU Cores: 64; # of Threads: 128; Max Boost Clock: 3.5 GHz Base Clock: **2.45 GHz; L3** Cache 256 MB**;** Default TDP / TDP: 280W; Mellanox HDR-100 **200Gb/s**
SPARTAN cluster at the Atos HPC, AI & QLM Benchmarking Centre – AMD EPYC Rome system with Mellanox ConnectX-6 HDR100 interconnect fabric
• **240 × AMD EPYC Rome 7742 / 2.25 GHz;** # of CPU Cores: 64; # of Threads: 128; Max Boost Clock: 3.35 GHz Base Clock**: 2.25 GHz; L3** Cache 256 MB; Default TDP / TDP: 225W; **Mellanox ConnectX-6 HDR 100 InfiniBand:** Memory: 256GB DDR4 2677MHz RDIMMs per node: **DDN lustre 7990 Storage, NFS**

AMD EPYC Genoa Clusters

Cluster / Configuration
Dell Minerva cluster at the Dell Technologies HPC & AI Innovation Lab – AMD Genoa sub-system with Mellanox NDR interconnect fabric running Slurm
• **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** • The 9354 systems are connected to NDR InfiniBand configured on a single switch.
AMD Genoa cluster 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 NDR 400Gb/s.**
• **63 'standard' compute nodes**, 384 GB RAM, 1x NDR200 Dual Port IB HCA: **10 'high mem' compute nodes, 1536 GB RAM**, 1x NDR200 Dual Port IB HCA; 4 'GPU' compute nodes, 2x AMD 9454 48C 2.75GHz CPUs, 768 GB RAM, 8x NVIDIA A100 80GB PCIe Gen4 Passive GPU, 1x NDR200 Dual Port IB HCA. Spectrum Scale (GPFS). SLURM 23.02.4.

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)***

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#*

The Performance Benchmarks

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

DL_POLY, LAMMPS, AMBER & GROMACS (MD)

VASP and CASTEP (ab initio Materials properties)

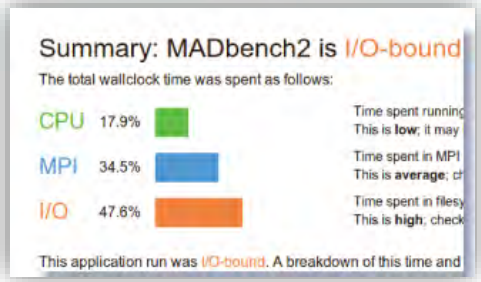
GAMESS-UK (molecular electronic structure)

FVCOM and NEMO (ocean modelling codes)

- 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.

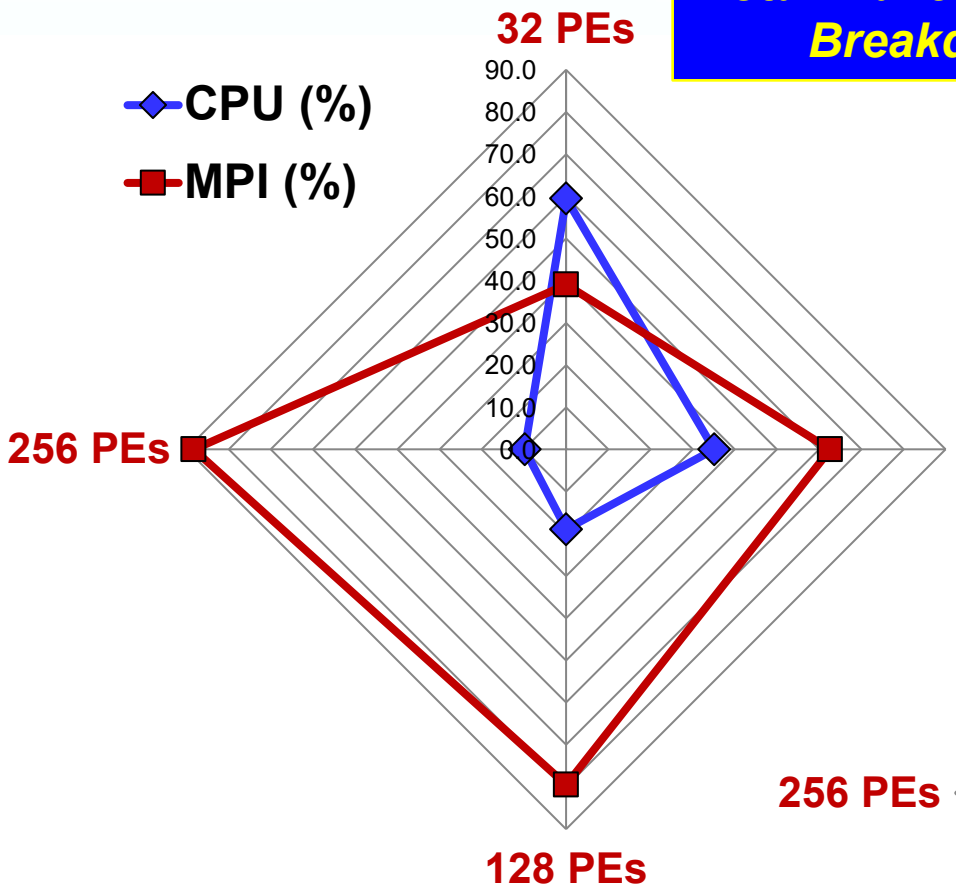
  perf-report mpirun $code

- ***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)**

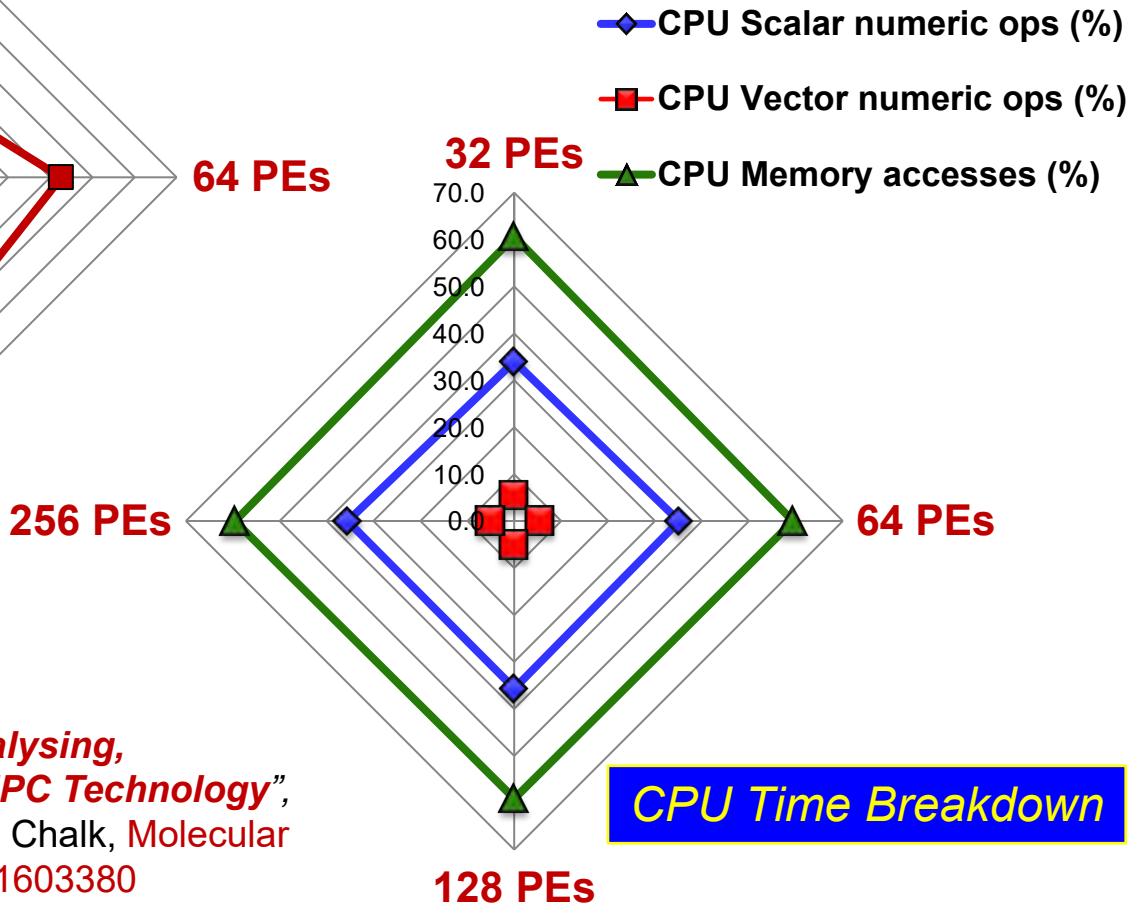
DLPOLY4 – Performance Report

Total Wallclock Time Breakdown

Performance Data (32-256 PEs)



Smooth Particle Mesh Ewald Scheme



CPU Time Breakdown

"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

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
```

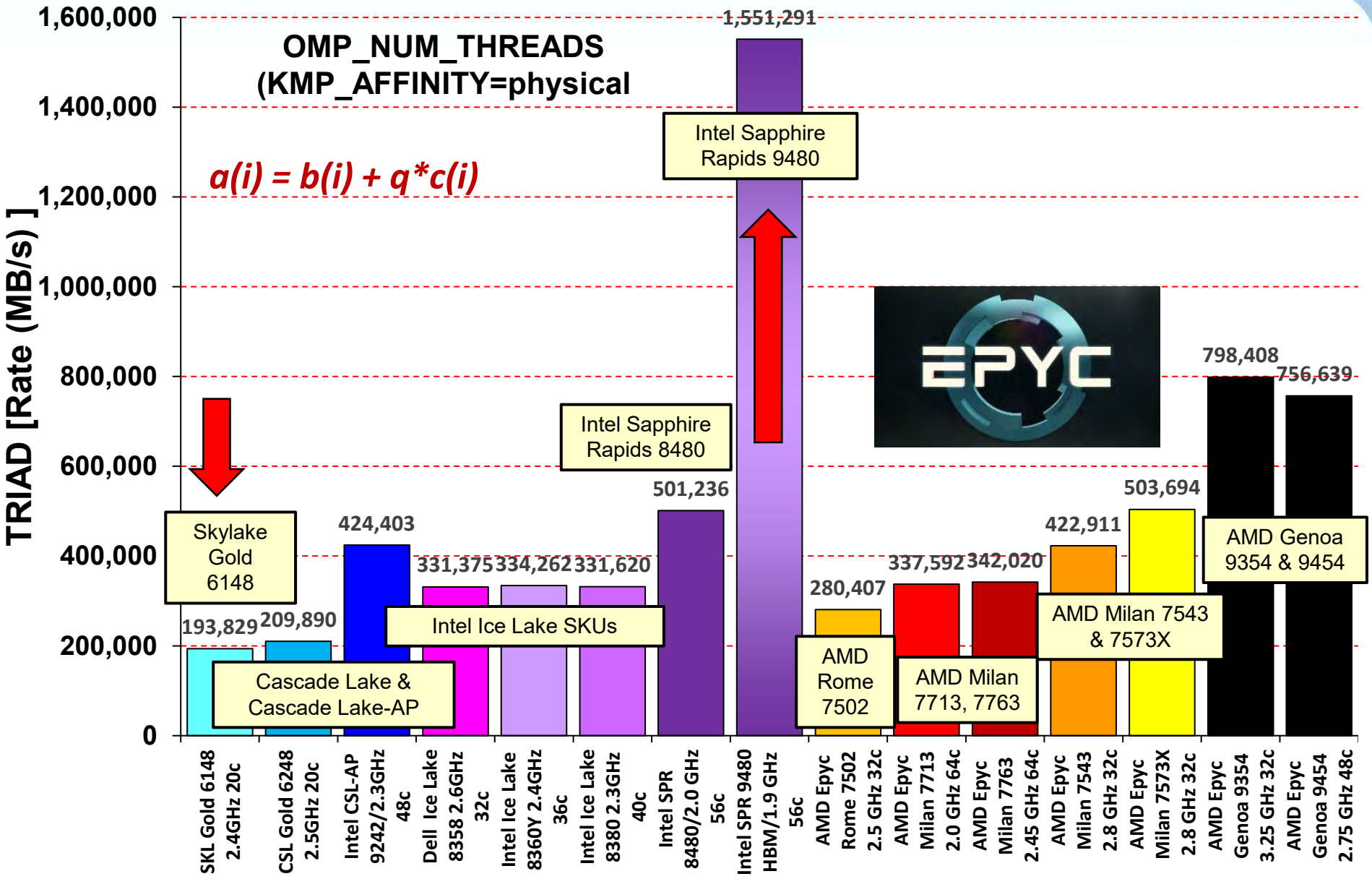
```
# Version of Intel compiler to use and way to source it
source /opt/intel/compilers_and_libraries_2020.2.254/linux/bin/compilervars.sh -
ofi_internal=1 intel64

# Increasing use of oneAPI: e.g., source /opt/intel/oneapi/setvars.sh

# Use of specific version of Intel MKL, further versions do not allow the setting
of AVX2 on non-Intel processors.
source /opt/intel/compilers_and_libraries_2019.6.324/linux/mkl/bin/mklvars.sh
intel64

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

# On AMD Rome/Milan when using Intel MKL
export MKL_DEBUG_CPU_TYPE=5
```

Compilation:

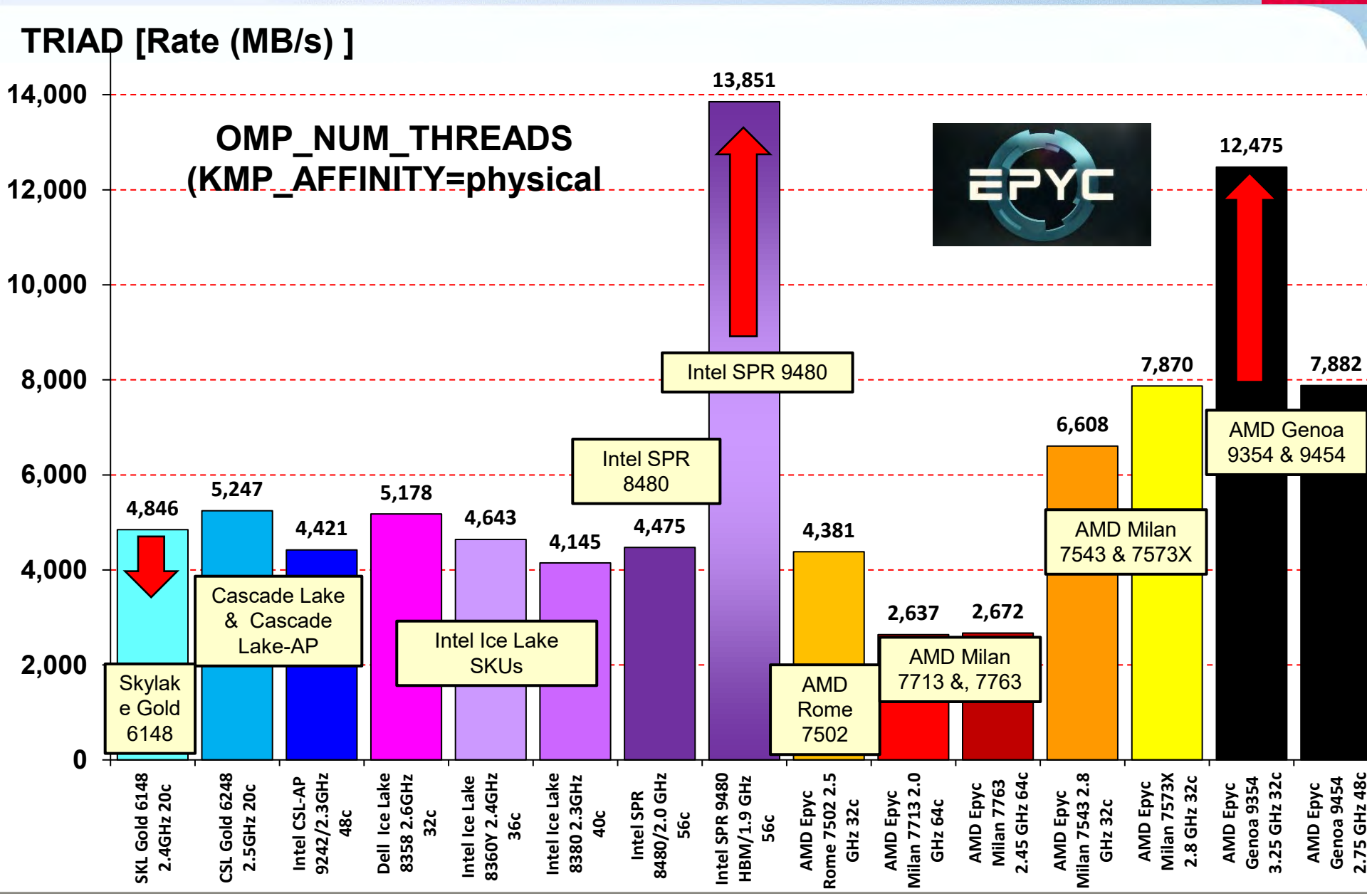
INTEL SKL: –O3 –xCORE-AVX512

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

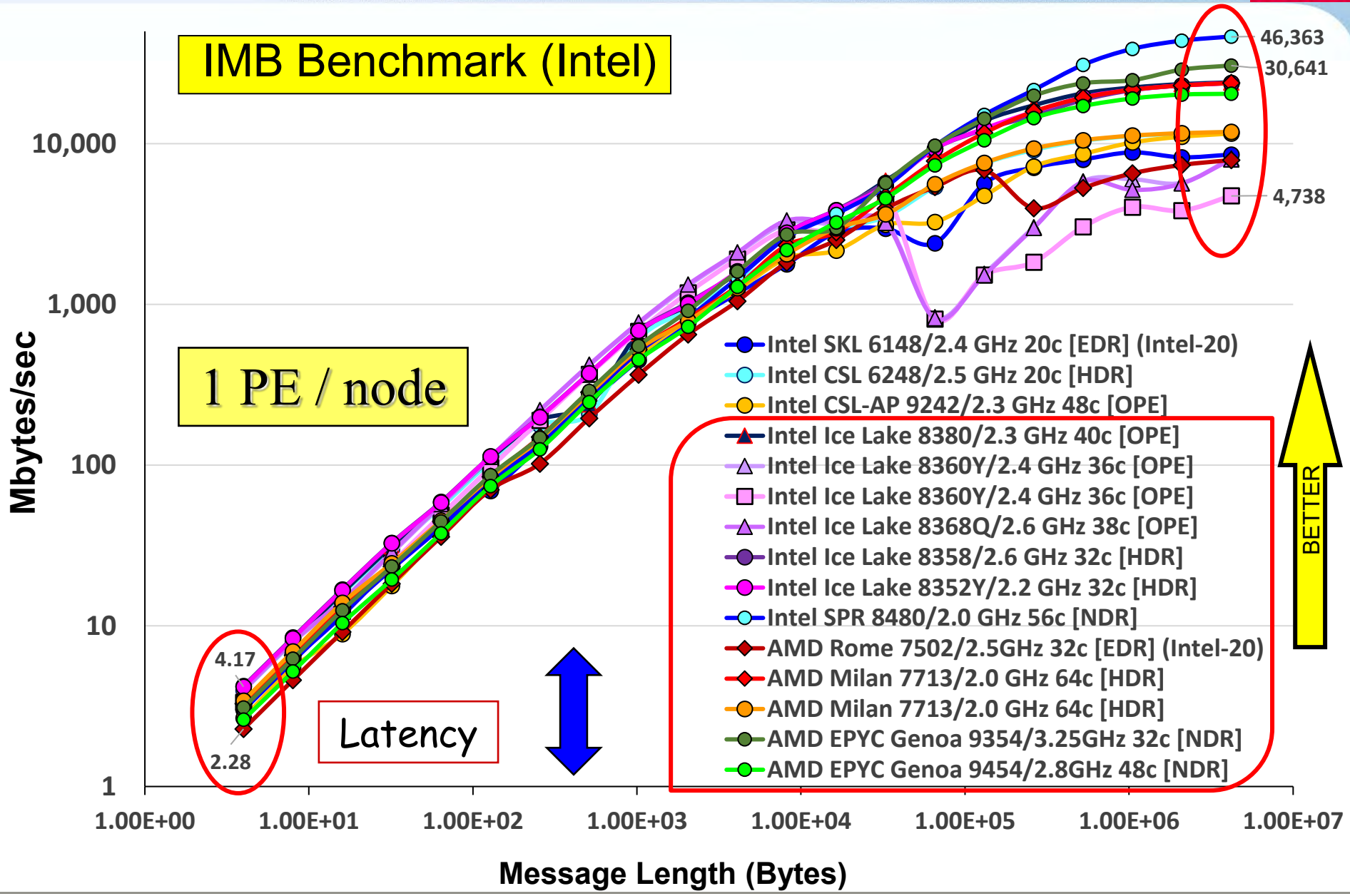
Memory B/W – STREAM performance



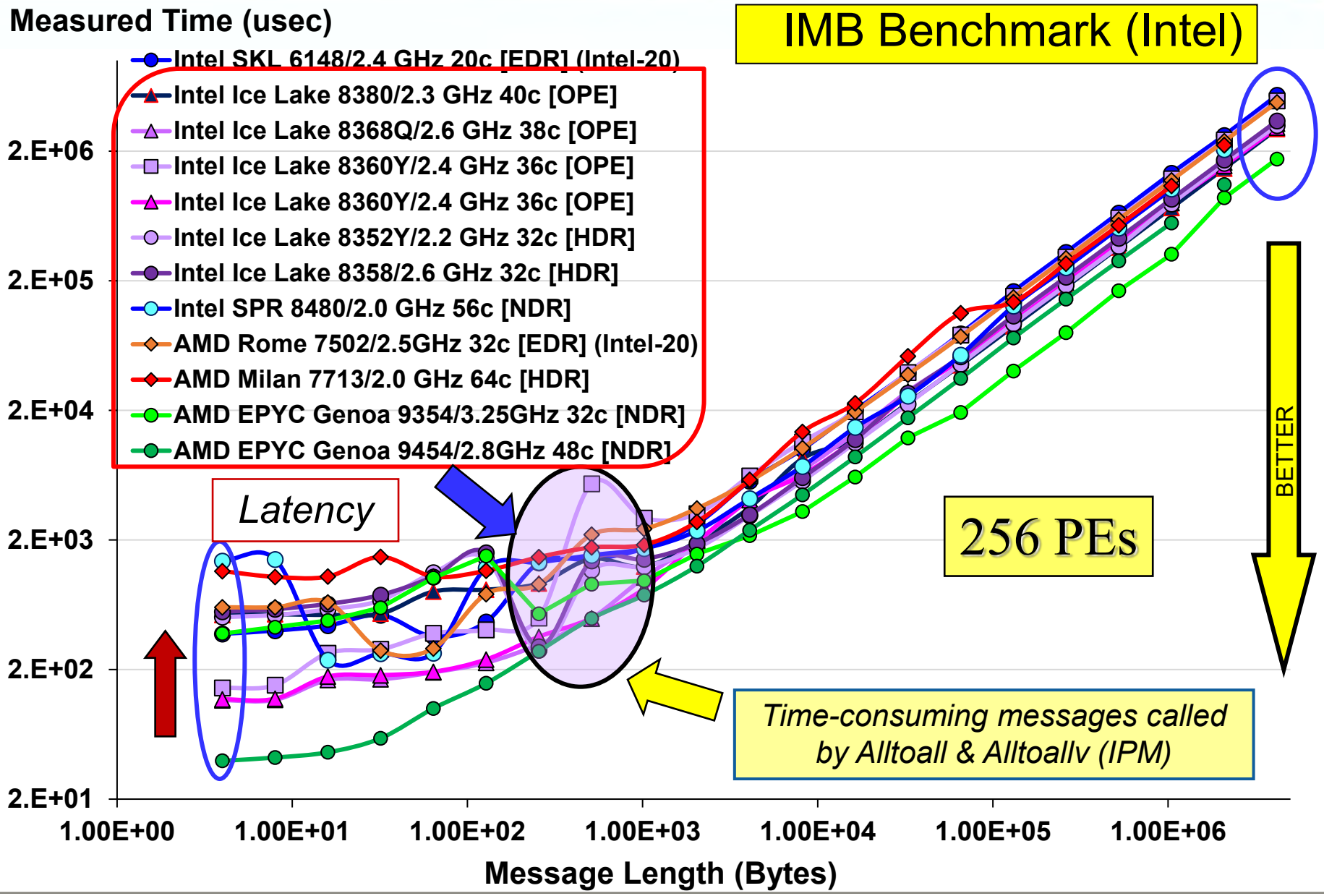
Memory B/W – STREAM / core performance



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 9354 nodes with 64 cores per node. [1-8 nodes]***
2	***Hawk - Dell |EMC Skylake Gold 6148 2.4GHz (T) EDR with 40 cores / node***	***Intel Xeon Sapphire Rapids 8480 nodes with 112 cores per node. [1-8 nodes]***

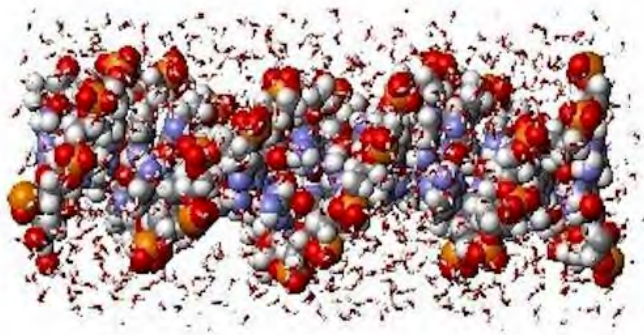
Performance of Computational Chemistry and Ocean Modelling Codes

Molecular Simulation;

1. DL_POLY

Molecular Simulation I. DL_POLY

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



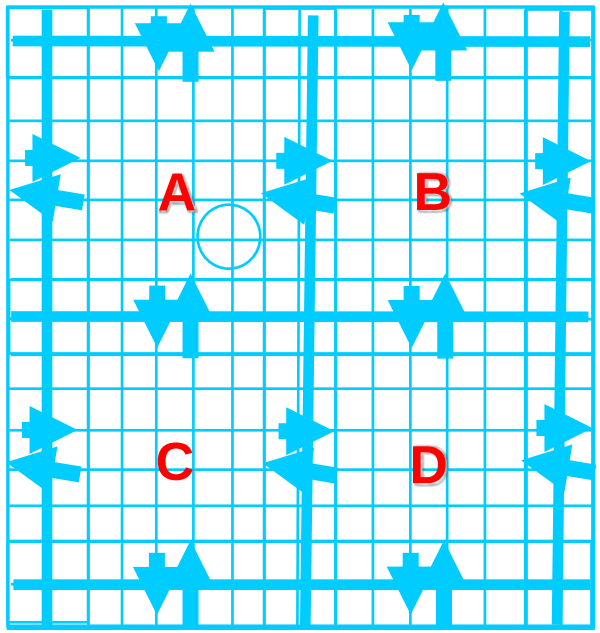
DL_POLY

- 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, 200 time steps, Cutoff=12Å**
2. **Gramicidin in water; rigid bonds + SHAKE: 792,960 ions, 50 time steps**

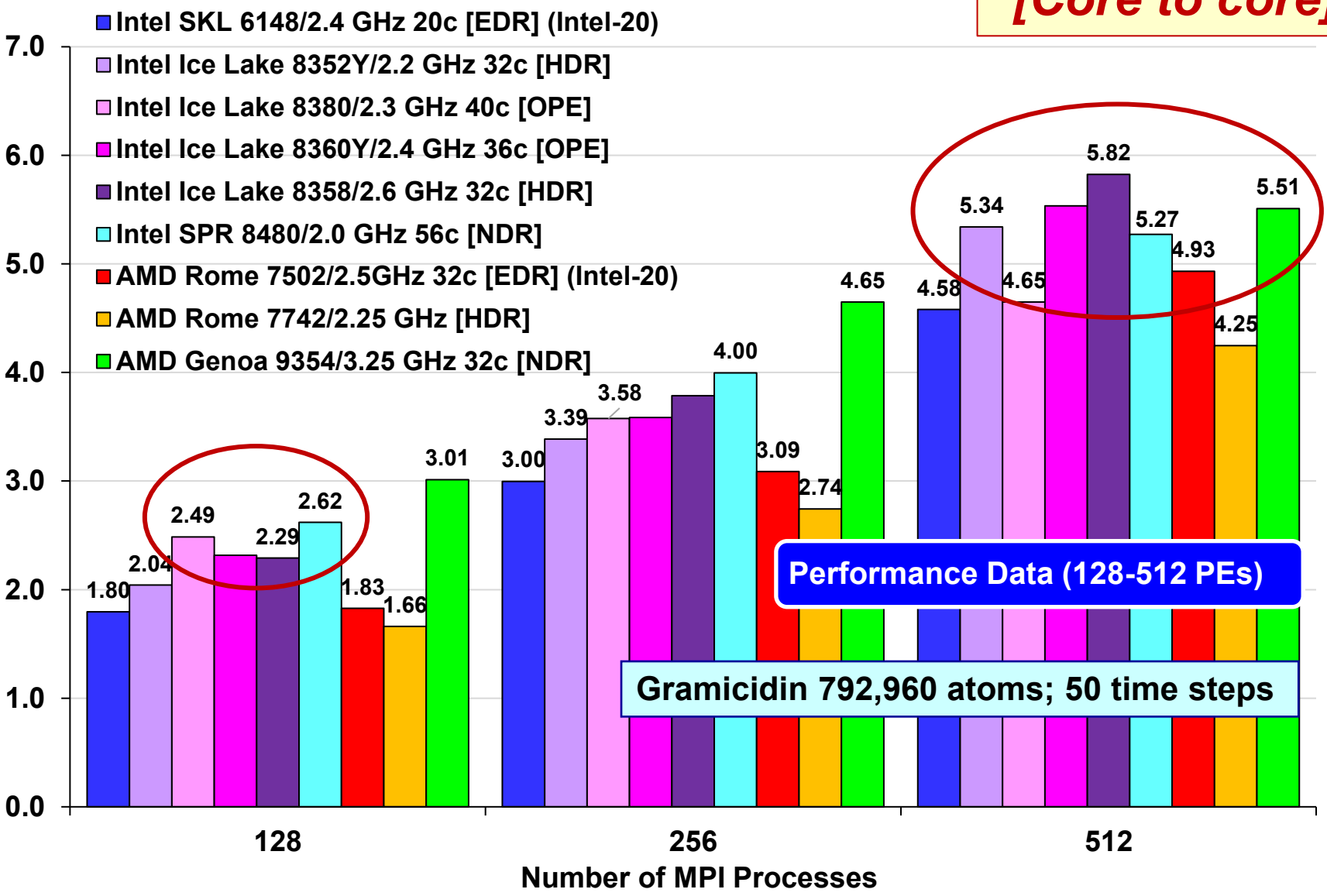
https://www.scd.stfc.ac.uk/Pages/DL_POLY.aspx

DL_POLY 4 – Gramicidin Simulation

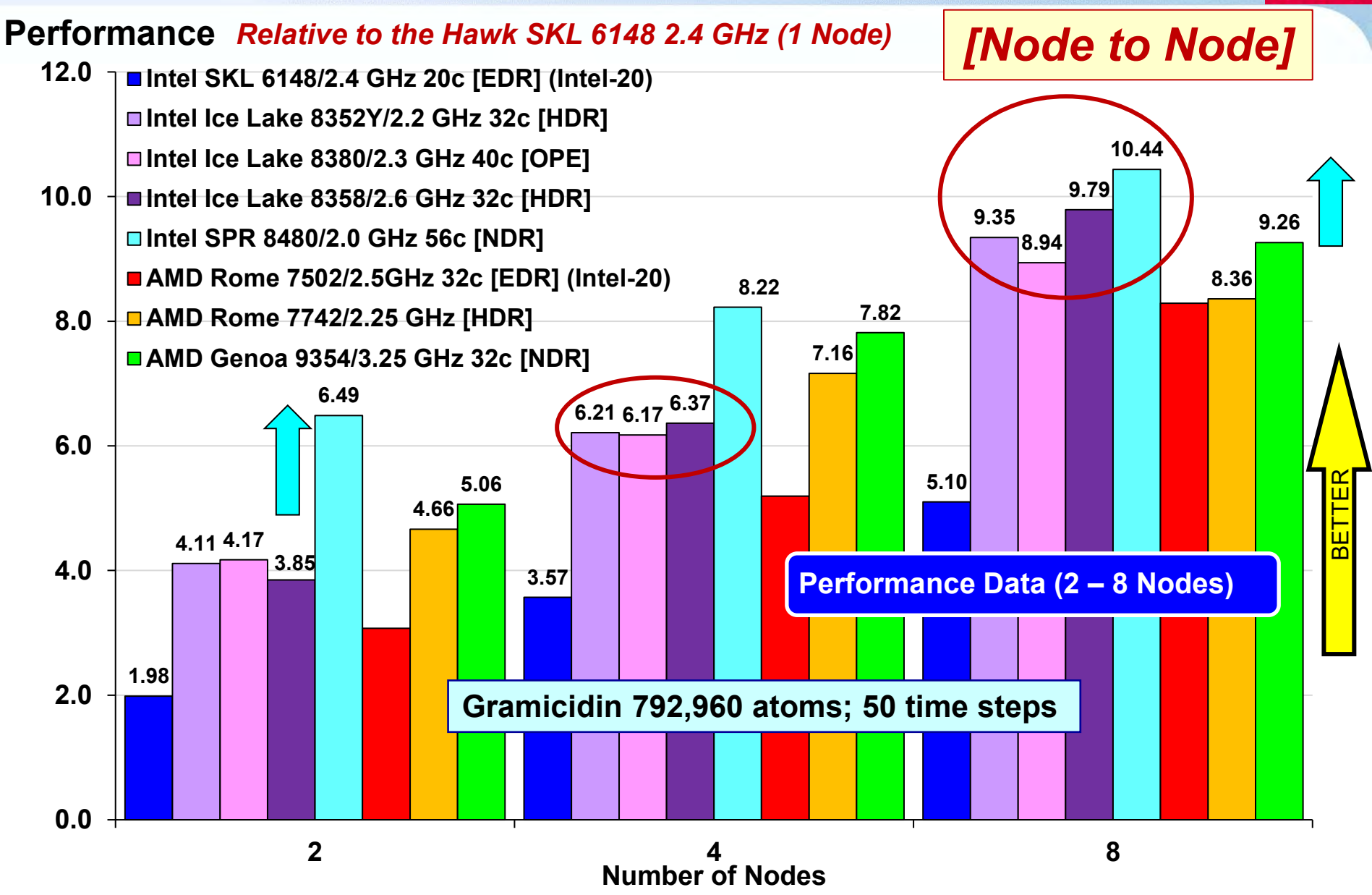
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

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

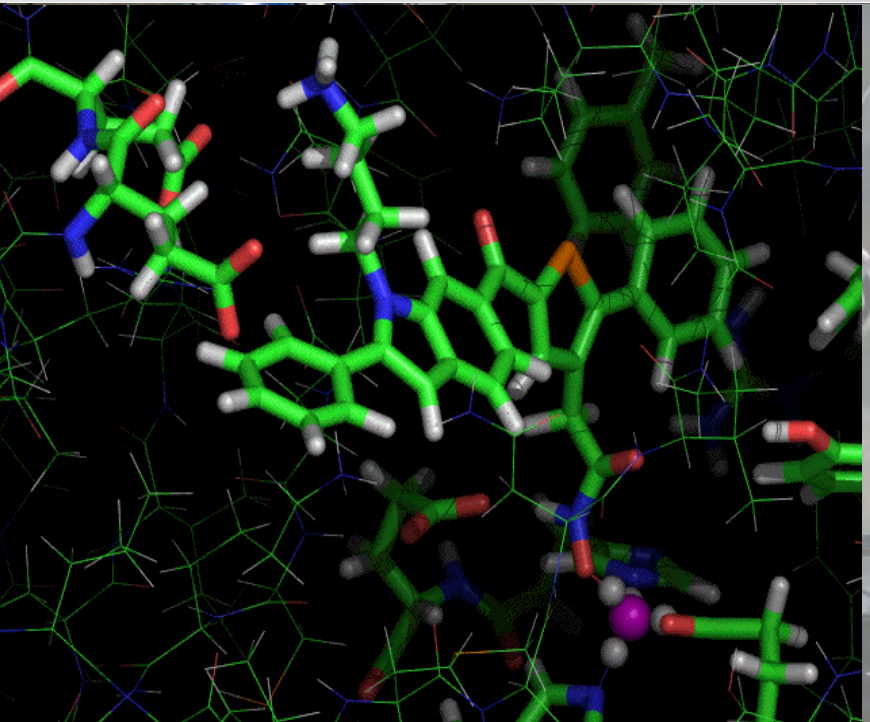
[Core to core]



DL_POLY 4 – Gramicidin Simulation



Performance of Computational Chemistry and Ocean Modelling Codes

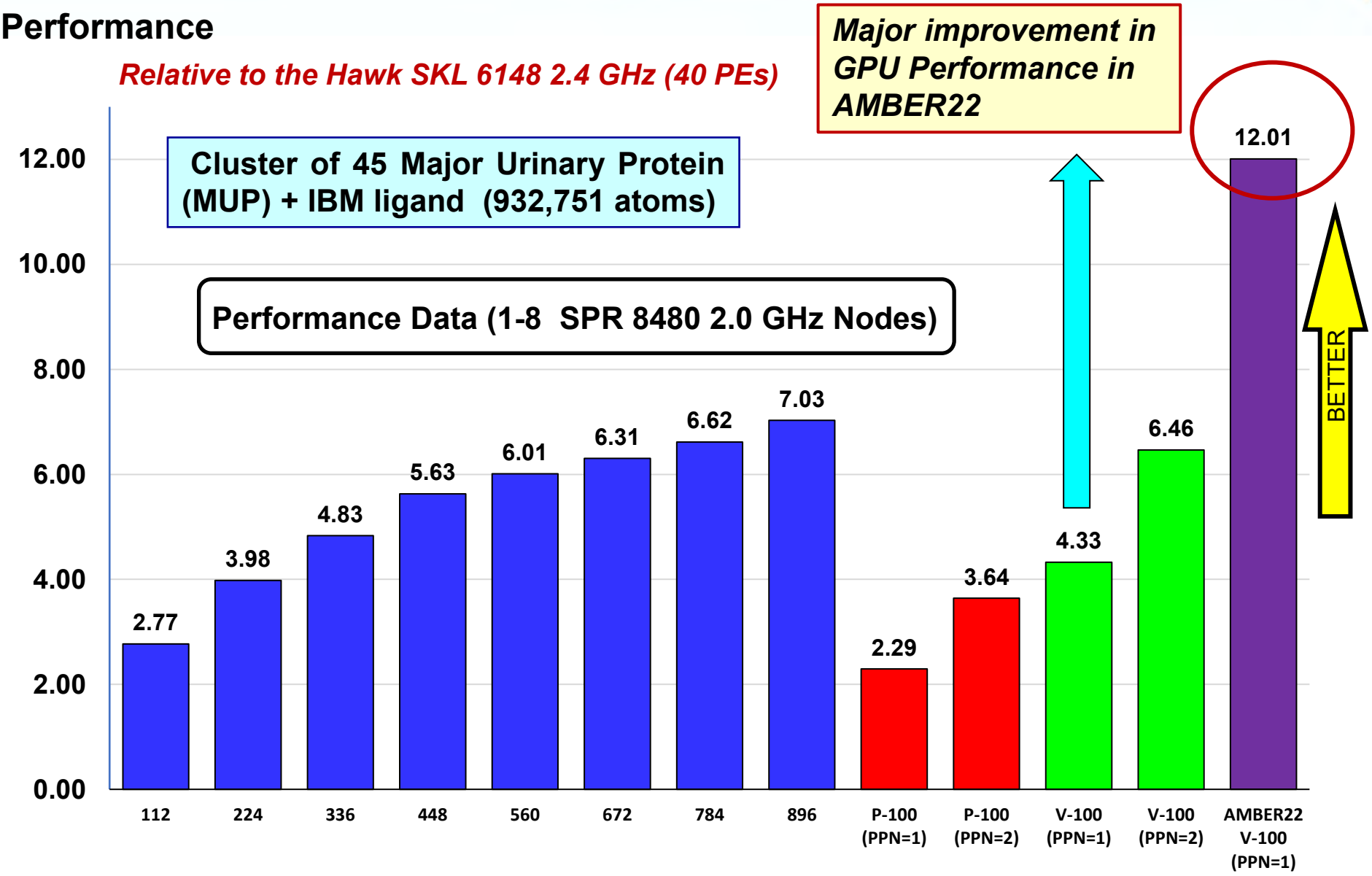


Molecular Simulation: 3. AMBER

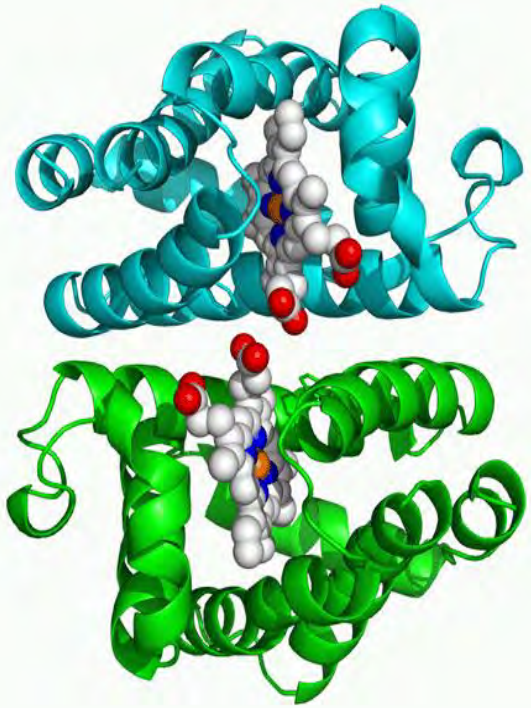
AMBER – GPU Performance M45 Simulation

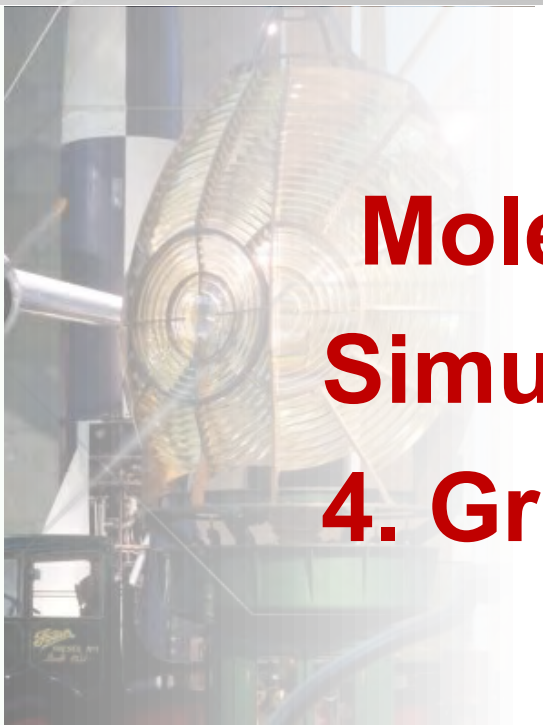
Performance

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



Performance of Computational Chemistry Codes





Molecular Simulation: 4. 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

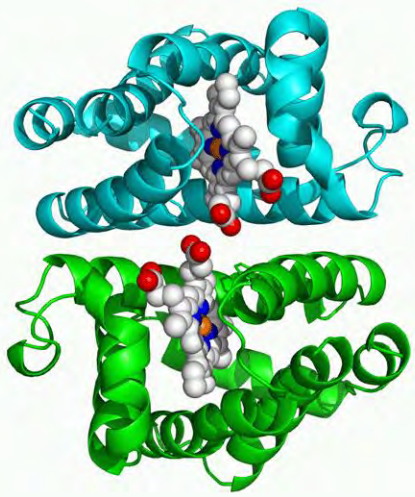
Version 2016.3 – 14 March 2017

Version 2018.2 – 14 June 2018

Version 2019.6 – 28 February 2020

Version 2020.1 – 3 March 2020

Version 2023.1 – 21 April 2023



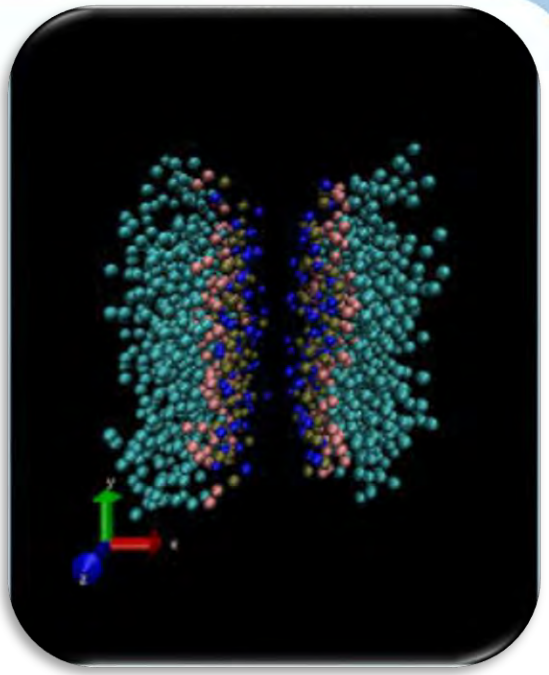
- **Berk Hess et al. "*GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*". *Journal of Chemical Theory and Computation* 4 (3): 435–447.**

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

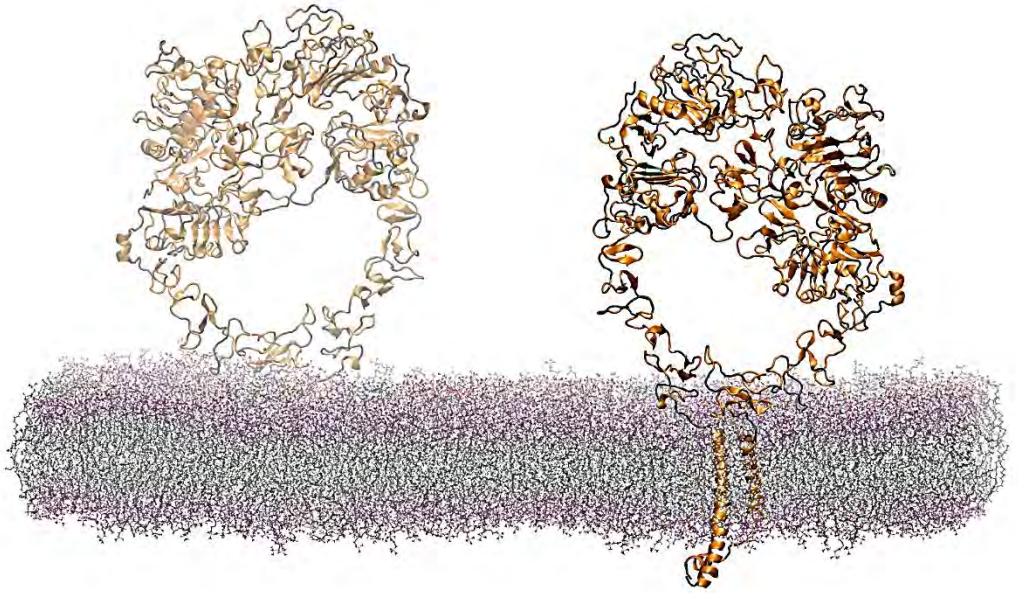


Lignocellulose

- 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

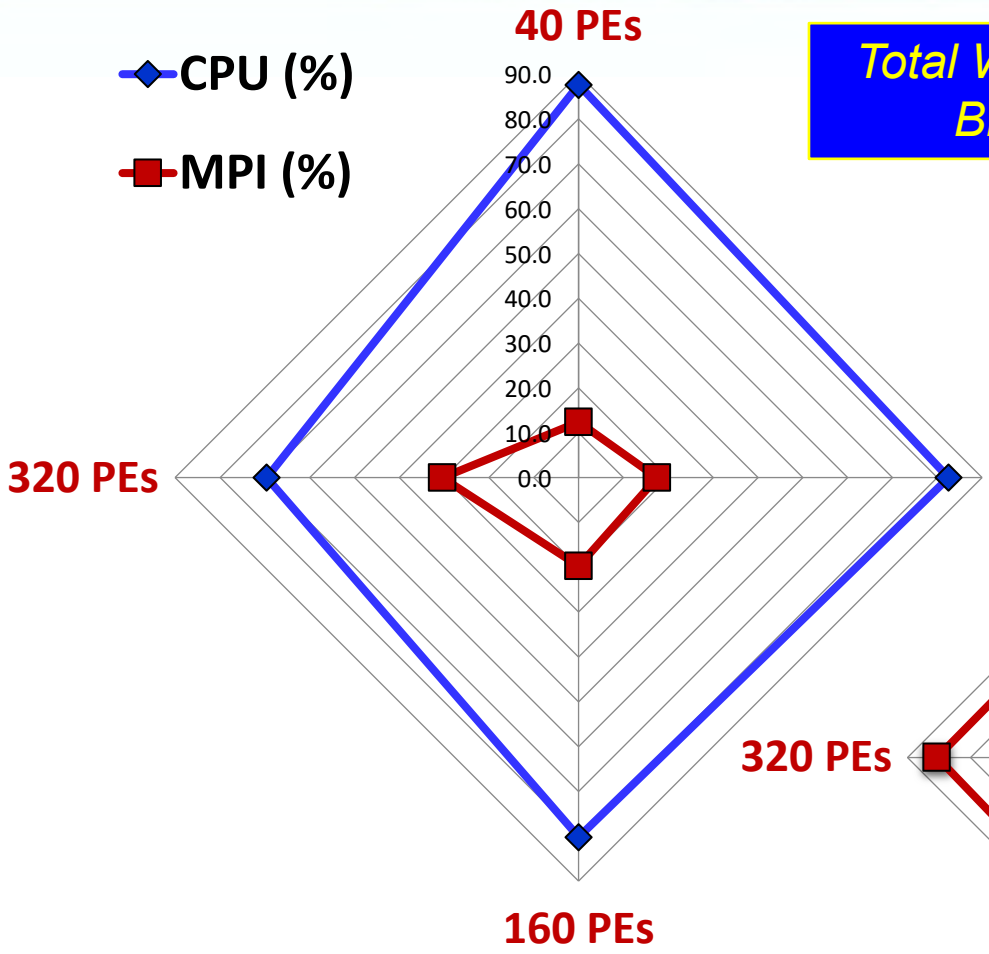
The HECBioSim Benchmarks



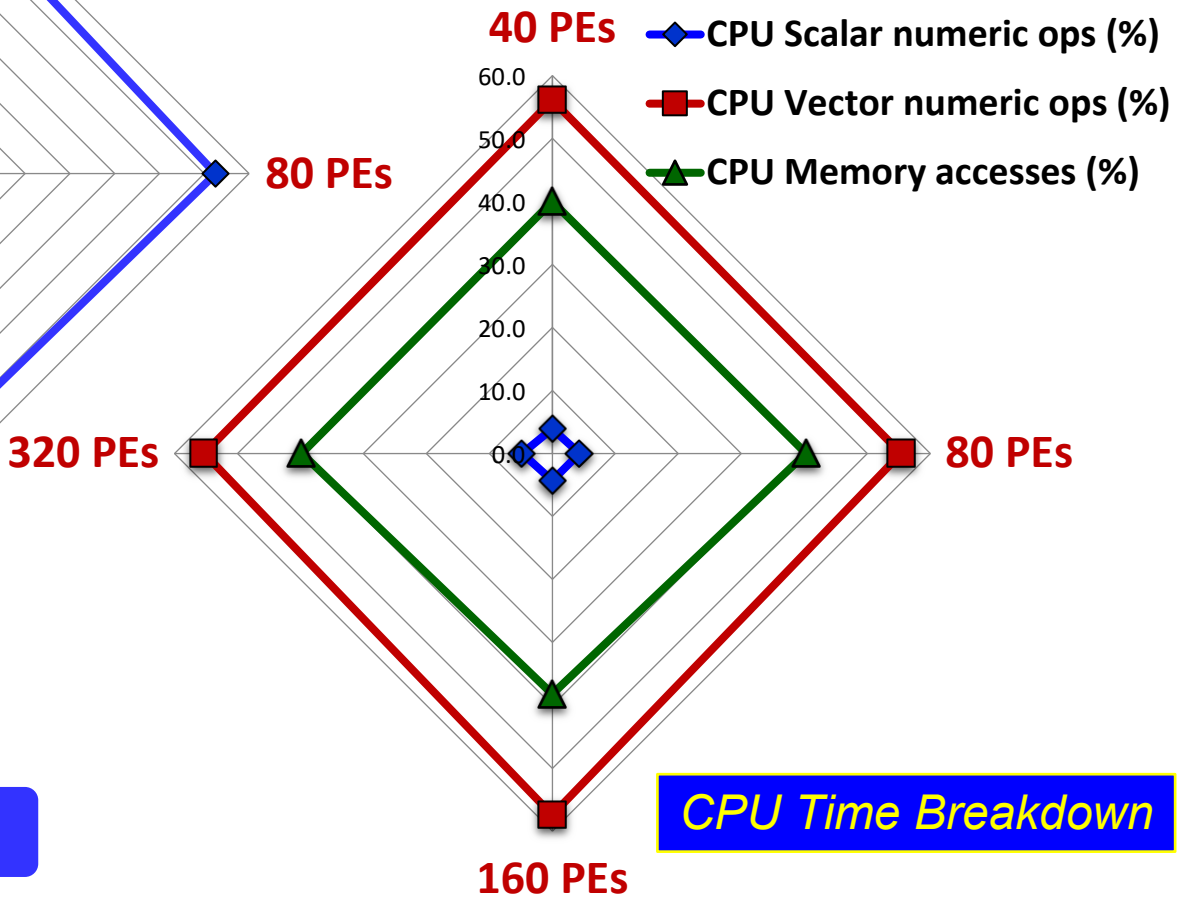
- PME simulation for 1.4M atom system - A Pair of Human Epidermal Growth Factor Receptor (hEGFR) Dimers of 1IVO and 1NQL
- Total number of atoms = **1,403,182**
- Protein atoms = 43,498 Lipid atoms = 235,304 Water atoms = 1,123,392 Ions = 986

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

GROMACS – HECBioSim Performance Report



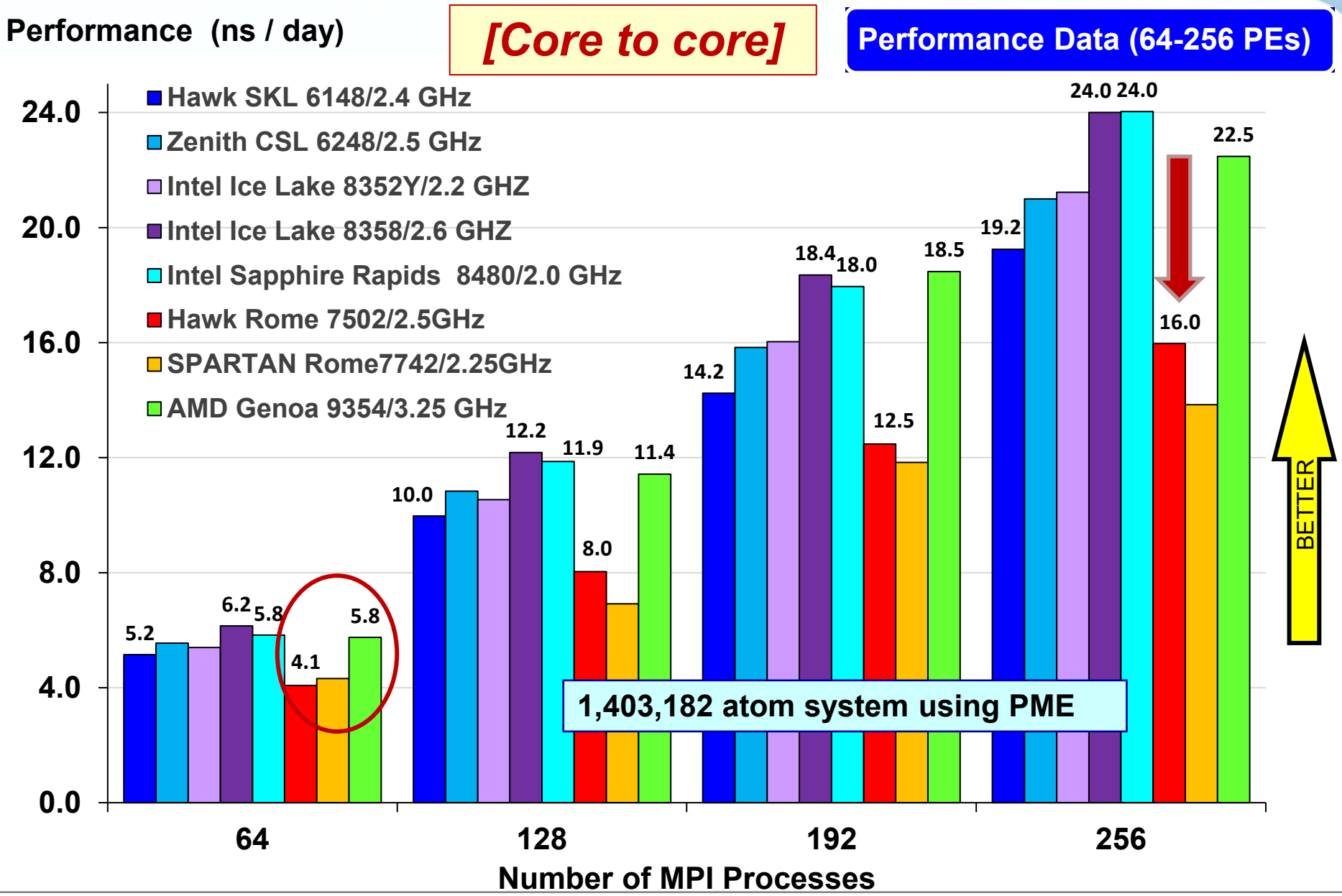
Total Wallclock Time Breakdown



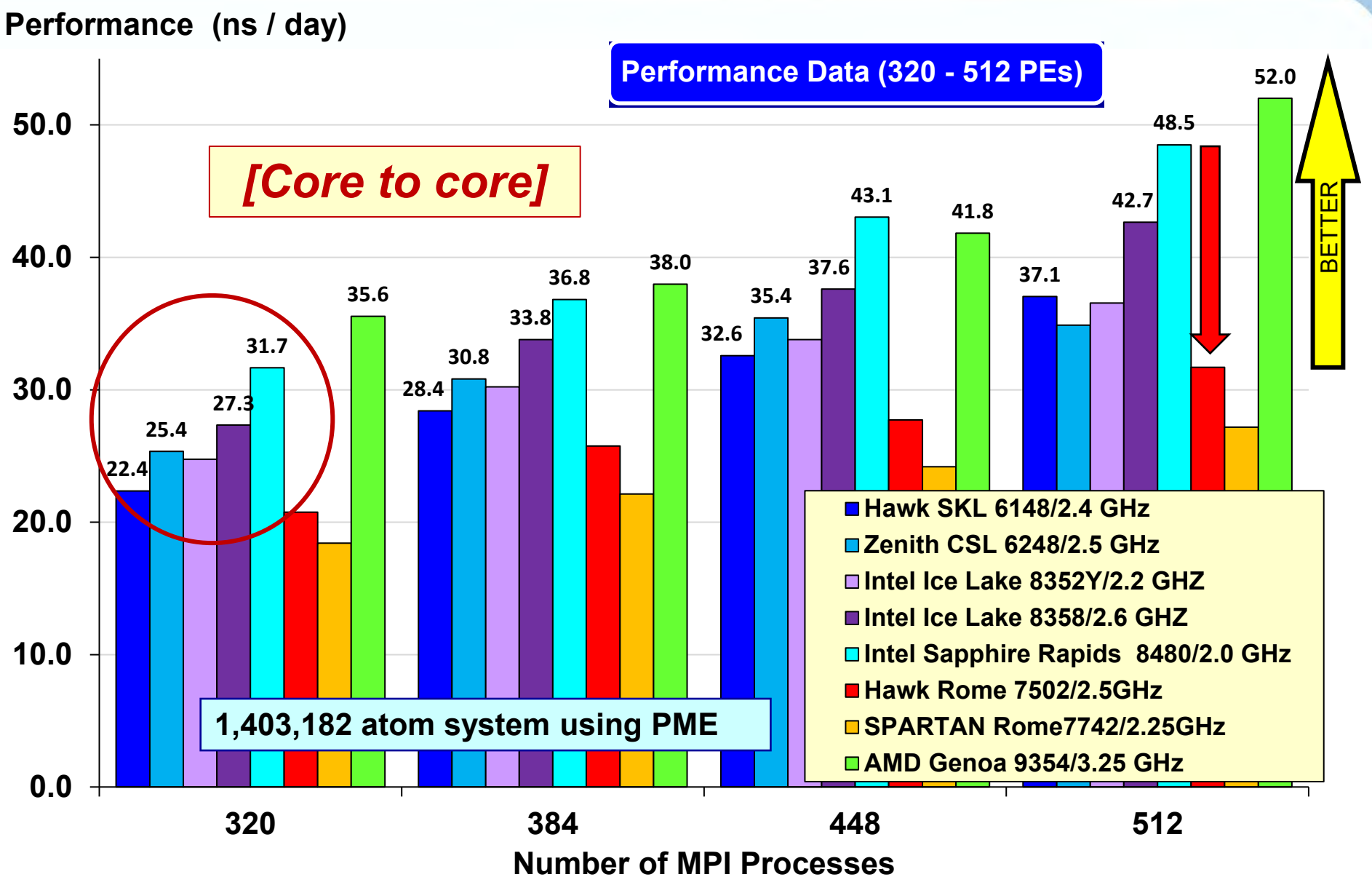
CPU Time Breakdown

Performance Data (40-320 PEs)

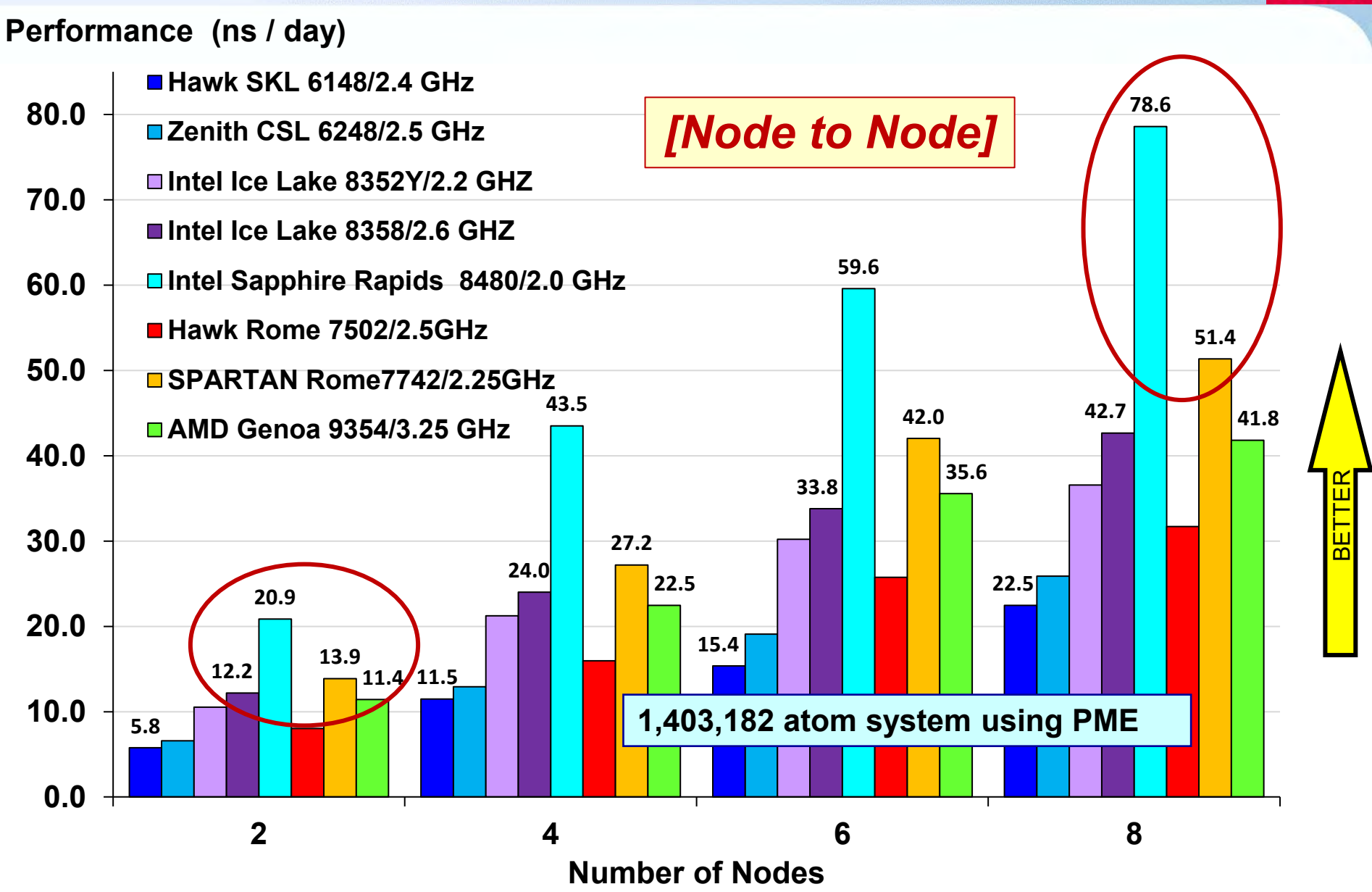
GROMACS – HECBioSim 1.4M Atom System



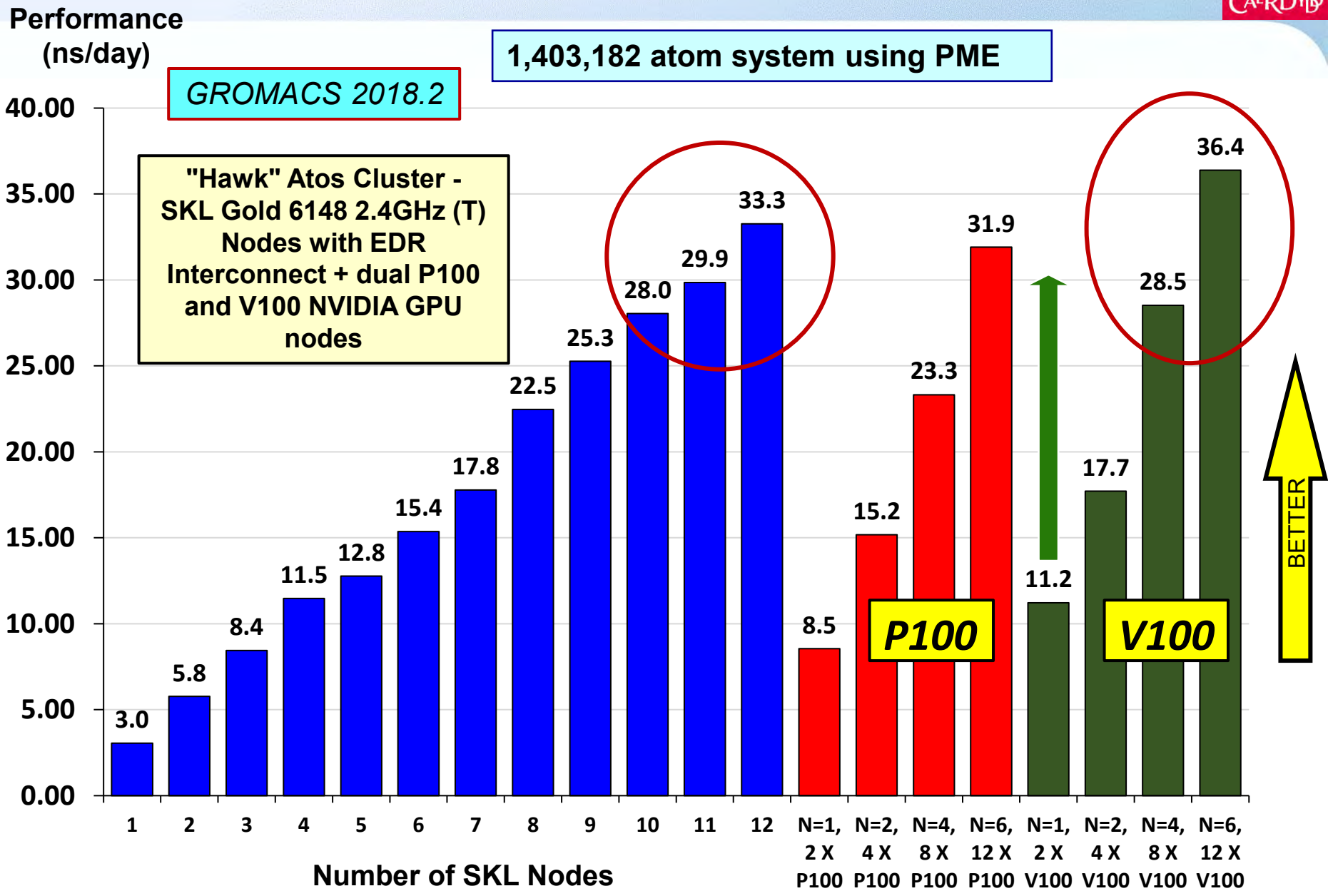
GROMACS – HECBioSim 1.4M Atom System



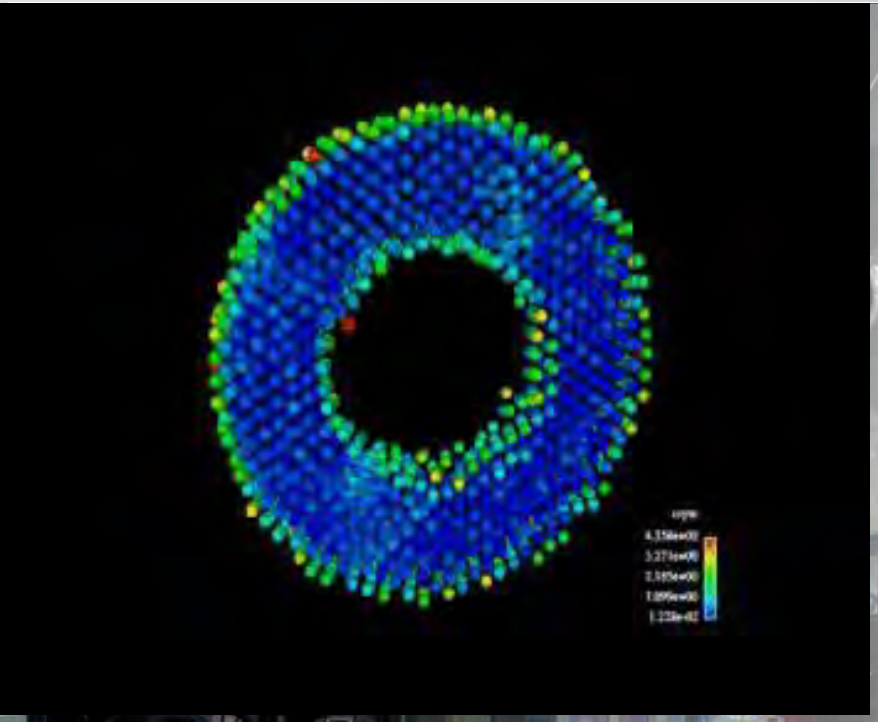
GROMACS – HECBioSim 1.4M Atom System

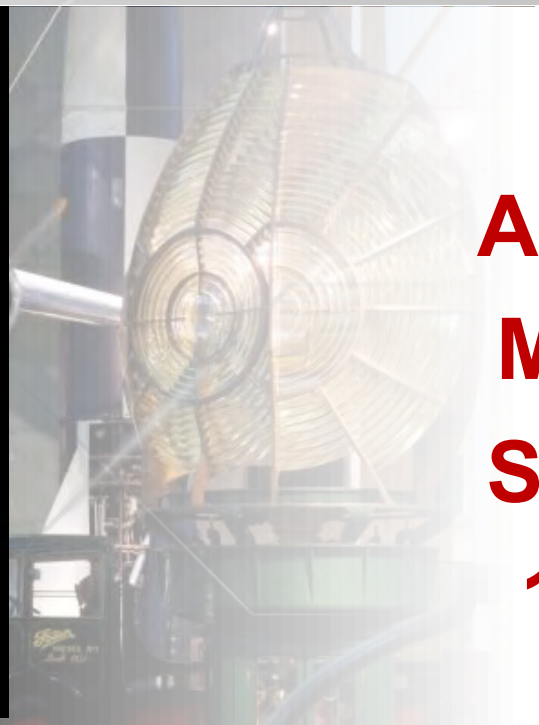


GROMACS – GPU Performance: HECBioSim Simulation



Performance of Computational Chemistry and Ocean Modelling Codes





Advanced Materials Software:

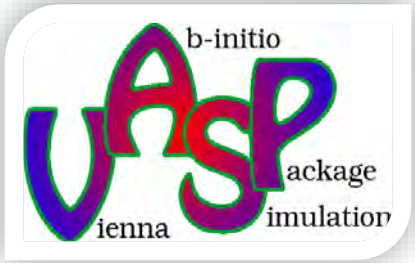
1. VASP

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



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

Archer Rank: 1

Pd-O Benchmark

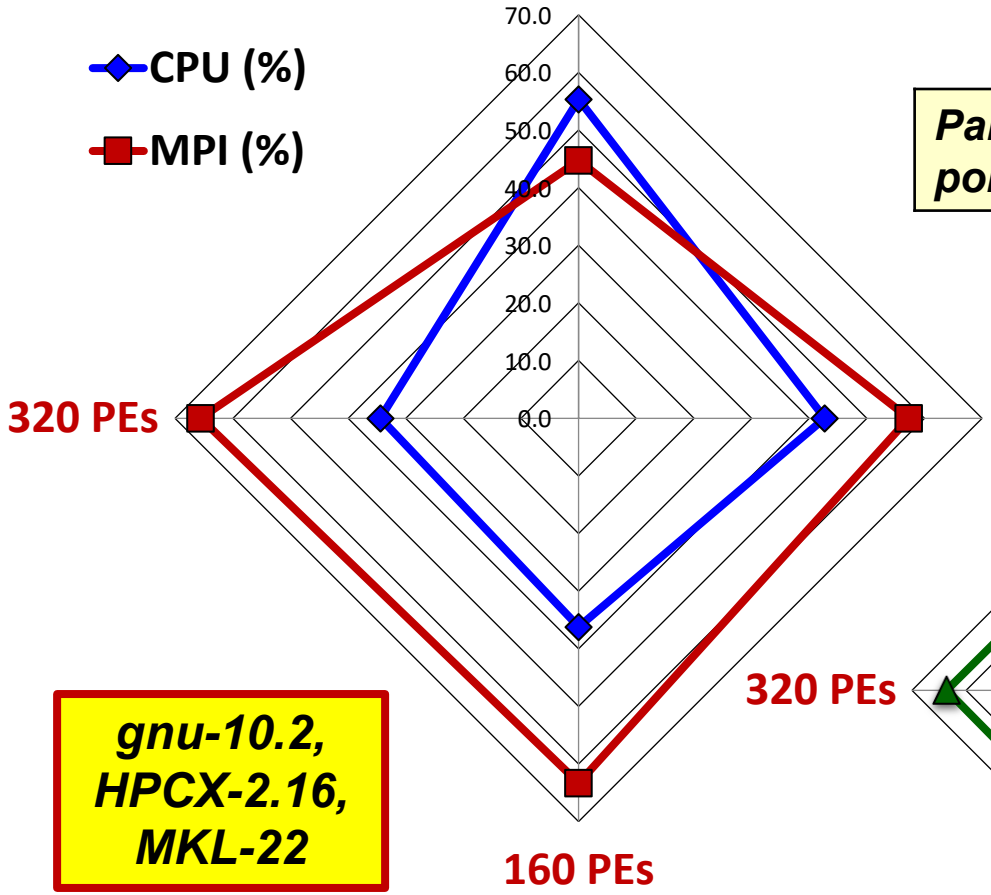
- Pd-O complex – $Pd_{75}O_{12}$, 5X4 3-layer supercell running a single point calculation and a planewave cut off of 400eV. Uses the RMM-DIIS algorithm for the SCF and is calculated in real space.
- 10 k-points; maximum number of plane-waves: 34,470
- FFT grid; NGX=31, NGY=49, NGZ=45, giving a total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

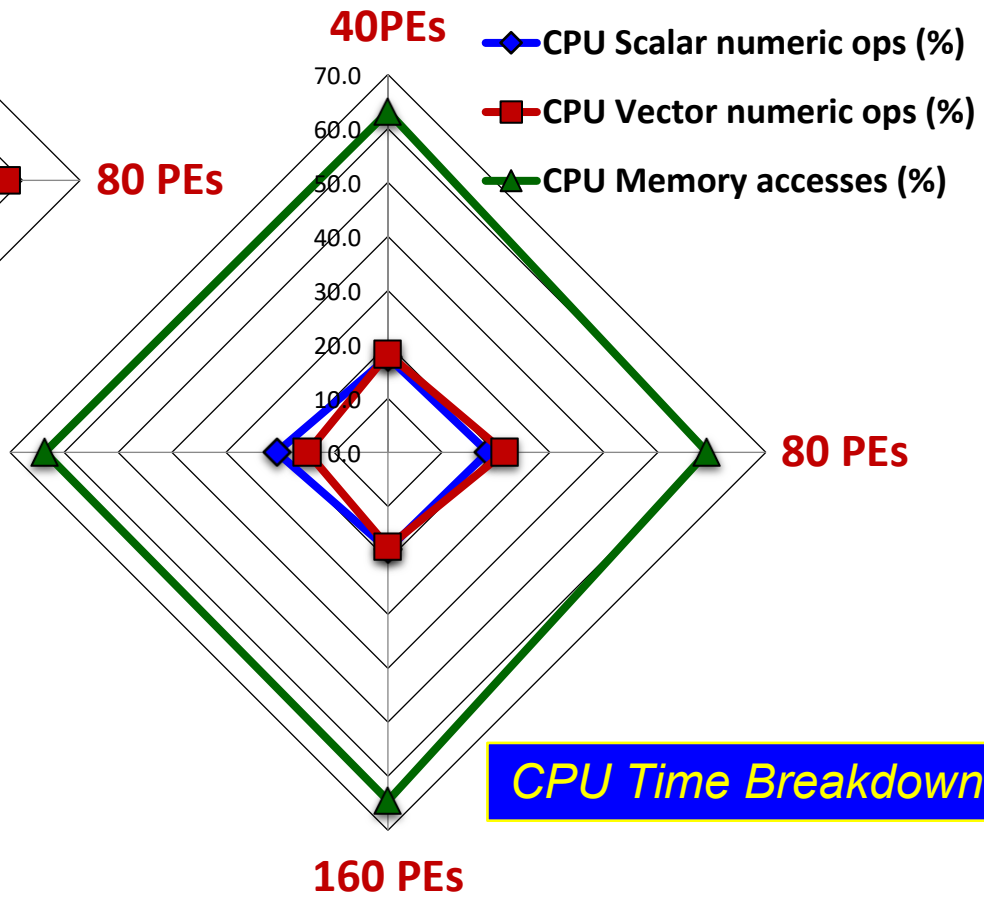
Performance Data (40-320 PEs)



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

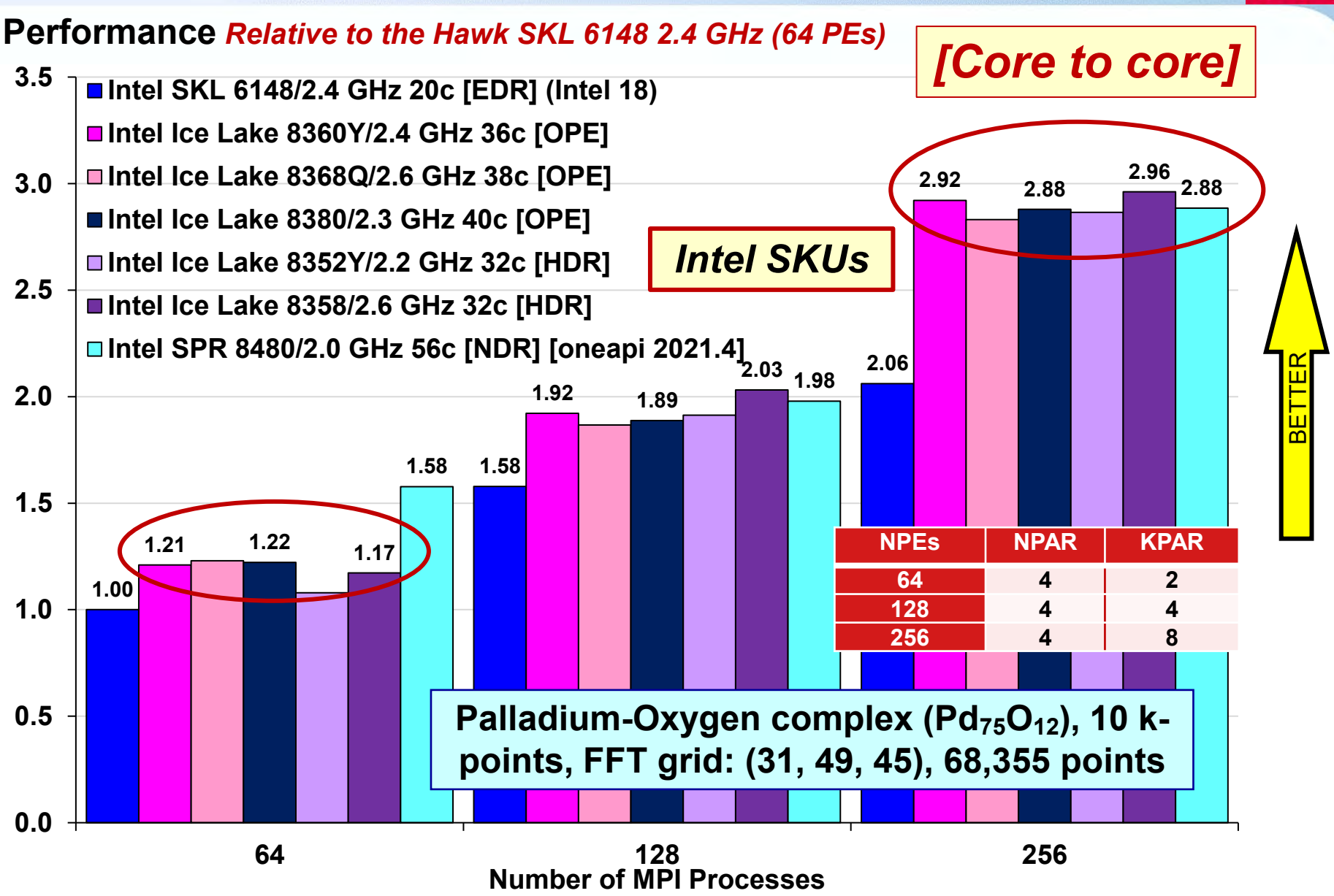
gnu-10.2, HPCX-2.16, MKL-22

Total Wallclock Time Breakdown

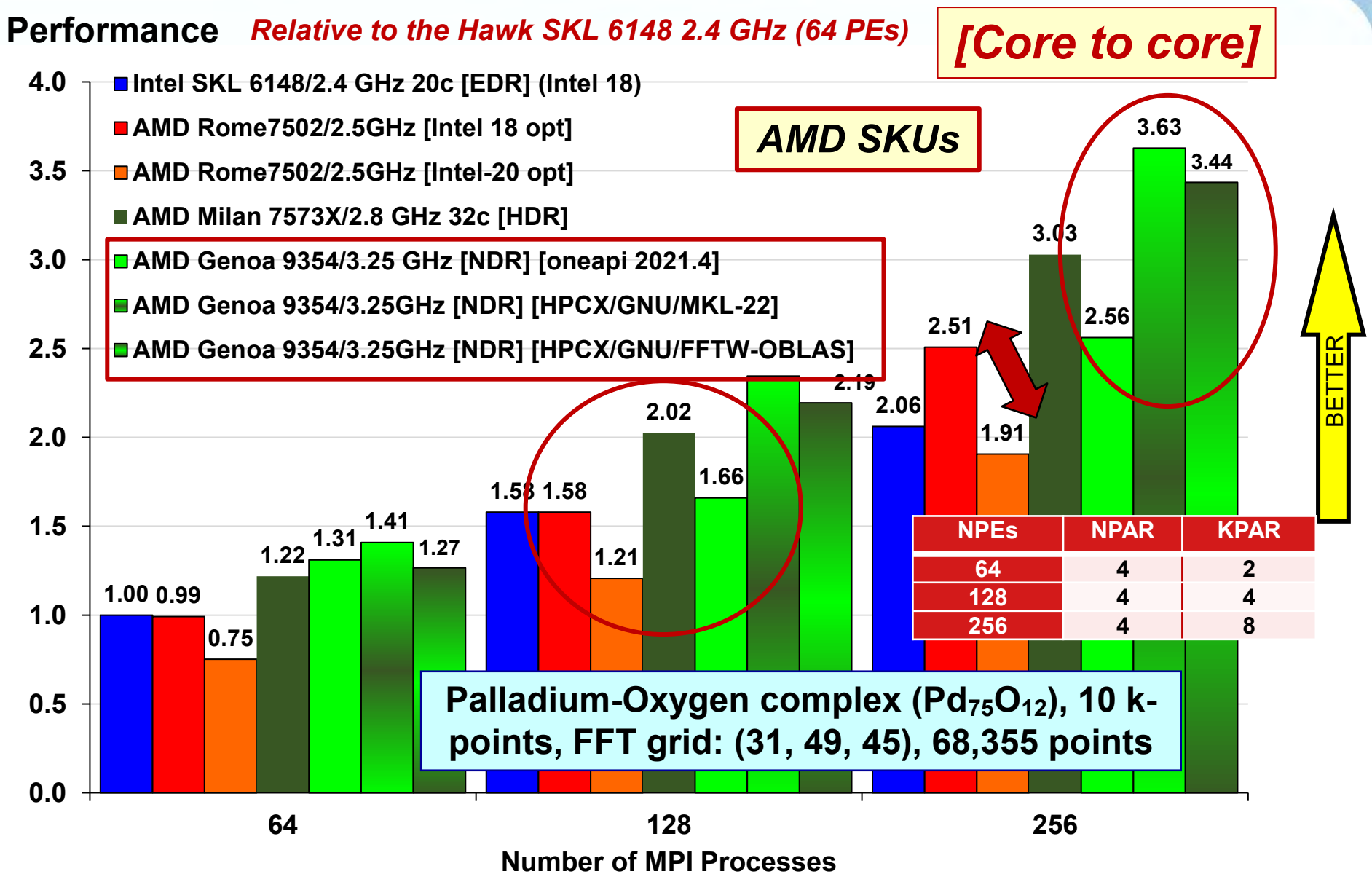


CPU Time Breakdown

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



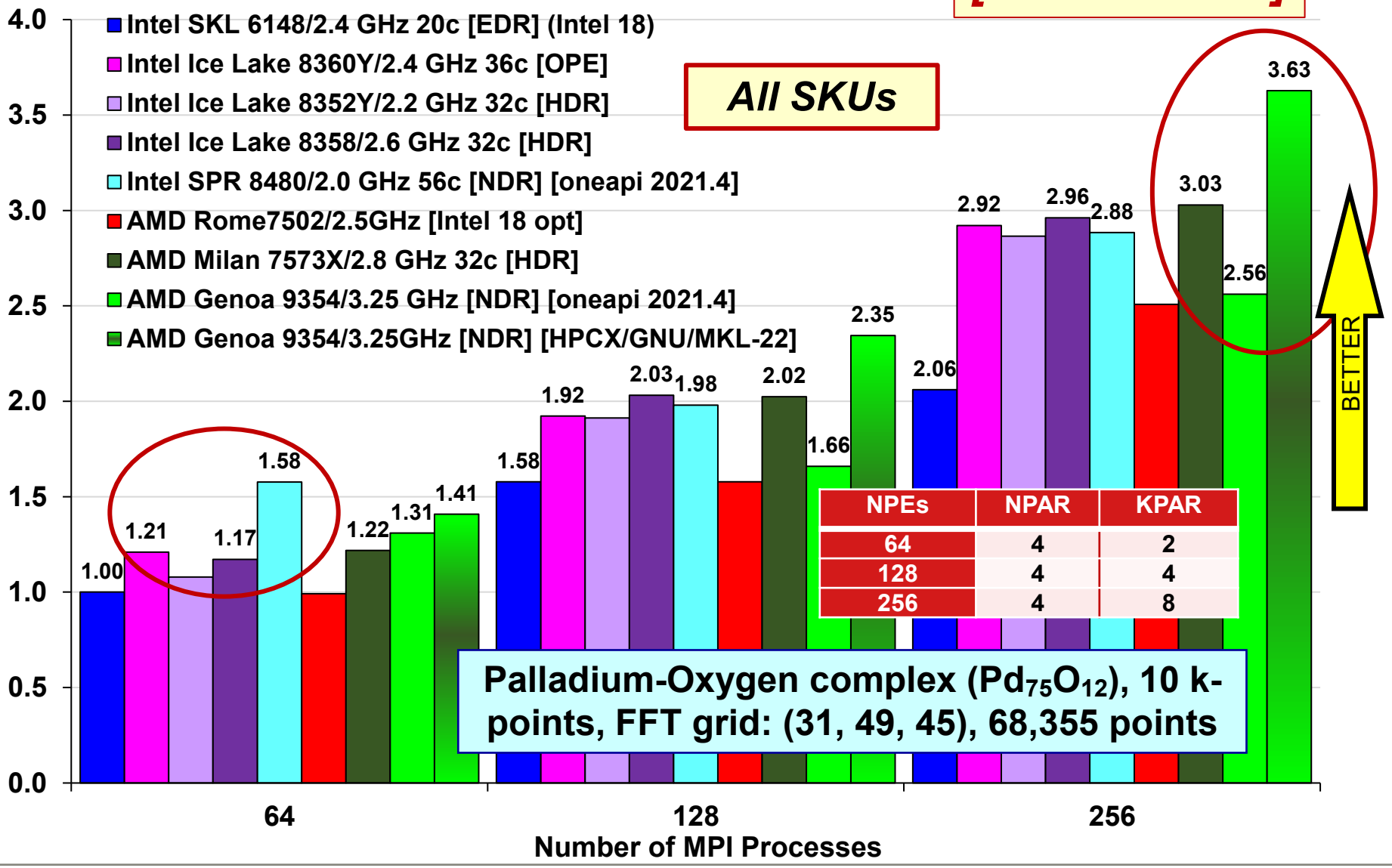
VASP 6.3 – Pd-O Benchmark - Parallelisation on k-points



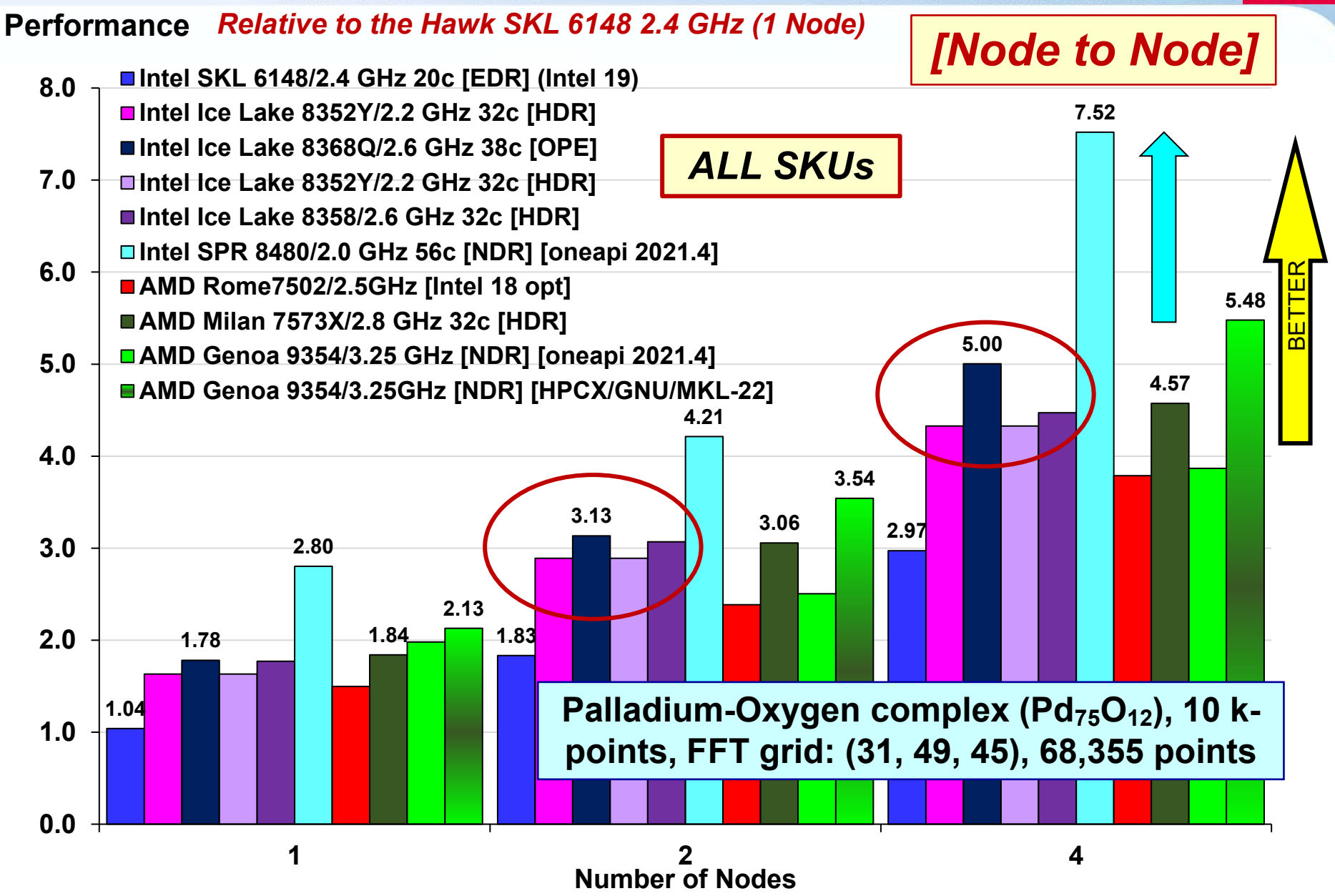
VASP 6.3 – Pd-O Benchmark - Parallelisation on k-points

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

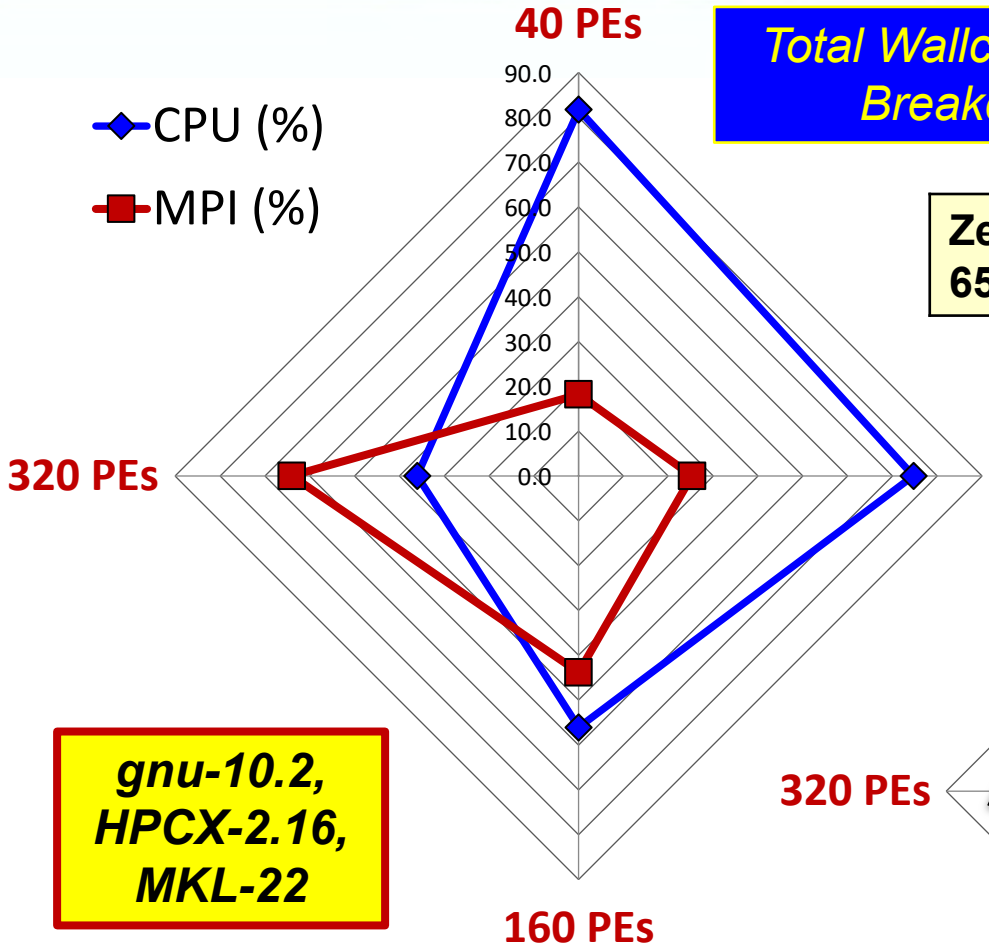
[Core to core]



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

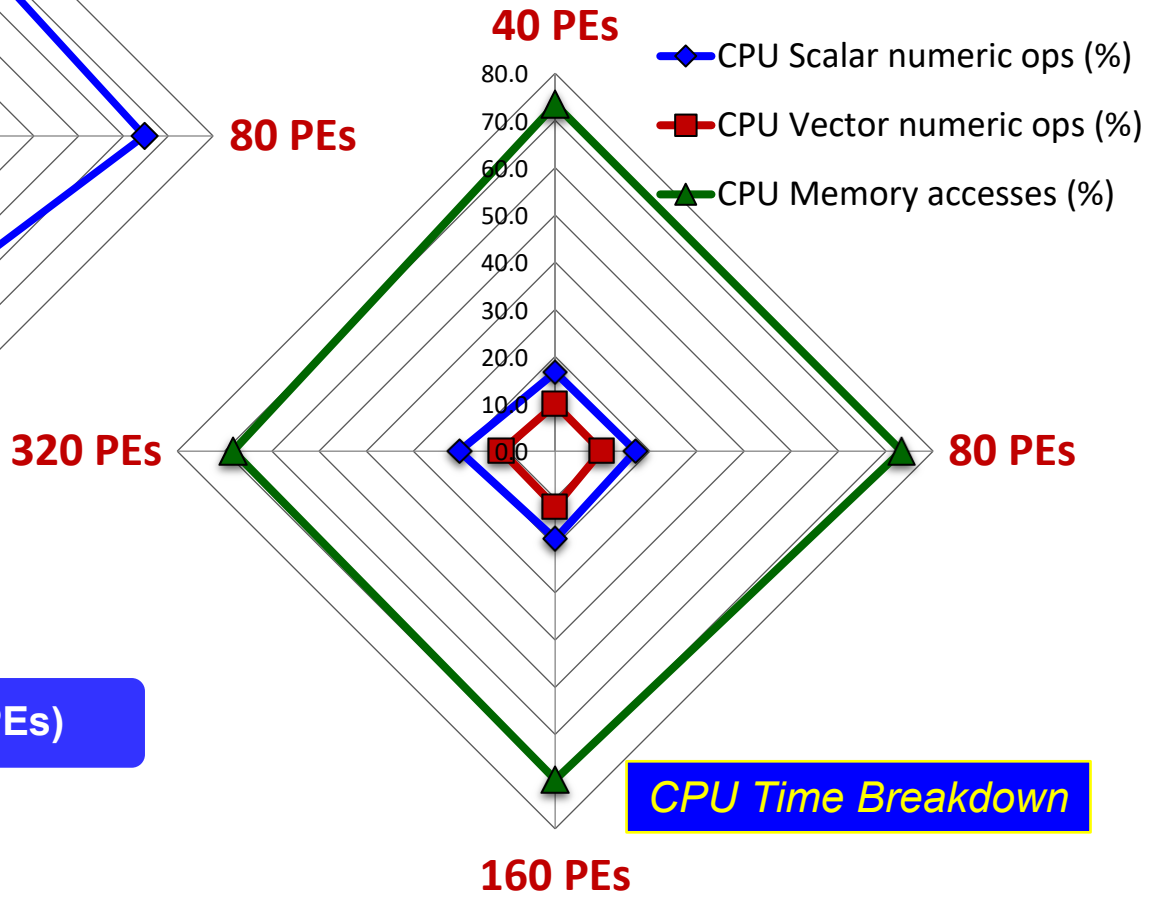


VASP – Zeolite Cluster Performance Report



Total Wallclock Time Breakdown

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



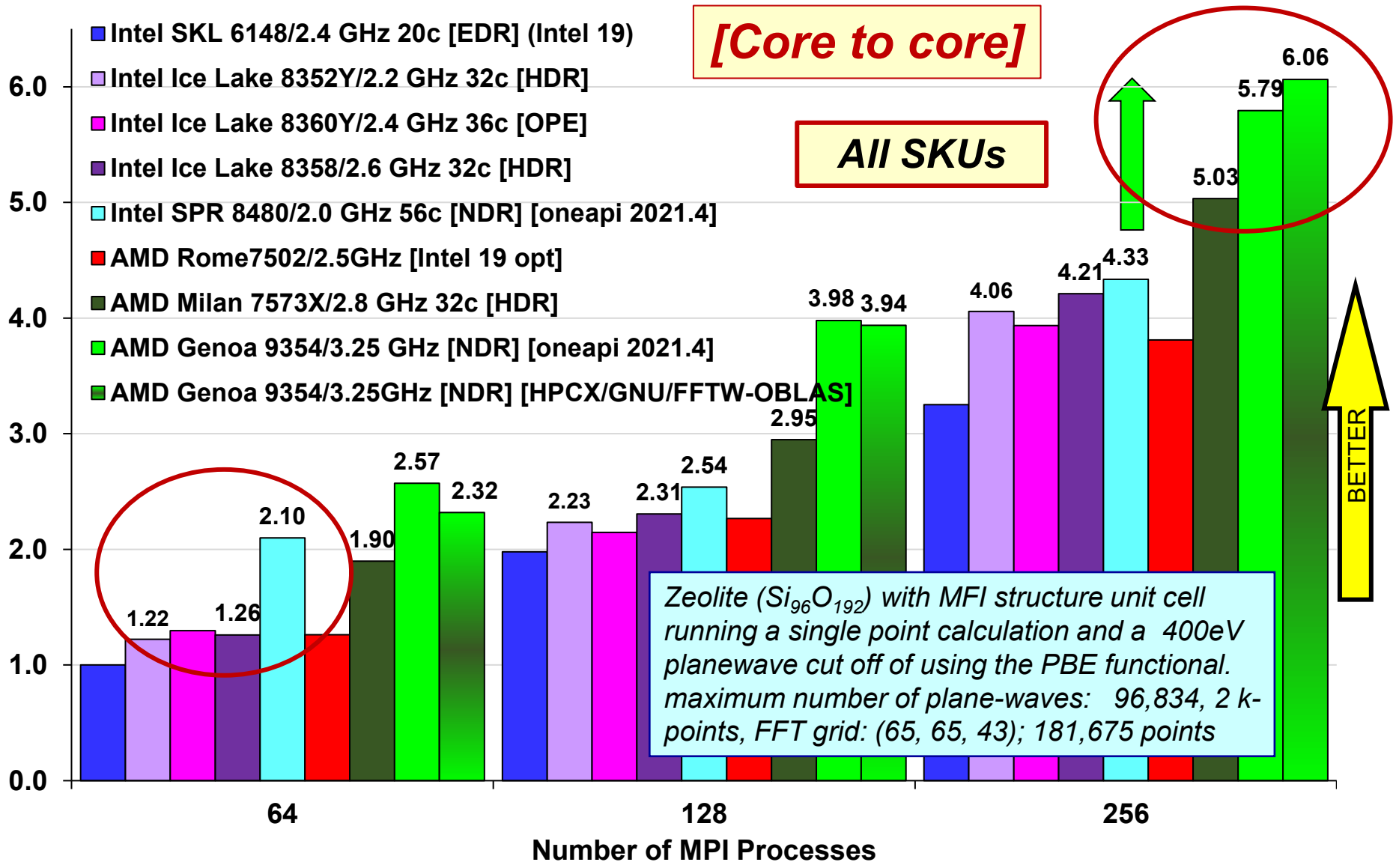
CPU Time Breakdown

gnu-10.2, HPCX-2.16, MKL-22

Performance Data (40-320 PEs)

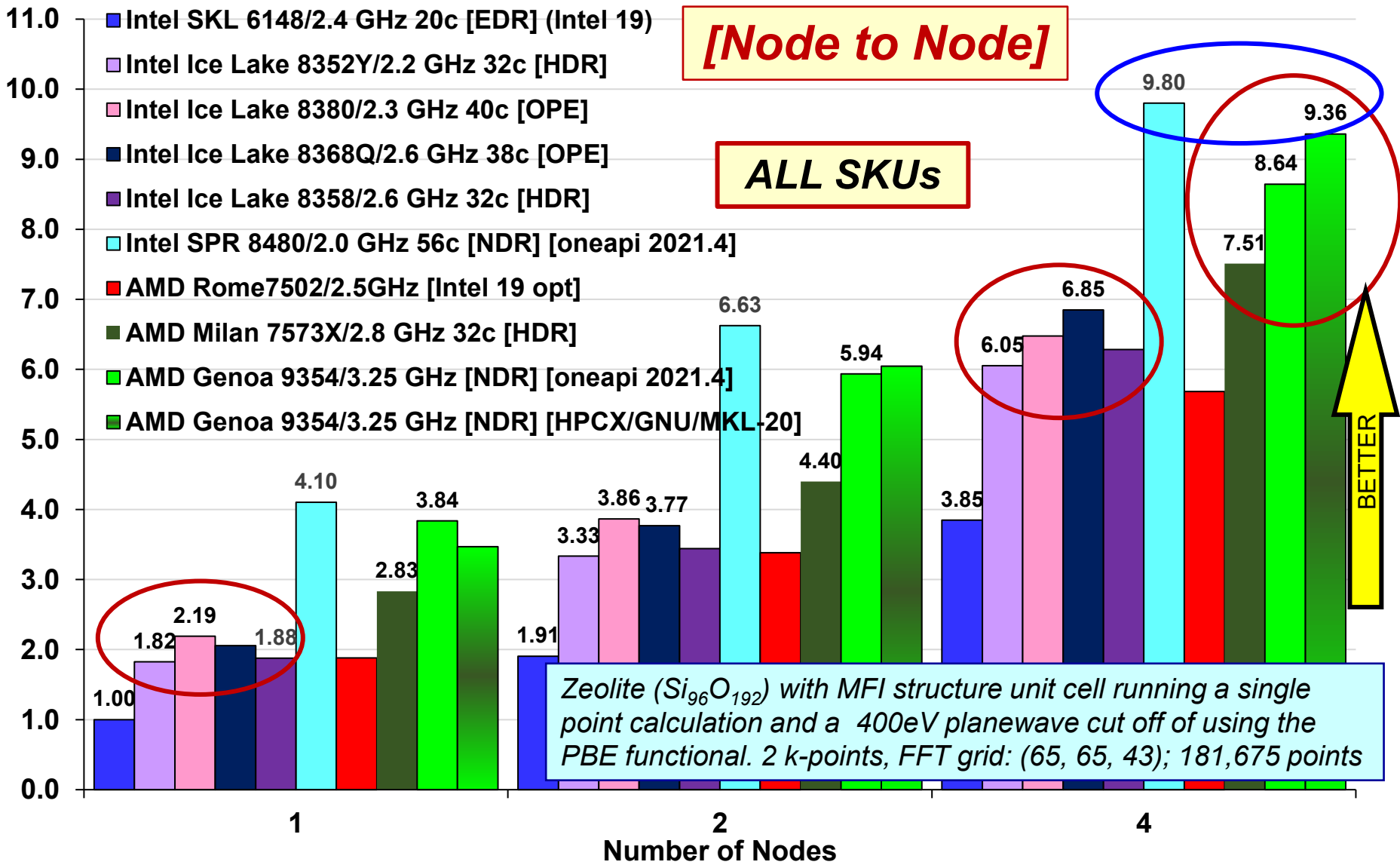
VASP 6.3 – Zeolite Benchmark - Parallelisation on k-points

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

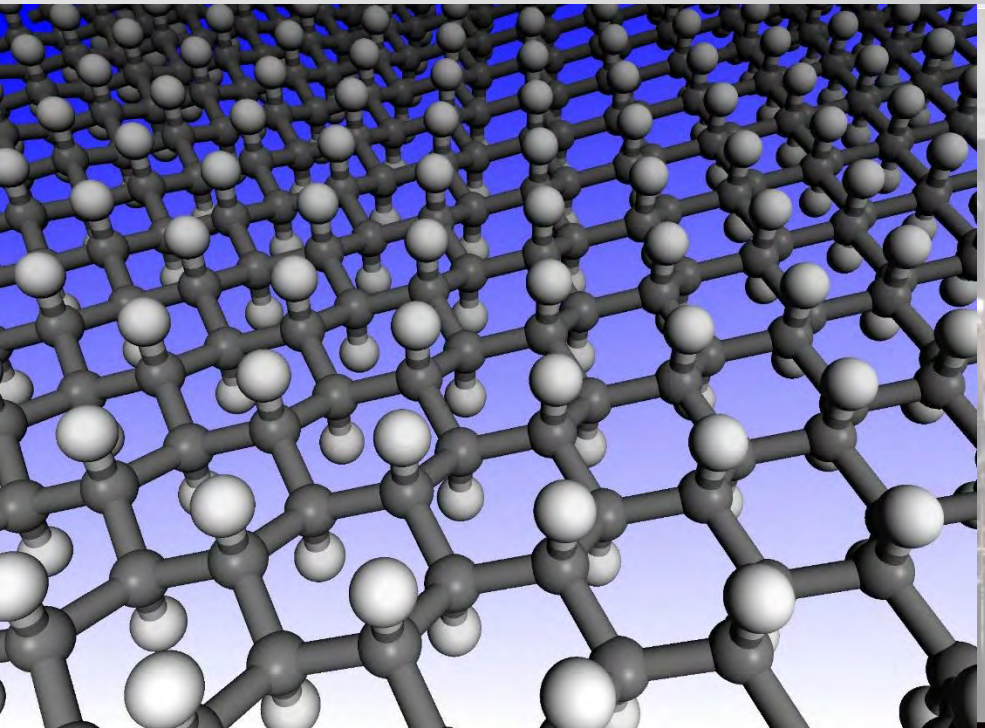


VASP 6.3 – Zeolite Benchmark - Parallelisation on k-points

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



Performance of Computational Chemistry and Ocean Modelling Codes



Advanced Materials Software:

2. CASTEP

CASTEP – Materials Modelling

- **CASTEP** is a full-featured materials modelling code based on a first-principles quantum mechanical description of electrons and nuclei. It uses the robust methods of a plane-wave basis set and pseudopotentials.
- Two versions of CASTEP used in this study, **Version 19.1.1** and the current academic release of CASTEP, **Version 21.1.1**.
- Parallelisation over g-vectors leads to a global data exchange to transpose the FFT grid in 3-dimensions i.e., **MPI_alltoallv.**

- **Al3x3 Benchmark**

  The al3x3 simulation cell comprises a 270-atom sapphire surface, with a vacuum gap. There are only 2 k-points, so it is a good test of the performance of CASTEP's other parallelisation strategies.

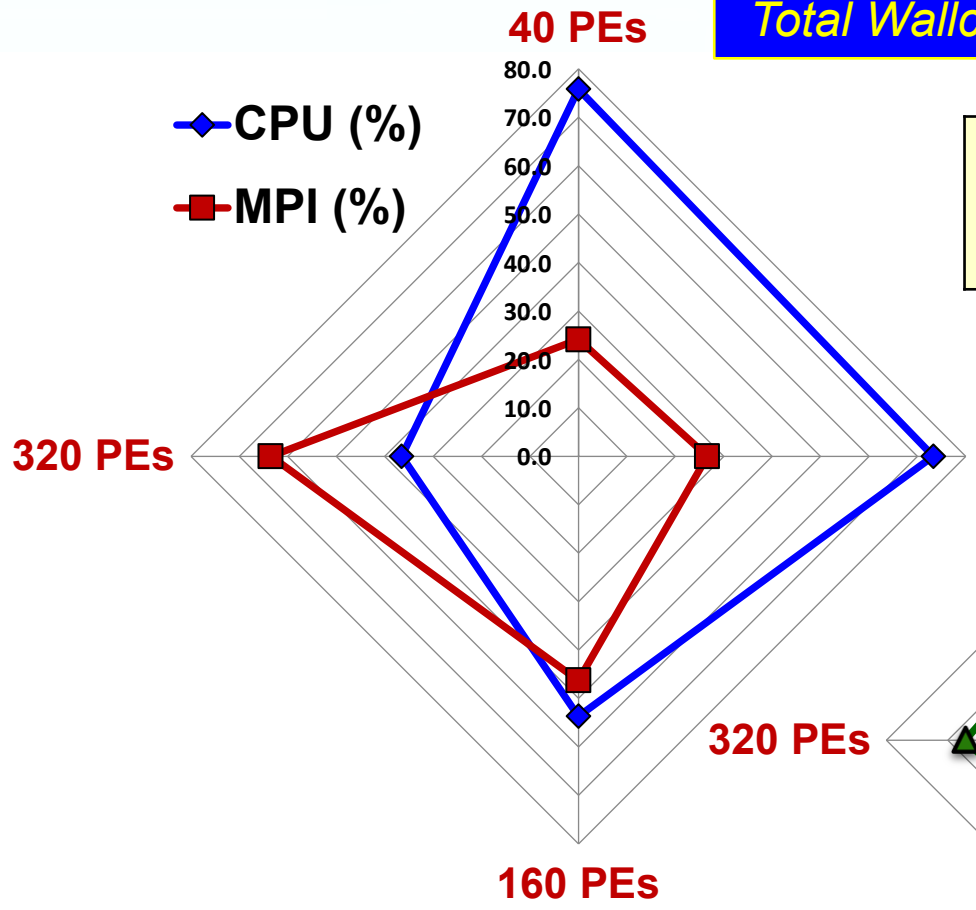
- **MnO_2 Benchmark**

  Bigger calculation (313 electrons and 64 ions) and involves MPI AllToAllV across all processors.

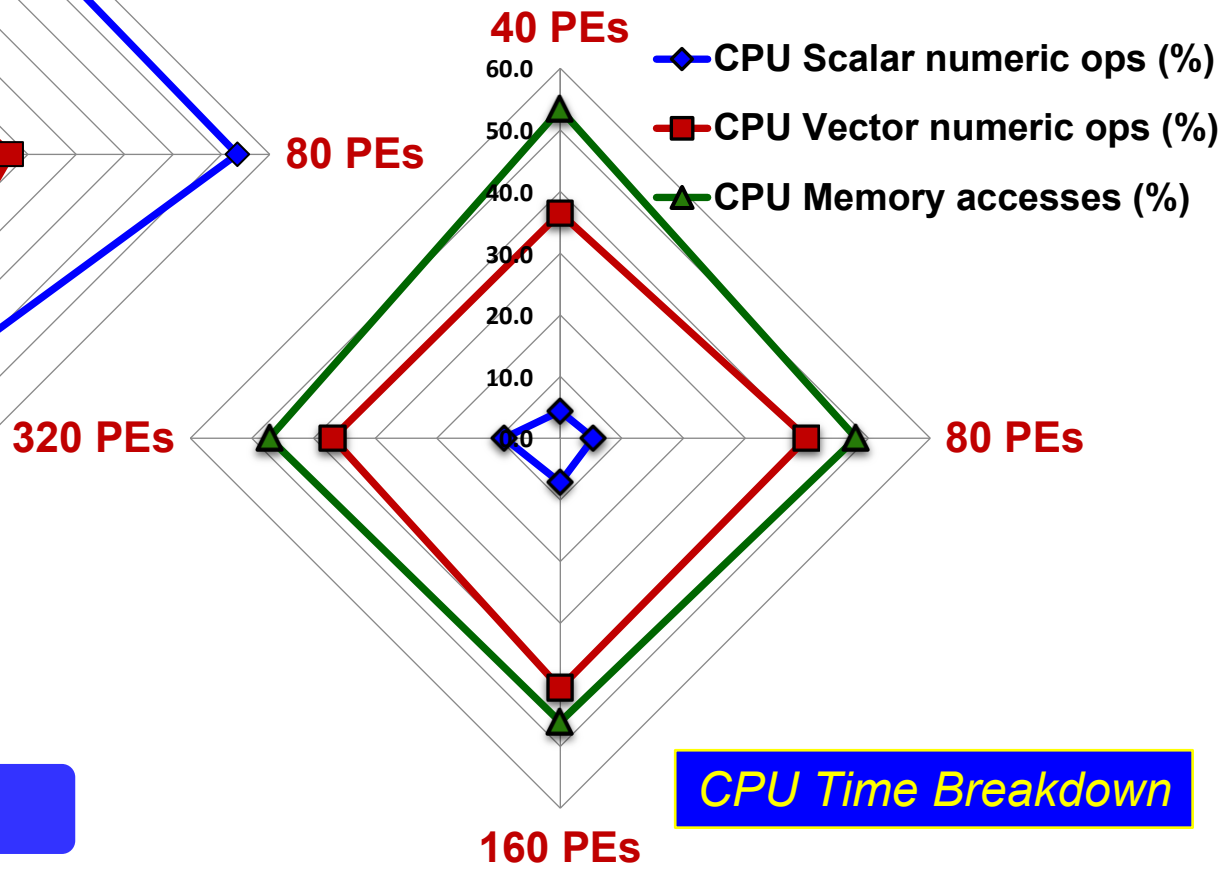
- **IDZ Benchmark**

  Longer MD calculation (1104 electrons and 404 ions) requiring several random initializations (16 MD iterations in total).

CASTEP 21 – al3x3 Benchmark Performance Report

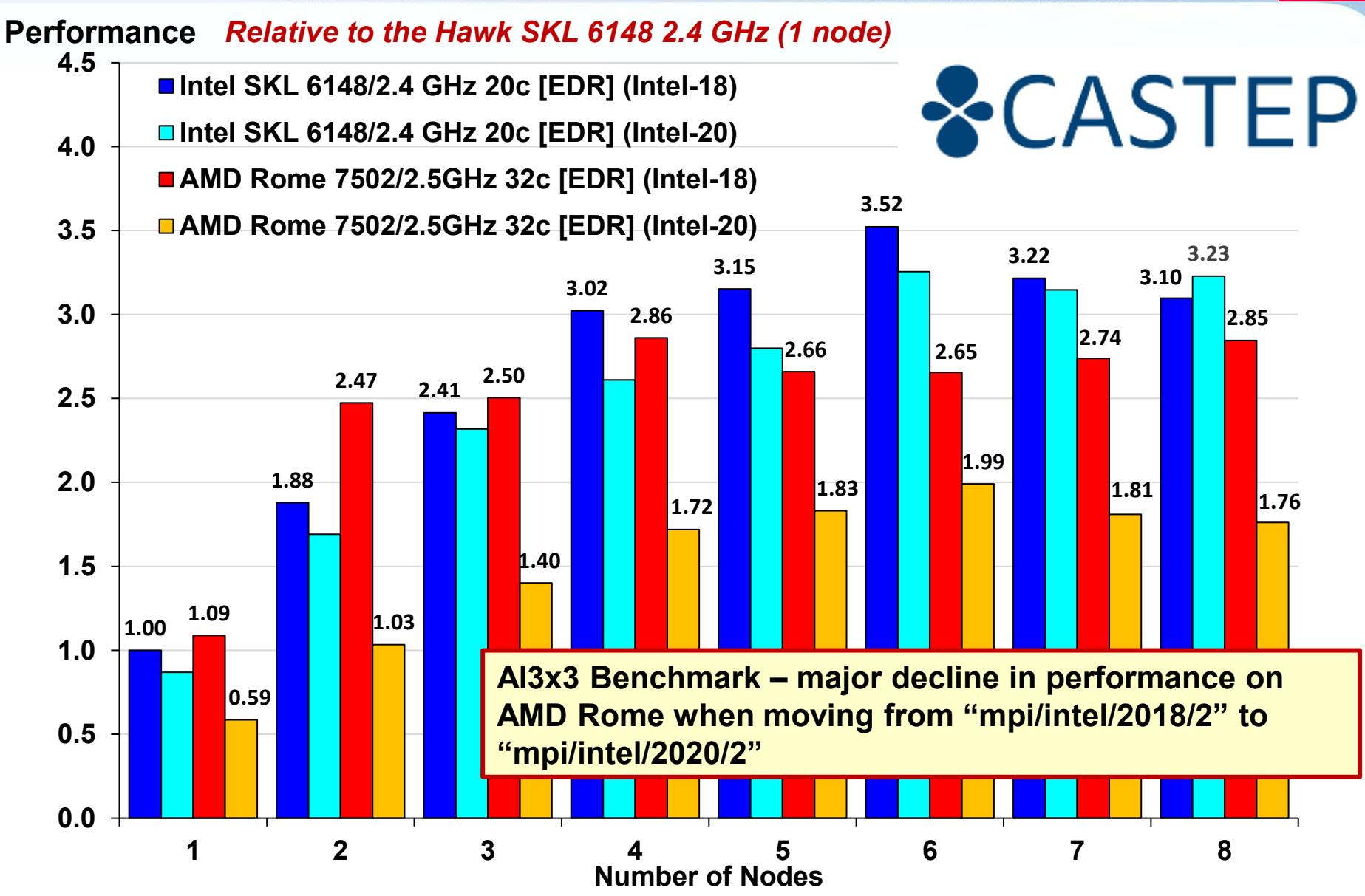


The al3x3 simulation cell comprises a 270-atom sapphire surface, with a vacuum gap. There are only 2 k-points.



Performance Data (40-320 PEs)

CASTEP – Impact of Intel MPI version on AMD clusters

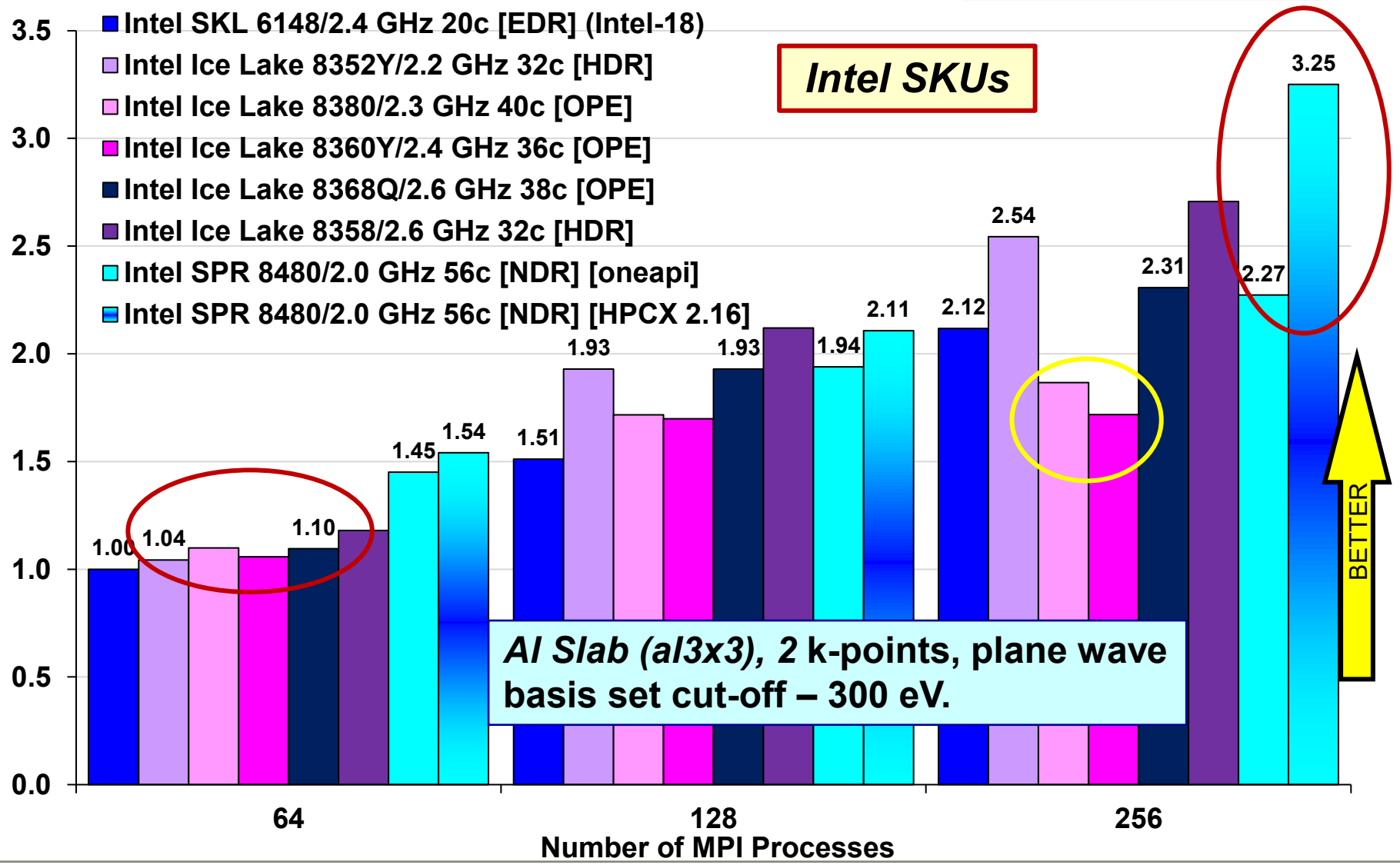


CASTEP 19 – Al Slab (al3x3) Benchmark

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

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

[Core to core]

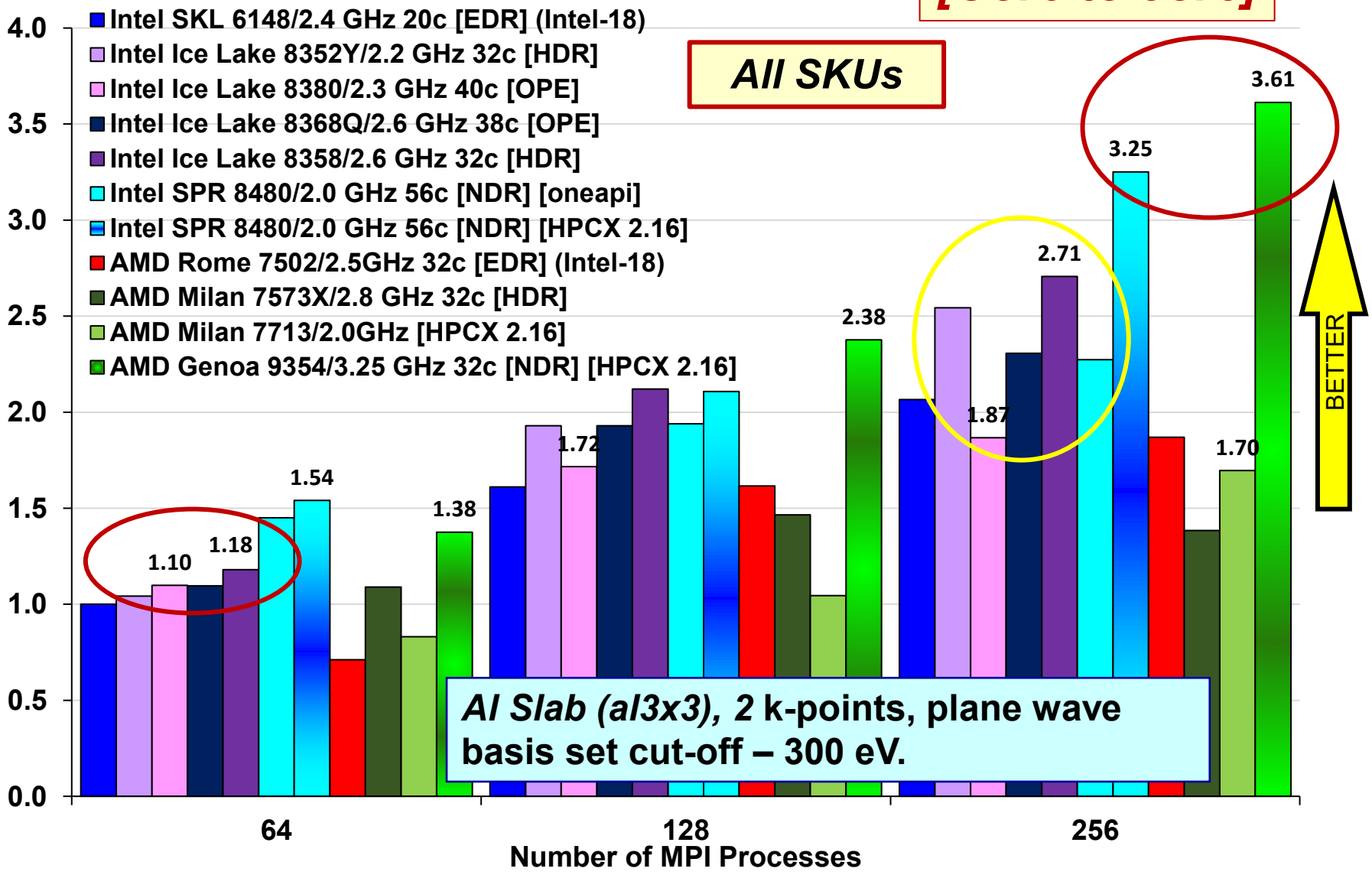


CASTEP 19 – Al Slab (al3x3) Benchmark

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

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

[Core to core]

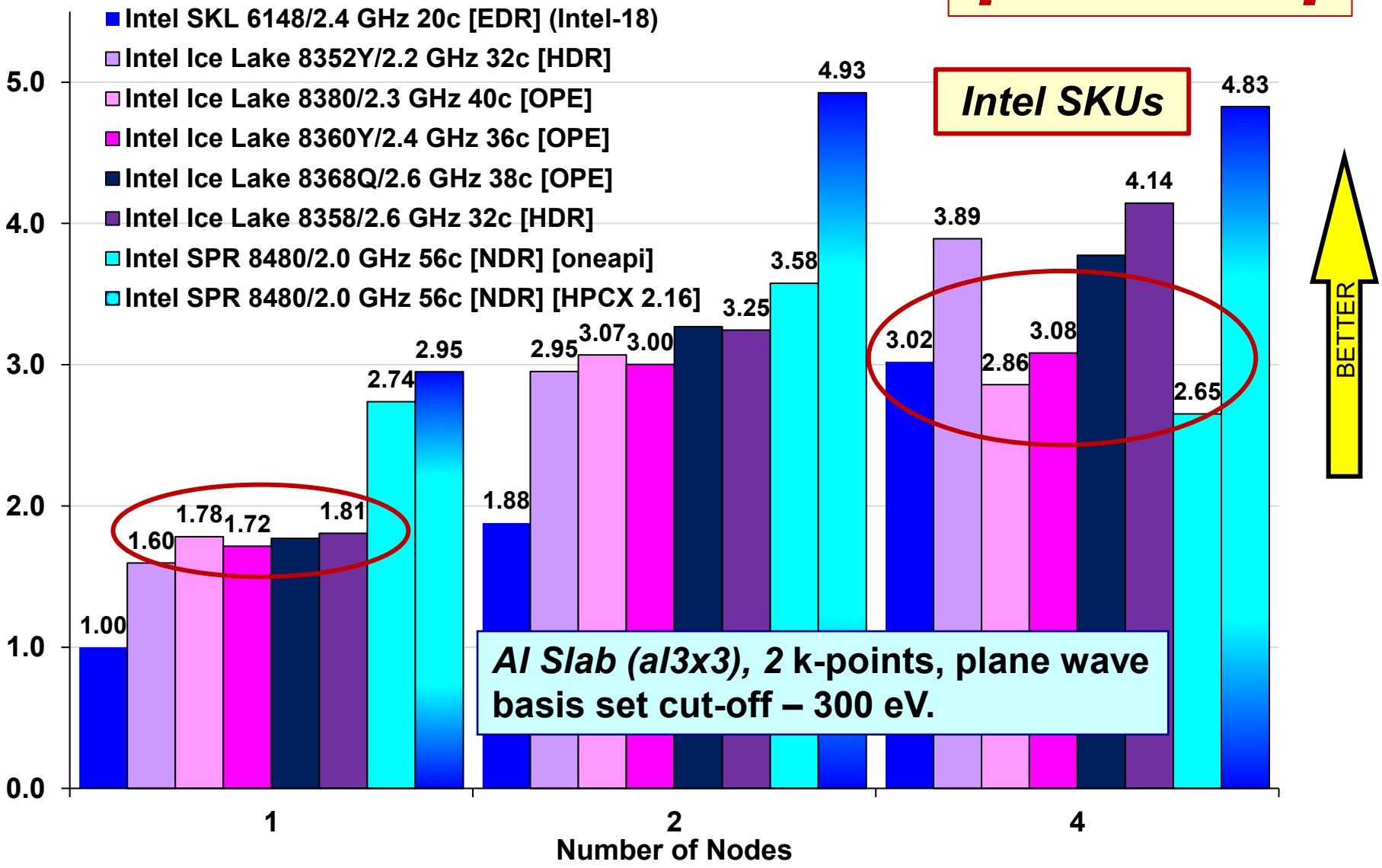


CASTEP 19 – Al Slab (al3x3) Benchmark

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

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

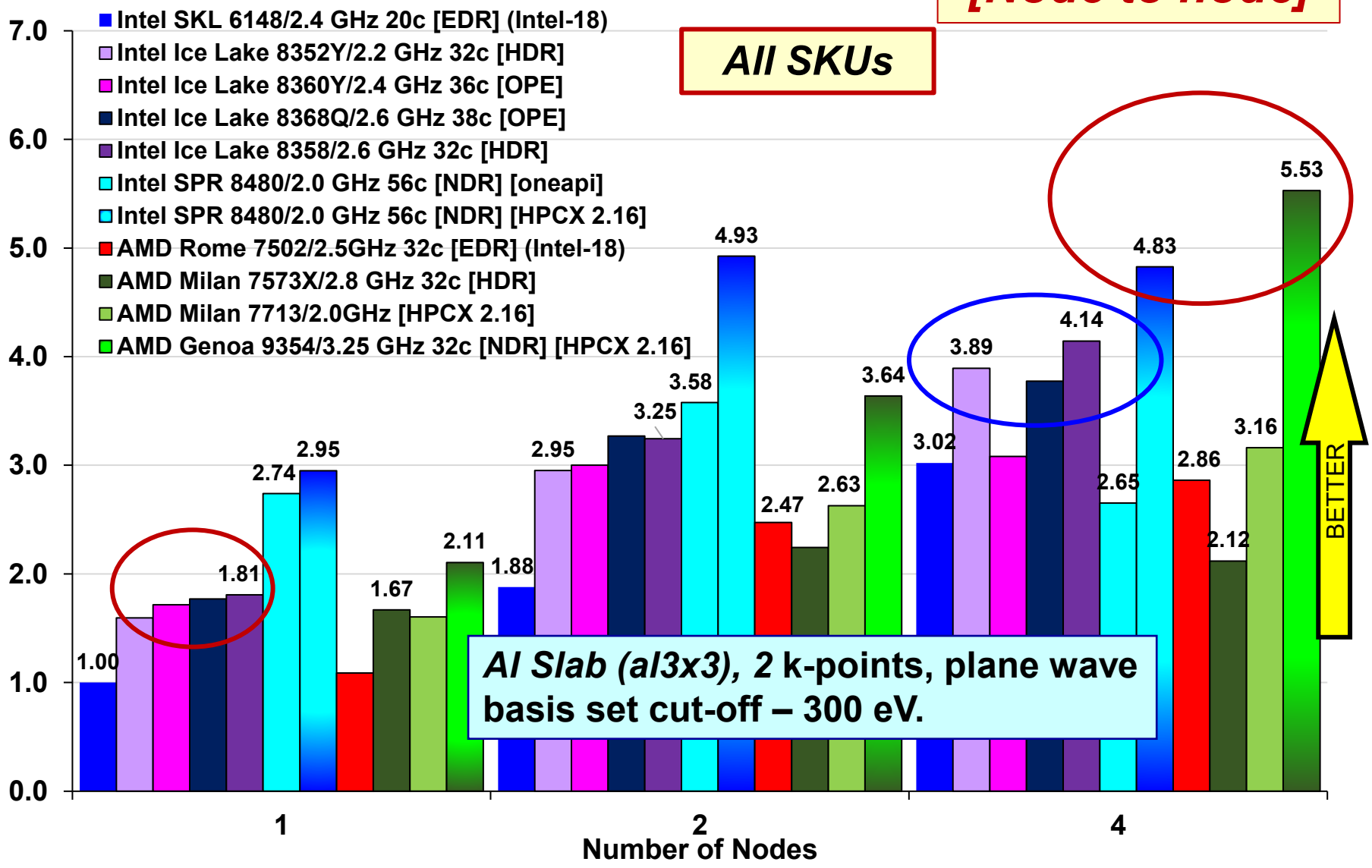
[Node to node]



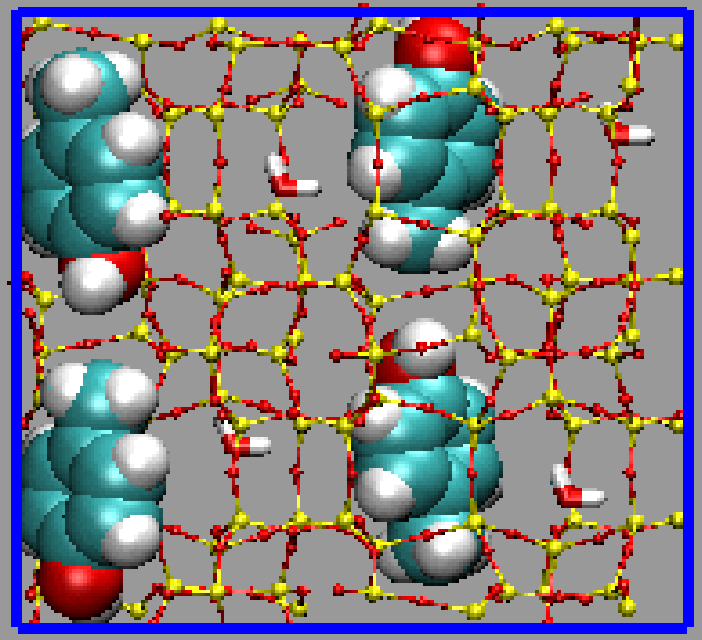
CASTEP 19 – Al Slab (al3x3) Benchmark

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

[Node to node]

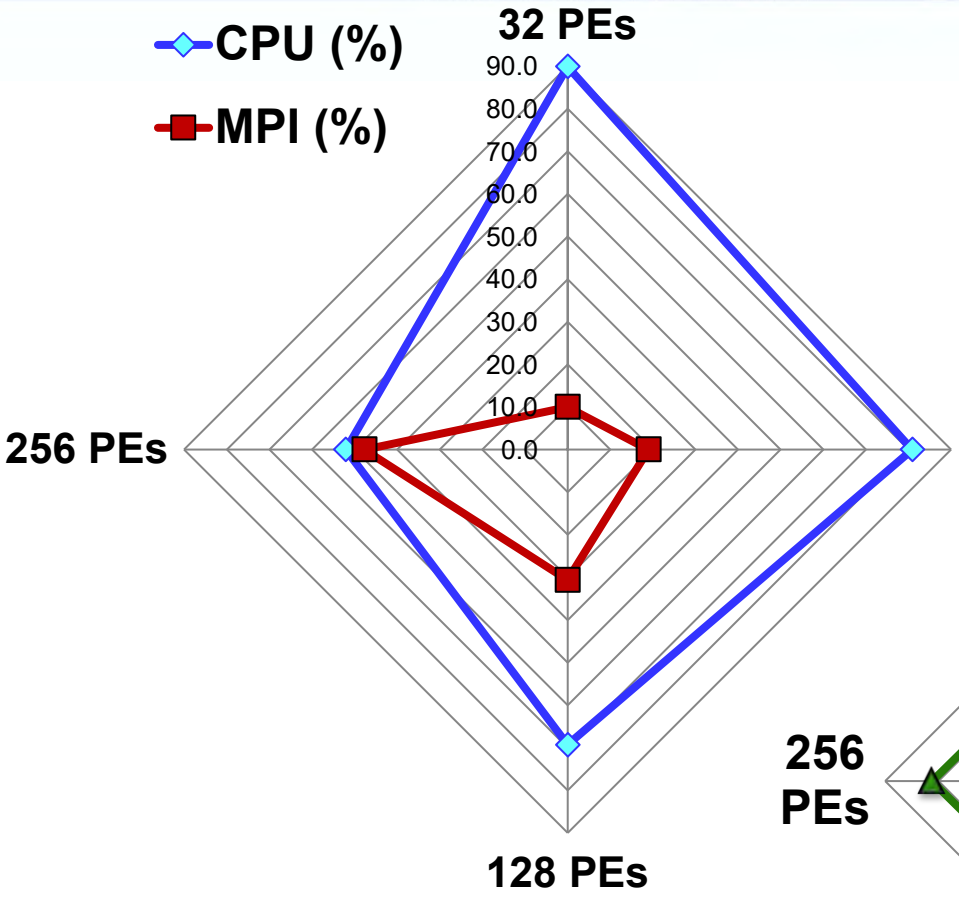


Performance of Computational Chemistry and Ocean Modelling Codes

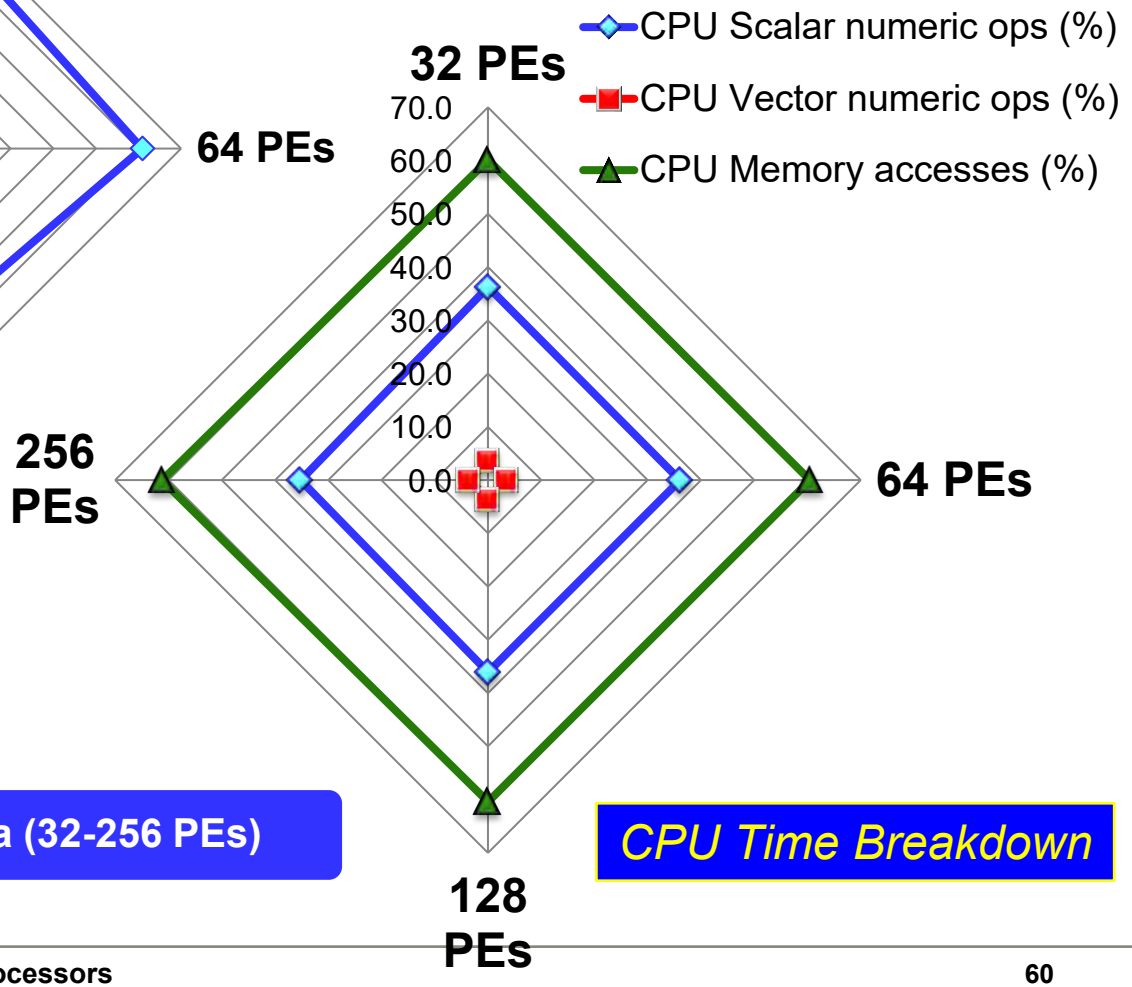


Electronic Structure GAMESS -UK

GAMESS-UK.MPI DFT – DFT Performance Report



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



Total Wallclock Time Breakdown

Performance Data (32-256 PEs)

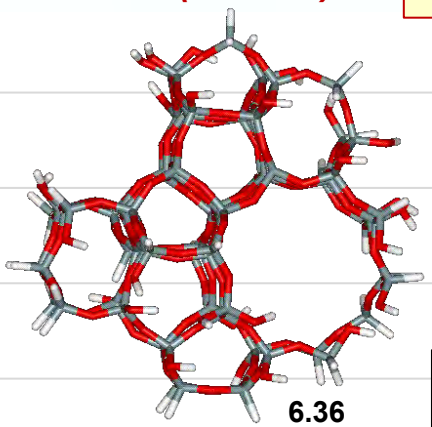
CPU Time Breakdown

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]



- Intel SKL 6148/2.4 GHz 20c [EDR]
- Intel Ice Lake 8352Y/2.2 GHz 32c [HDR]
- Intel Ice Lake 8380/2.3 GHz 40c [OPE]
- Intel Ice Lake 8360Y/2.4 GHz 36c [OPE]
- Intel Ice Lake 8358/2.6 GHz 32c [HDR]
- Intel SPR 8480/2.0 GHz 56c [NDR]
- AMD Rome 7502/2.5GHz 32c [EDR]
- AMD Milan 7713/2.0 GHz 64c [HDR]
- AMD Genoa 9354/3.25 GHz 32c [NDR]

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0

1.80 2.41 2.66 3.15 2.23 3.62 3.29 4.32 4.84 5.79 4.04 6.36 5.53 7.38 8.32 9.70 6.71 9.97

BETTER

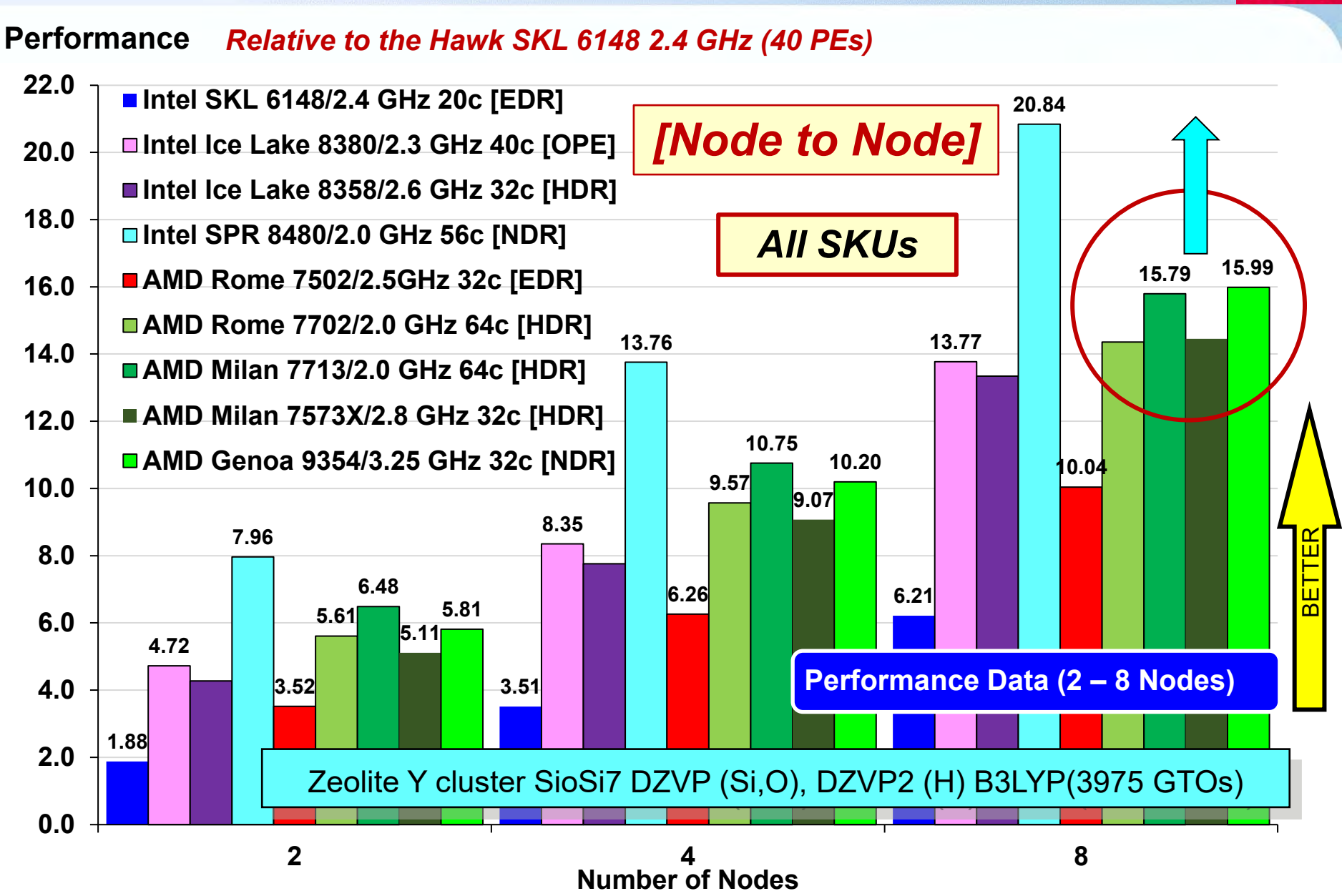
Performance Data (128 - 512 PEs)

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

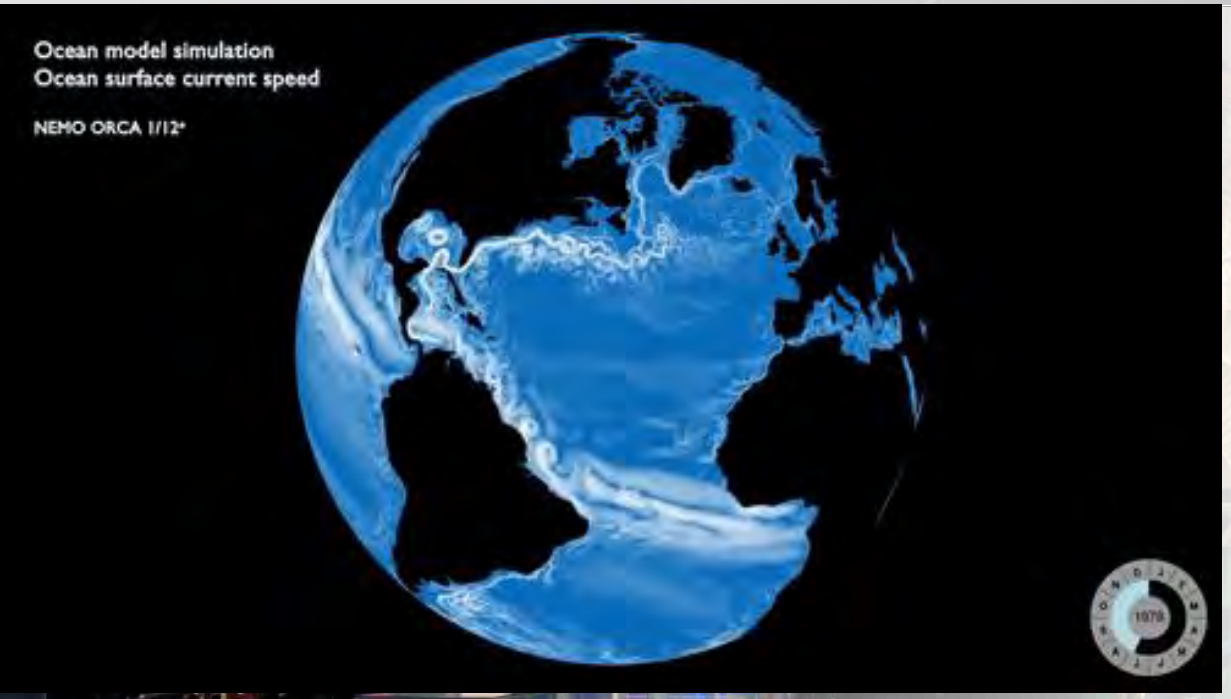
128 256 512

Number of MPI Processes

GAMESS-UK Performance - Zeolite Y cluster



Performance of Computational Chemistry and Ocean Modelling Codes



Ocean Modelling: NEMO and FVCOM

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.**

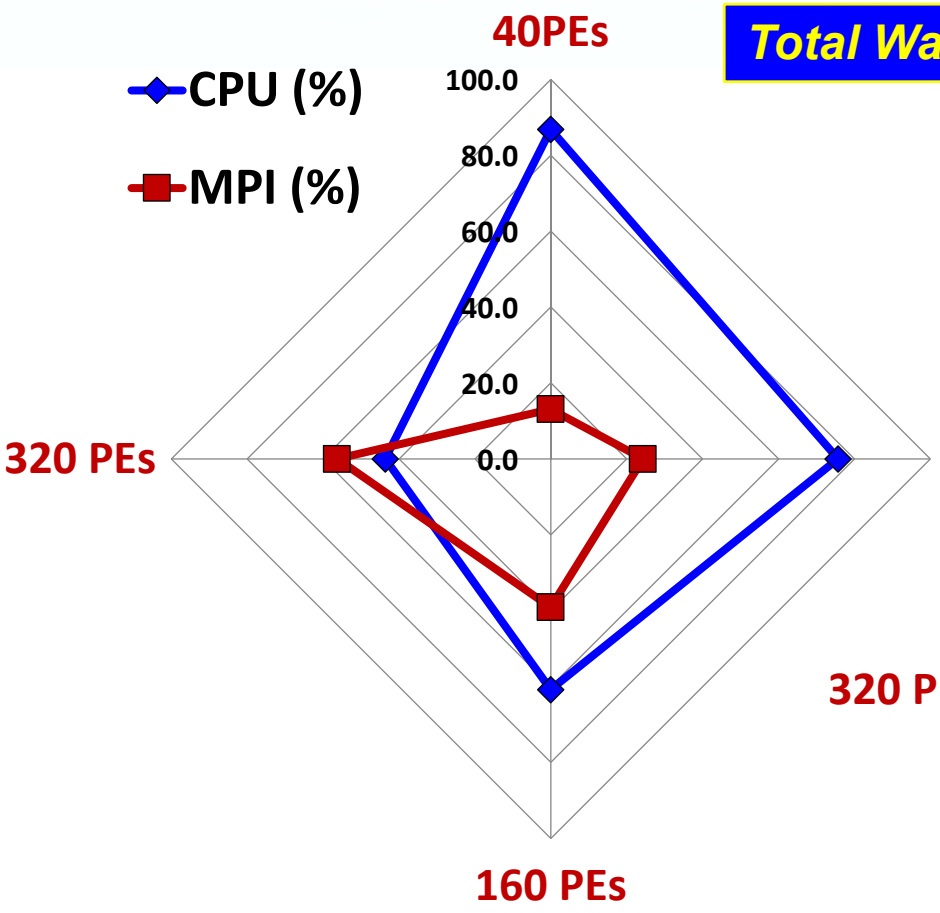
- Both models are often run with a **biogeochemical model called ERSEM -** significantly increases the compute & memory requirements.
- To be run efficiently, both models require a CPU based HPC system

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:

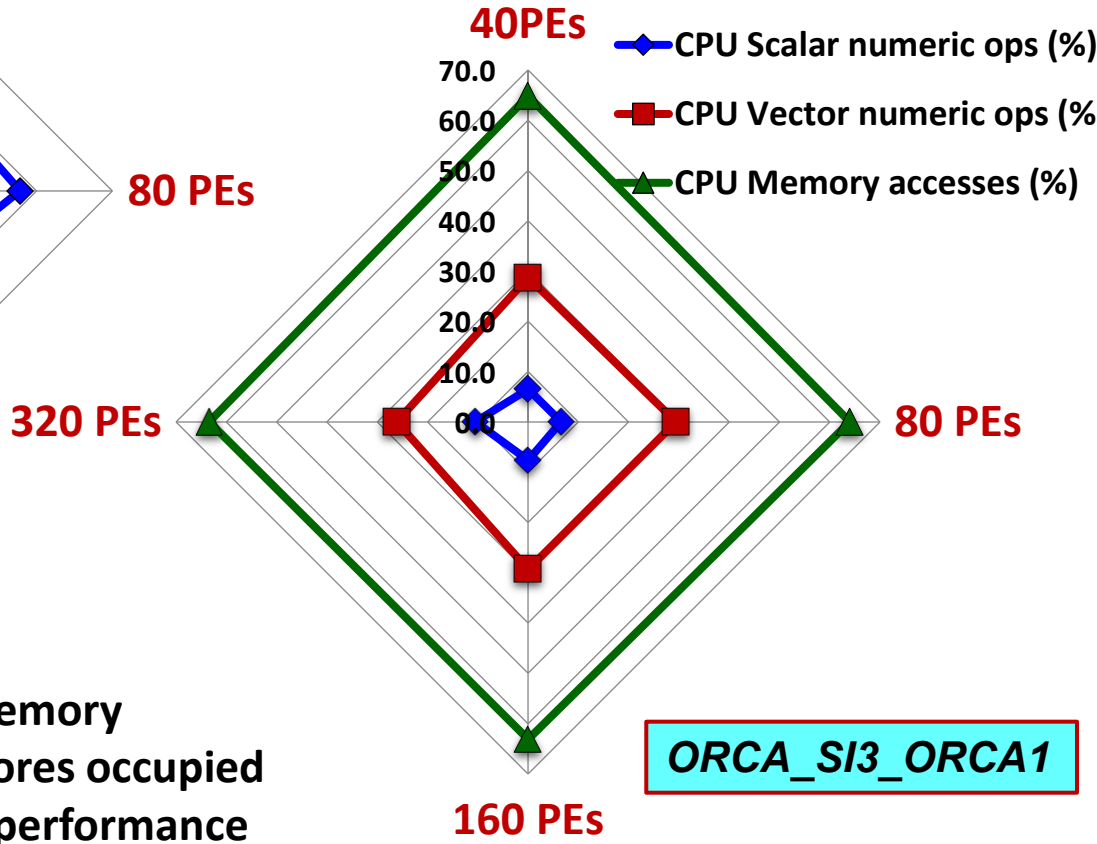
mpirun -n *$XIOSCORES* $code_xios : -n *$OCEANCORES* $code_nemo

NEMO – ORCA_SI3 Model Performance Report



Total Wallclock Time Breakdown

horizontal resolutions of 1-degree

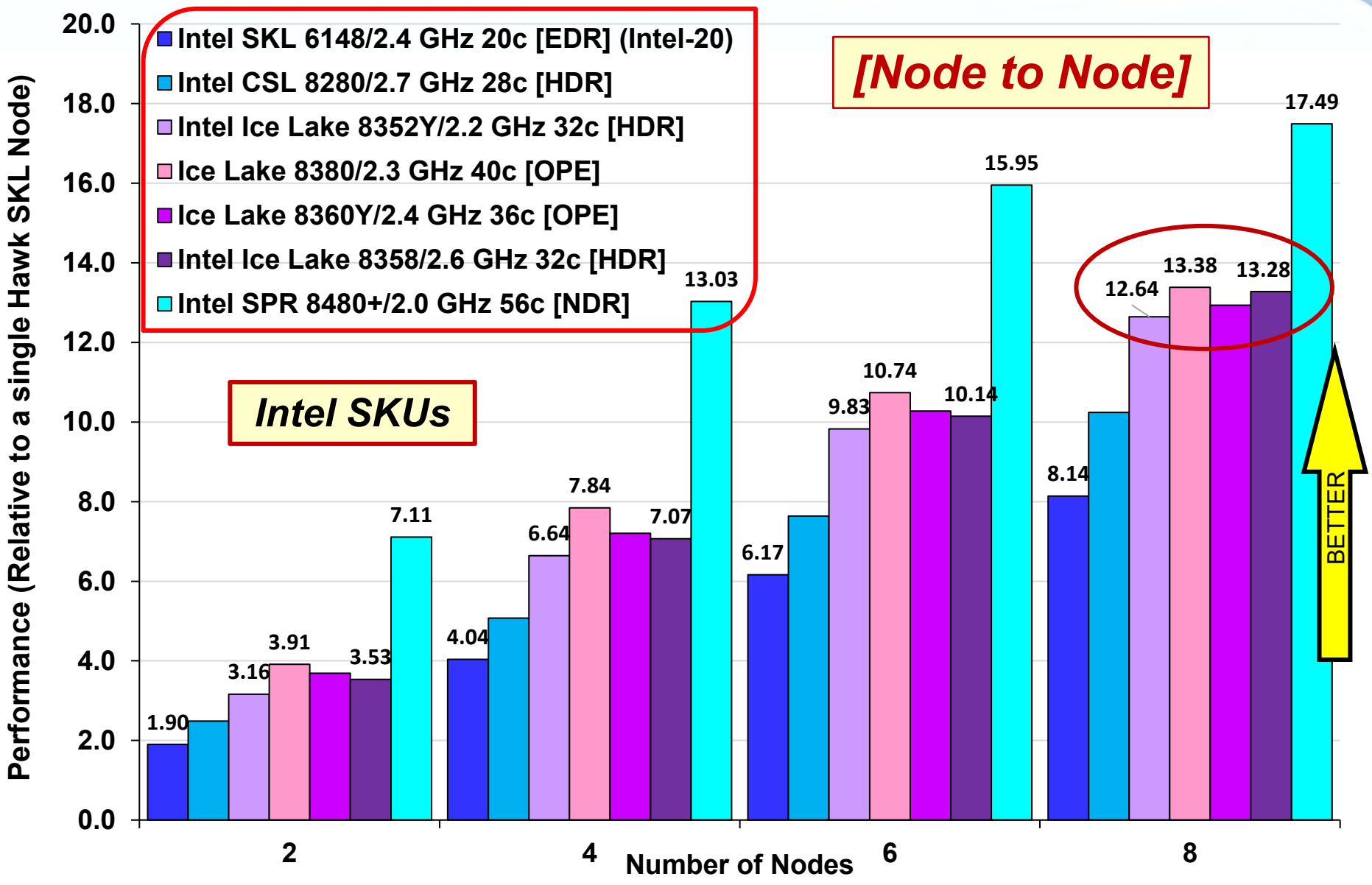


ORCA_SI3_ORCA1

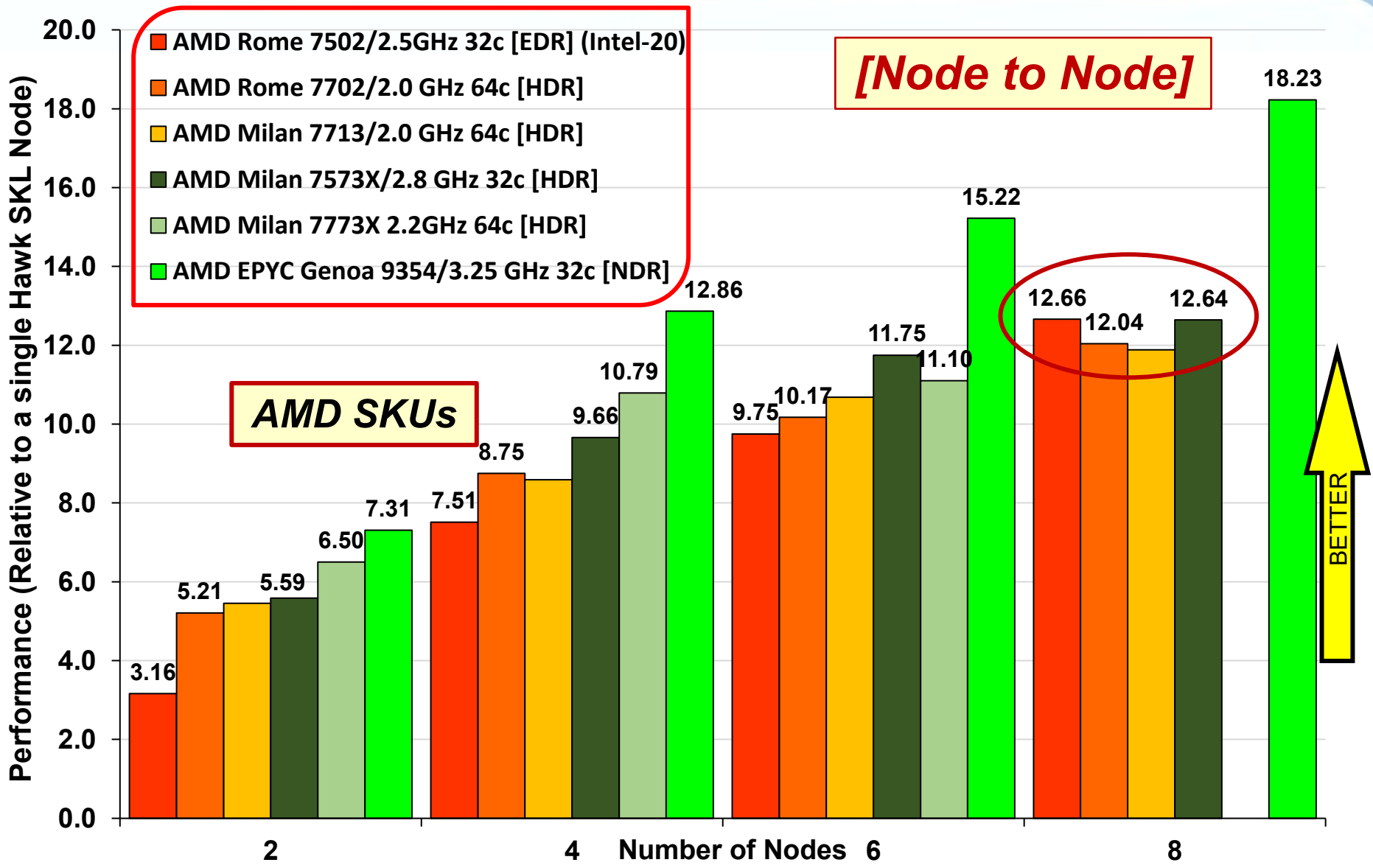
NEMO performance is dominated by memory bandwidth – running with 50% of the cores occupied on each Hawk node typically improves performance by ca. 1.6 for a fixed number of MPI processes.

CPU Time Breakdown

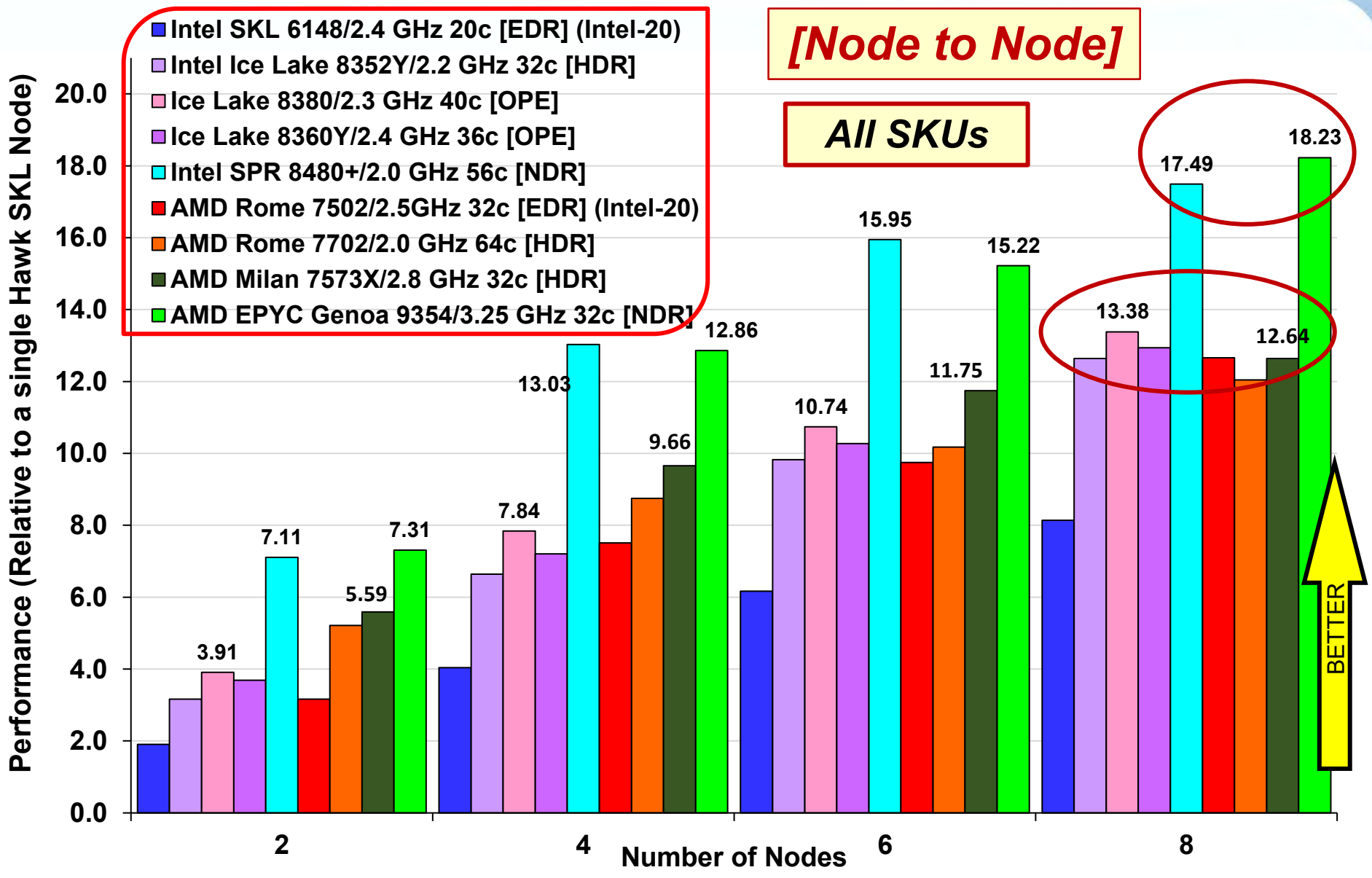
NEMO-FABM-ERSEM (AMM7) – Node Performance



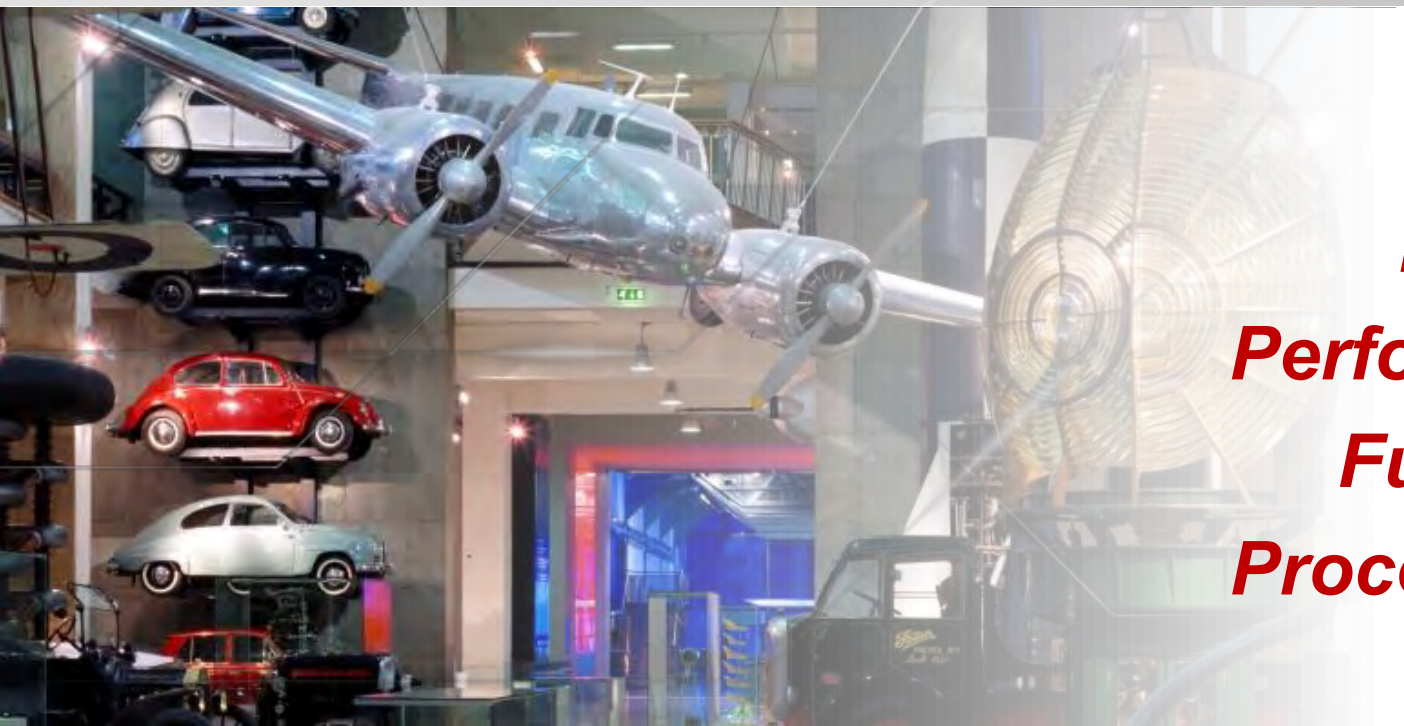
NEMO-FABM-ERSEM (AMM7) – Node Performance



NEMO-FABM-ERSEM (AMM7) – Node Performance

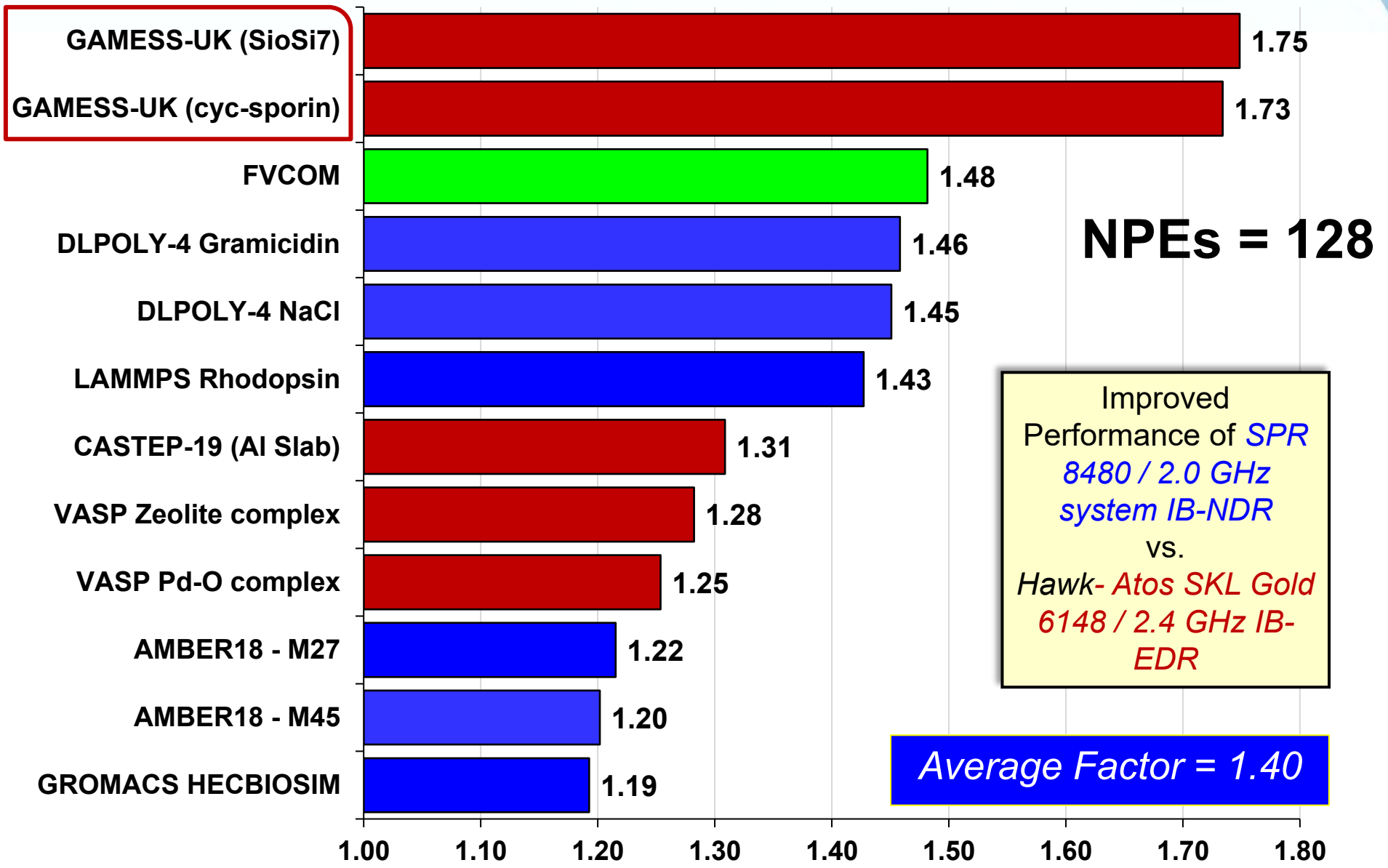


Performance of Computational Chemistry and Ocean Modelling Codes

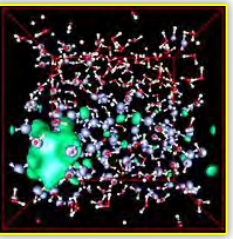


Relative Performance as a Function of Processor Family

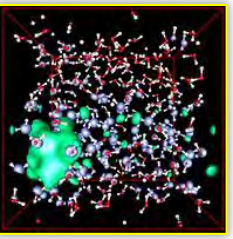
Sapphire Rapids 8480 2.0 GHz NDR vs. SKL 6148 2.4 GHz EDR



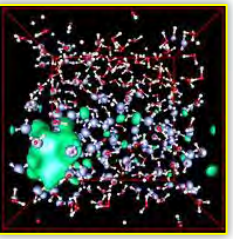
Target Codes and Data Sets – 128 PEs



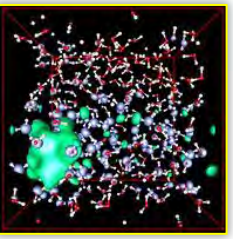
Target Codes and Data Sets – 256 PEs



Target Codes and Data Sets – 2 Nodes



Target Codes and Data Sets – 4 Nodes



4 Nodes Performance [Applications]

Conclusions – Core-to-Core Comparisons

- ***Core-to-Core* comparisons** suggests that the **AMD Genoa 9354 32c 3.25 GHz** outperforms the **Intel SPR 8480 2.0 GHz SKU** in **most cases**, **The exceptions being the** Gromacs 1.4M atom HECBIOSIM & DLPOLY4 NaCl simulations**.**
- The **Intel SPR 8480 2.0 GHz SKU** outperforms all other Intel SKUs (cf. CASTEP), with relative performance sensitive to use of AVX instructions. Low utilisation of AVX-512 leads to weaker performance of the SKL, CSL & Ice Lake CPUs and **better performance of the AMD Milan-based clusters** e.g. DLPOLY, GAMESS-UK, LAMMPS.
- Superior performance of **AMD Genoa 9354** compared to their Milan predecessors.
- Major performance improvement of CASTEP when using the **HPC-X MPI** library on **both Intel and AMD clusters**.
- With **significant AVX-512 utilisation**, Intel **Ice Lake systems** outperform the **AMD Milan systems** e.g.. Gromacs. **Exception** is the **AMD Milan 7573X / 2.8 GHz that outperforms the Intel Ice Lake SKUs** in a number of applications.
- With the possible exception of the **Intel Ice Lake 8358**, there is little to choose between the variety of Intel-based Ice Lake SKUs used in this study.
- Baselined in part across the **V100** NVIDIA GPU performance.

Conclusions – Node-to-Node Comparisons

- Given superior core performance, a ***Node-to-Node comparison*** typical of the performance when running a workload shows the **SPR 8480** delivering **far superior performance** compared to (i) the SKL Gold 6148 (112 cores vs. 40 cores). Average improvements factors of **3.2** (2-node) and **2.8** (4-nodes) across all applications.
- A ***Node-to-Node comparison*** *shows the* ***SPR 8480*** delivering on average ***superior performance*** *compared to* ***the*** **AMD Genoa 9354 32c** (112 cores vs. 64 cores – **1.75**) of **1.30** (2-nodes) and **1.25** (4-nodes). The **NEMO-AMM7** and **CASTEP-19 (Al-slab)** position the Genoa 9354 ahead.
- Performance of the **AMD Milan 7713, 7763 and 7773X (128 core nodes)** is disappointing.
- In contrast to the core-to-core comparisons, the higher core count **Ice Lake systems** – **38c 8368Q & 40c 8380** – now perform on a par with the **32c 8358**.
- Relative to the Ice Lake systems, the **32c AMD Milan 7573X** is ranked first in four of the 4-node application benchmarks.
- **Pricing** – remains of course a key issue, but lies outside the scope of this presentation.

Acknowledgements

- **Joseph Stanfield, Joshua Weage,** Martin Hilgeman, Benoit Lodej, Mark Mendez & Dave Coughlin for access to, and assistance with, the variety of AMD EPYC & Intel Xeon SKUs at the Dell Benchmarking Centre.
- **Toby Smith, Ian Lloyd and Adam Roe** for access to and assistance with the CXL-AP and Ice Lake clusters at the Swindon Benchmarking Lab
- **Erwin James and John Swinburne** for implementing the NETCDF and XIOS-5 libraries on the Endeavour cluster for testing both the NEMO and FVCOM applications
- ***Okba Hamitou, Luis Cebamanos and Chrisophe Bertherlot*** for access to the SPARTAN and Ice Lake & Milan systems (Genji) at the Atos HPC, AI & QLM Benchmarking Centre
- **Jim Clark, Dale Partridge, Gary Holder and Jerry Blackford** at Plymouth Marine Laboratory for discussions on NEMO & FVCOM performance.

Summary

- Focus on systems featuring **processors from Intel** (Sapphire Rapids & Ice Lake SKUs) and **AMD** (Genoa & Milan 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.
  - ❖ **Two** Intel Sapphire Rapids clusters – the 56-core Platinum 8480 and Platinum HBM 9480 plus **five** Intel Xeon Ice Lake clusters, and their Cascade Lake & Cascade Lake-AP counterparts.
  - ❖ **Four** AMD EPYC Milan clusters featuring the 64-core **7713** & **7773X** and the 32-core **7543** & **7573X**. **Two** AMD Genoa systems, the 9354 & 9454.
- **Performance** of both synthetic and **end-user applications,** including molecular simulation (**DL_POLY, AMBER, LAMMPS & GROMACS MD codes**), materials modelling (**CASTEP, VASP**), & electronic structure (**GAMESS-UK**), plus the **NEMO** and **FVCOM** ocean modelling codes.
- **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.

Any Questions?

Martyn Guest *GuestMF@Cardiff.ac.uk*

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

Thomas Green *Greent10@cardiff.ac.uk*

CIUK 2023 Keynote Presentation

Melyssa Fratkin (Industry Programs Director, Texas Advanced Computing Center, UT Austin)

Women in Advanced Computing: Leveling the Playing Field

Abstract: The HPC industry is changing rapidly. We are all faced with expanding our horizons and finding talented staff to keep up with the changes – changes that can bring new insights and different perspectives. But 'Leveling the Playing Field' in HPC is not just about hiring from a more diverse talent pool. We also need to talk about Equity, Inclusion, and Belonging, or we'll never escape the status quo. What strategies can we employ to find that new talent? And how can we build a more inclusive culture to ensure that our new hires stick around? This talk will offer some strategies and ideas for DEIB efforts that can be implemented in any organization.

Bio: Melyssa Fratkin is the Industrial Programs Director at the Texas Advanced Computing Center (TACC) at The University of Texas at Austin. In this role, she oversees TACC's corporate partnerships and government relations activities. Melyssa focuses on developing and managing strong collaborations between industry and academia with the Science & Technology Affiliates for Research (STAR) program.

Melyssa is the founding co-chair of Texas Women in HPC, an organization aimed at raising awareness and broadening diversity in HPC by supporting women and under-represented populations in high performance computing in industry, academia and government across the state of Texas. She also serves as Vice President of the AMD HPC User Group, and Communications Coordinator of the ACM Special Interest Group on High Performance Computing (SIGHPC).

Melyssa received a BA from Rutgers University and MBA from the Robert H. Smith School of Business at the University of Maryland.

WOMEN IN ADVANCED COMPUTING: LEVELING THE PLAYING FIELD

Melyssa Fratkin

Co-Chair, Texas Women in HPC

Industry Programs Director, TACC

December 7, 2023 - CIUK

ABOUT ME

- National Cathedral School for Girls
- Rutgers University BA, Spanish
- Universidad de Salamanca (fluent in Spanish)
- University of Maryland Robert H Smith School of Business, MBA eCommerce
- 25+ years in high tech and HPC

TODAY'S TALK

- Women in HPC – stats
- Does Diversity Matter?
- Diversity is Uncomfortable
- Attracting and Keeping New Talent
- Build Trust
- Be Curious!

TEXAS WOMEN IN HPC

- Texas Women in HPC brings together a diverse community of professionals in industry, academia, and government, from the advanced computing community across the state.

- The mission of TXWHPC is to provide a venue for knowledge-sharing, networking, support, and visibility for women and minorities, by engaging in initiatives to raise awareness and broaden diversity in HPC.

NOT MANY WOMEN IN HPC

- Women make up ~17% of our community
- Female attendance peaked at 16% for SC19
- SC22: 13% Female, 0.5% non-binary or other gender identity
- SC23: 14% Female, 0.5% non-binary or other gender identity*

- First paper on female participation in HPC conferences:
  - Women represent only 10% of all HPC authors

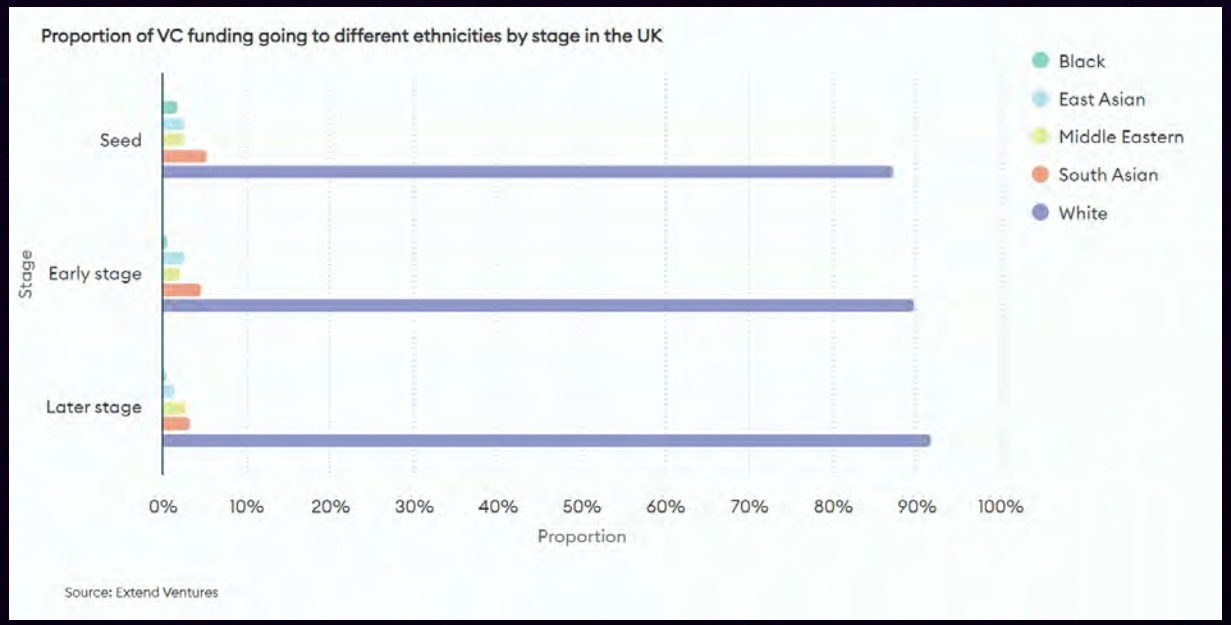
NOT MUCH DIVERSITY EITHER

- Black people make up 12%of the US workforce but only 8 percent of employees in tech jobs, and just 3% of technology executives in the C-suite are Black.
- Black students earned only 7% of STEM bachelor's degrees in 2018, compared with 10%of all bachelor's degrees

- At Intel, 28% of Intel employees are women, while 72% are men. The most common ethnicity at Intel is White (52%). 19% of Intel employees are Asian
- At Dell, 34% of employees are women (2% better than 2021). 10% of employees are Hispanic or Latino, 15% Asian, 6% Black.

TECH DIVERSITY IN THE UK

- **In the UK, 26% of workers in tech are women** (overall it's 50%)
- There are more BAME people in tech than the labour market as a whole, 11.8% for all occupations, and **15.2% for tech**. But 20% of people in the UK are BAME.



- In the UK, BAME Entrepreneurs received a total of **1.7% of VC investment**. On the other hand, 76% of VC investment went to all white founding teams.
- **43% of VC funding** went to teams where at least one founding member was from an elite university, defined as the **University of Cambridge, the University of Oxford, Harvard University, or Stanford University** including their business schools

DOES DIVERSITY MATTER?

“This is a business decision. By 2025, we are going to be a millennial and Generation Z workforce [that is] inclusive and diverse. If your business is not, you are going to get bottom-of-the-barrel workers.”

Increasing diverse participation is not a women's issue or an issue that is *only* relevant to women and other underrepresented groups

WHY DIVERSITY MATTERS - ECONOMICS

- McKinsey found that gender-diverse companies are **25% more likely to outperform** their non-gender diverse counterparts.
  - This is up from 21% in 2017 and 15% in 2014
  - Even in online collaboration, teams who communicated a lot, participated equally, and possessed good “emotion-reading” skills worked smarter.
- Ethnically diverse companies are 36% more likely to outperform their non-diverse peers

NOBEL PRIZE IN ECONOMICS

- Claudia Goldin has documented the journey of American women from holding jobs to pursuing careers
- **Women have outpaced men in education, poured into the labor force, and found meaning in their work.**
- Yet women still lag behind men in pay, in their workforce participation, and the share who reach the top of professions.
- **She has disproved the conventional wisdom that women are paid less because they choose lower-paying careers.**
- “We’re never going to have gender equality until we also have couple equity,”

https://www.nytimes.com/2023/10/11/upshot/claudia-goldin-nobel-prize.html

WHY DIVERSITY MATTERS – THE FUTURE

Who is on your bench?

- Organizations that are developing more high-potential leaders from diverse backgrounds are also:
  - 11X more likely to have high-quality leaders overall.
  - 10X more likely to have a strong leadership bench.
  - 3.2X more likely to engage and retain top talent.

LEGAL/POLITICAL ISSUES

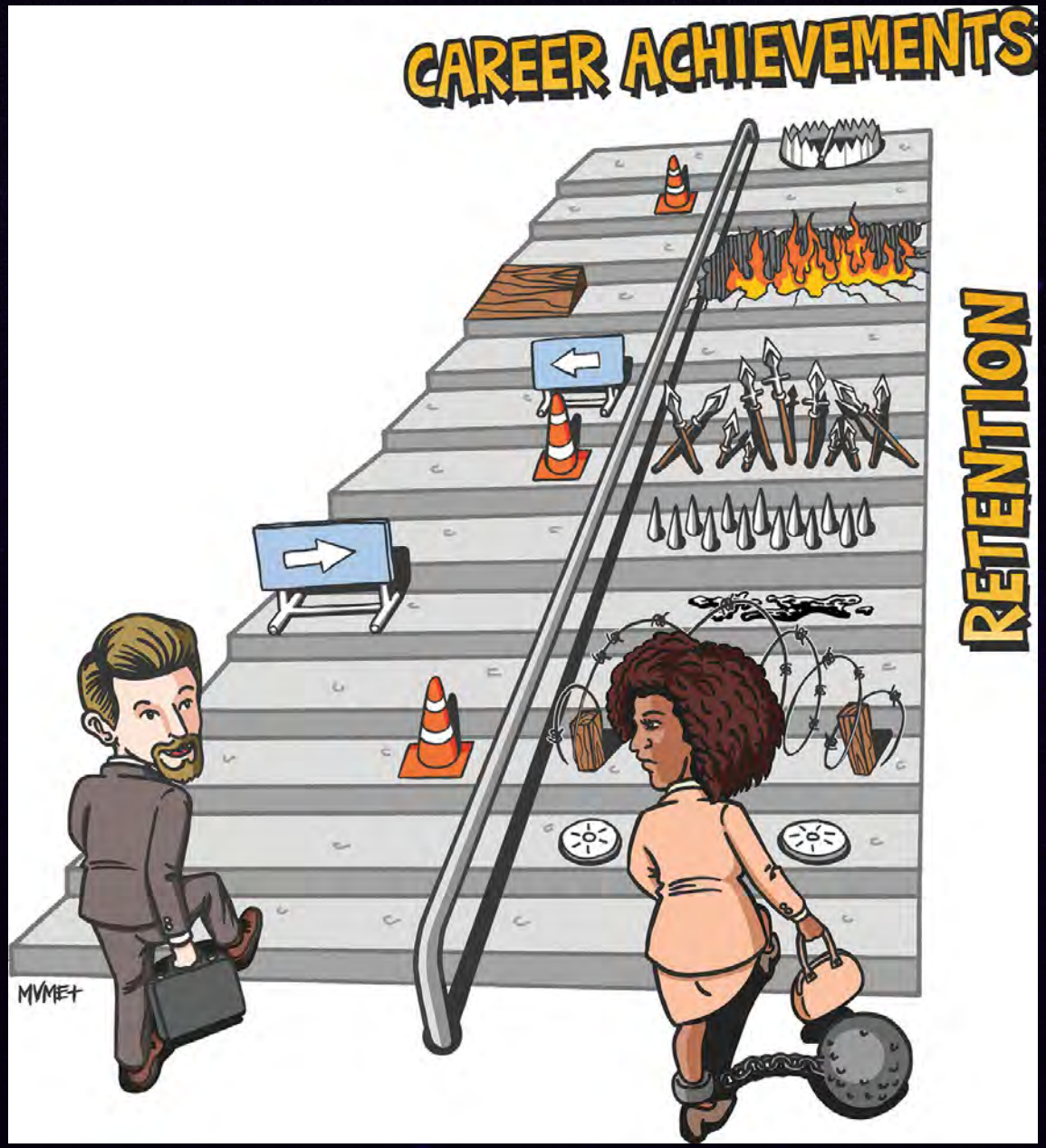


Image From: **Scientists from historically excluded groups face a hostile obstacle course**
Nature Geosciences
23 December 2021
https://www.nature.com/articles/s41561-021-00868-0

LEGAL/POLITICAL ISSUES – WOMEN'S RIGHTS

- There is a strong link between **reproductive rights and workforce participation**.
- In a global market, an empowered worker is one who can migrate.
- Researchers estimated that restrictions on abortion cost states **$105 billion a year**
- Lifting these restrictions could add **half a million women** to the workforce.

LEGAL/POLITICAL ISSUES – THIS TALK!

Leila Saidane/The Texas Tribune

- State Universities in Texas had to **close their diversity, equity, and inclusion offices,** which were used to try to boost faculty diversity and *help students from all backgrounds succeed*.
- It will "undermine the ability of educators and administrators to create a diverse and welcoming campus and threaten the quality of higher education in Texas."
- It "negatively impacts student outcomes," such as retention and graduation rates for students from underrepresented communities.


NOW
WHAT?

FIRST: ADMITTING THERE'S A PROBLEM

- "If I have to hire a 'diverse' candidate, won't the candidates be less qualified?"
- "The HPC community cannot influence diversity – the problem is elsewhere"
- "I don't think we deliberately treat women differently, but I think it's very easy for us to implicitly assume that a decent female student will take a more applied path, while we force her male colleague to start programming."

“ADDRESSING DIVERSITY” IS NOT QUICK

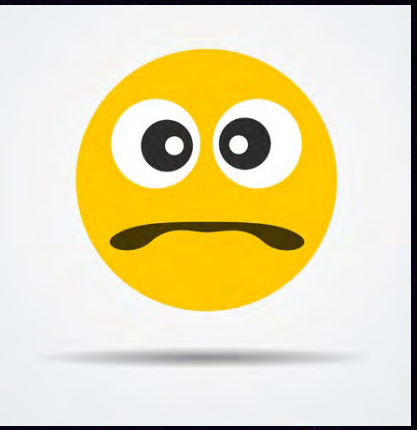
- It’s a multi-level challenge. Start younger, and continue after people are hired.
  - Teenagers (or younger): sponsor summer camps & offer paid internships, apprenticeships
    - *Only 5% of HS students in Texas take CS*
  - Recent graduates: recruit in new places, provide training
  - Current employees: support engagement & mentorship

“ADDRESSING DIVERSITY” IS NOT QUICK

- A commitment to increasing/improving your organization’s diversity **has to come from the top**, with allies at all levels. And it must really be a commitment, not just lip service. It has to be baked into your culture.
- Applicants are looking for people who look like them, who are *succeeding* in the company.

- Inclusion and equity work are actually about **reorganizing the world around us — and re-distributing power and resources**. It’s hard work.

DIVERSITY IS *UNCOMFORTABLE*



- Homogenous teams just *feel* more effective.
- **Working on diverse teams produces better outcomes** *because* it's harder -- no pain, no gain.
- Requires a strong sense of team and **organizational inclusion.**
- The debate and unfamiliarity that come with diversity are an **important catalyst for creativity** and deep thinking

RECRUITING: HOW TO ATTRACT DIVERSE TALENT

- Have a Diverse talent acquisition team
- Ensure inclusive wording in job descriptions
- No-name Resume Screening
- Emphasize value-based hiring
- Give a sample work test
- Standardize interviews
- N8 CIR Checklist

https://n8cir.org.uk/news/diversity-checklist/

WHERE TO FIND TALENT:

- Society of Women Engineers - http://societyofwomenengineers.swe.org/
- National Society of Black Engineers - http://www.nsbe.org/home.aspx
- MAES – Latinos in Science & Engineering http://mymaes.org/
- National Girls Collaborative Project https://ngcproject.org/
- Girls Who Code – http://www.girlswhocode.com
- Diversify Tech - https://www.diversifytech.com/
- Diversity in Tech UK: https://www.diversityintech.co.uk/
- Women in HPC job listings: https://womeninhpc.org/community/jobs
- Broaden your recruiting & collaborations to include HBCUs, MSIs, Tribal Colleges in your area

WOMEN: APPLY FOR THAT JOB!

- Organizations *expect new people to grow into the position*. They want new hires to ask a lot of questions, seek out mentoring, and even make a few mistakes as they get acclimated to a role.
- Look for positions that will stretch you, not ones where you can already tick all the boxes. Other people have had to learn on the job too – they don't know everything.
- **Ignore the Imposter Syndrome (or Discriminatory Gaslighting, or your own inner demons) and APPLY!**

“CULTURE ADD”

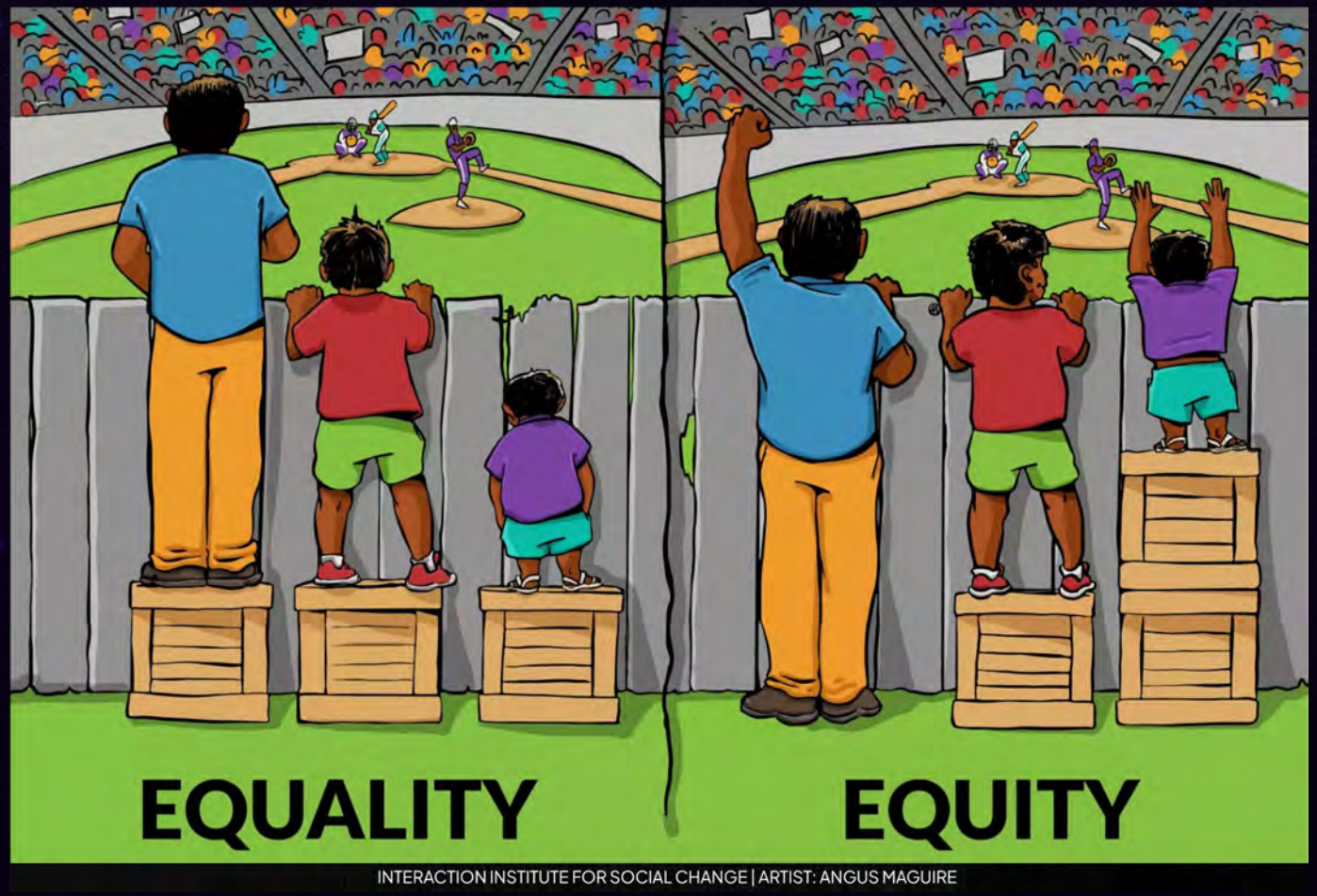
- Move away from “Culture Fit” and towards “***Culture Add***” – Look for people who are different and will add diversity to your team.

- Engagement affects the bottom line. Engaged employees generate more sales.
- Gen Z is 57% more likely to say that diversity, culture, and environment are important.

THEY'RE HERE!...

NOW WHAT?

EQUITY IN THE WORKPLACE



- Equality means that all employees have access to the same opportunities, resources, and treatment.
- **Equity** recognizes that each person has different circumstances and allocates the exact resources and opportunities needed to reach an equal outcome.
- Companies with highly engaged employees are more productive and more profitable.
- Equity improves retention and reduces burnout.

MAKING YOUR WORKPLACE MORE EQUITABLE

- **Hire for skills, not degrees**. Can a candidate prove their qualifications with experience and skills?
- **Accommodate health conditions or disabilities**. Remote work is just one possible accommodation. Every person is different, so this will vary based on individual needs.
- **Offer flexible work schedules**. What roles at your company could be accomplished with flexible hours?
- **Provide training during work hours**. Record presentations so any employees who can't attend the live session can watch the replay.

https://insightglobal.com/blog/equity-in-workplace/

NEXT STEP: INCLUSION

- Diversity and Equity are good, but without inclusion, you're left with a room full of people without connectedness and without the necessary relationships that promote belonging.
- Inclusion ties it all together. It promotes meaningful contributions, and it allows for the use of all skills, strengths, and abilities. It develops a sense of identity and belonging.
- Inclusion ensures that the individual feels appreciated, valued, respected and involved.

INCLUSION WILL CHALLENGE YOUR BIASES

Unconscious bias is a hidden preference or predisposition for or against something

- Governs our pre-existing beliefs and attitudes towards particular groups
- More problematic in workplaces dominated by a single gender or group
- Inadvertently disadvantages those who are in an underrepresented group

INVITE PEOPLE TO BE THEIR AUTHENTIC SELVES AT WORK

https://www.youtube.com/watch?v=B6uulHpFkuo

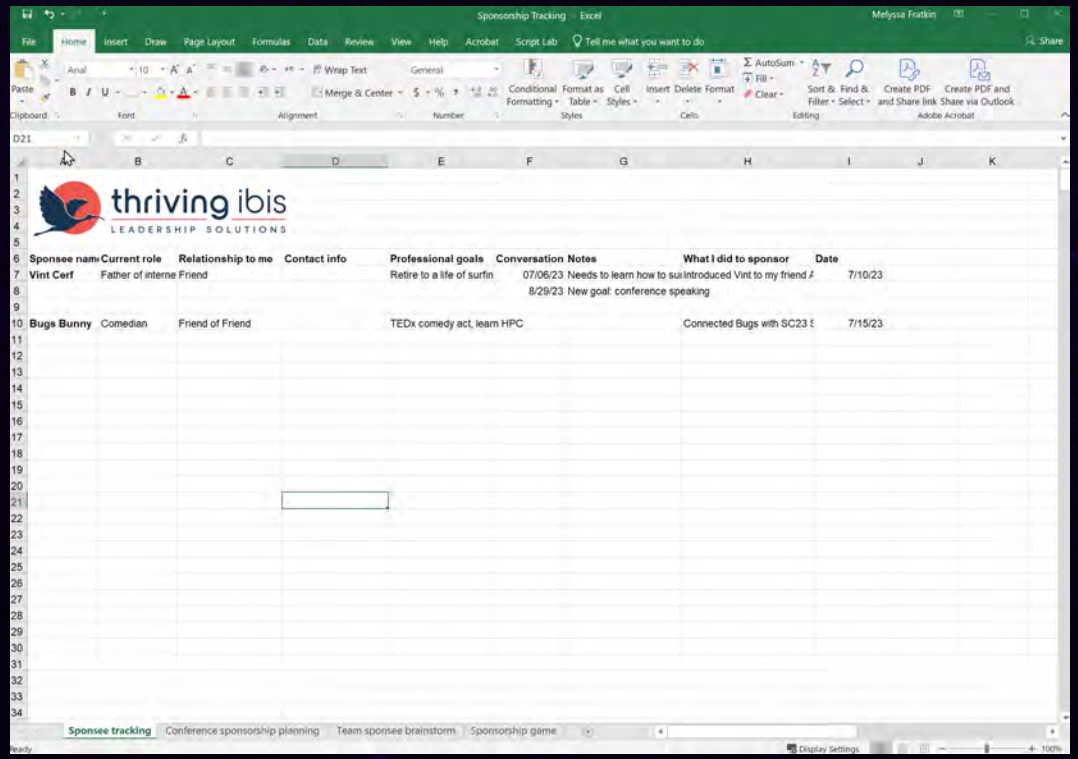
DIVERSITY -> EQUITY -> INCLUSION -> RETENTION

RETENTION EFFORTS – MENTORS & SPONSORS

- A mentor is someone who shares knowledge and provides guidance. A sponsor is someone who actively promotes growth, provides access to opportunities at work, and/or advocates for career advancement
- “Women of color do not need special accommodations to excel (not even close); what we need is the **unwavering belief in our potential to succeed** and being offered opportunities to prove our capabilities”
- “Black women are more ambitious and more likely to say that they want to advance in their companies than their white women counterparts but are less likely to find mentors who will aid their climb up the corporate ladder,”

ACTIVE SPONSORSHIP

- Sponsoring someone's career means mentioning their name when projects or opportunities arise.
- **It means advocating for someone to get a raise or to get sent a job description.**
- With intentionality, it can result in boosting the careers of people who sometimes do not get as many opportunities as others.
- Think about whose career you will talk up while you cruise the exhibit hall and attend the many events we all use to build our network.



RETENTION: HPC CONFERENCES ARE NOT ‘WELCOMING’

- People of different races, ethnicities, or genders have different inclusivity experiences at HPC conferences
- Perceptions of how welcoming HPC conferences and events were indicated a lack of understanding of the experiences of white women and women of color, as well as men of color.

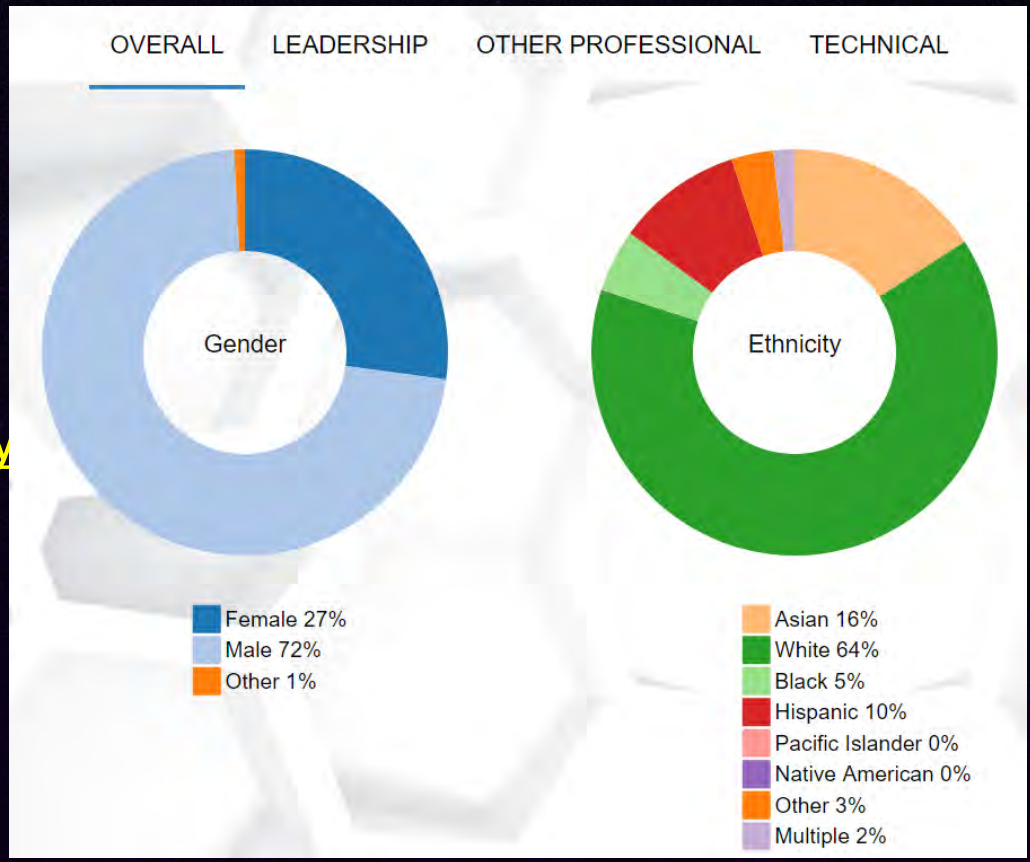
- Different experiences of inclusivity at HPC conferences impact attendees’ thoughts of ***leaving the HPC field***
- Paper on Inclusivity at HPC conferences: https://dl.acm.org/doi/10.1145/3569951.3597580

ACCOUNTABILITY IS KEY

- In order to foster fair, inclusive workplaces, diversity initiatives must incorporate accountability.
- They must be more than “colorful window dressing” that unintentionally angers a substantial portion of the workforce.
- Diversity policies must be researched, assessed for effectiveness, and implemented with care so that everyone in the workplace can feel valued and supported.

DIVERSITY REPORTING

- Create a diversity page
- Google's example at www.google.com/diversity
- TACC's example https://www.tacc.utexas.edu/about/diversity
  - (We are surveying again this year)
- You already know who works for you and where they come from; just start counting
- Provide the data in aggregate to protect the individuals behind the data from a public identification that they may not yet be ready to make



WHAT ELSE CAN YOU DO?

- Pronounce Names Correctly
  - Be comfortable asking for correction
- Look around – notice the demographics of the room
  - What can I do differently? Whose stories and viewpoints are we including and whose are missing?
- Be an Ally!
  - Speak up when you see something, or report it
- Look for and/or Provide Mentorship and Sponsorship
- Review family-friendly working policies (and hybrid work policies)
  - And don't penalize women for using them

THE IMPORTANCE OF MALE ALLIES

THE IMPORTANCE OF MALE ALLIES

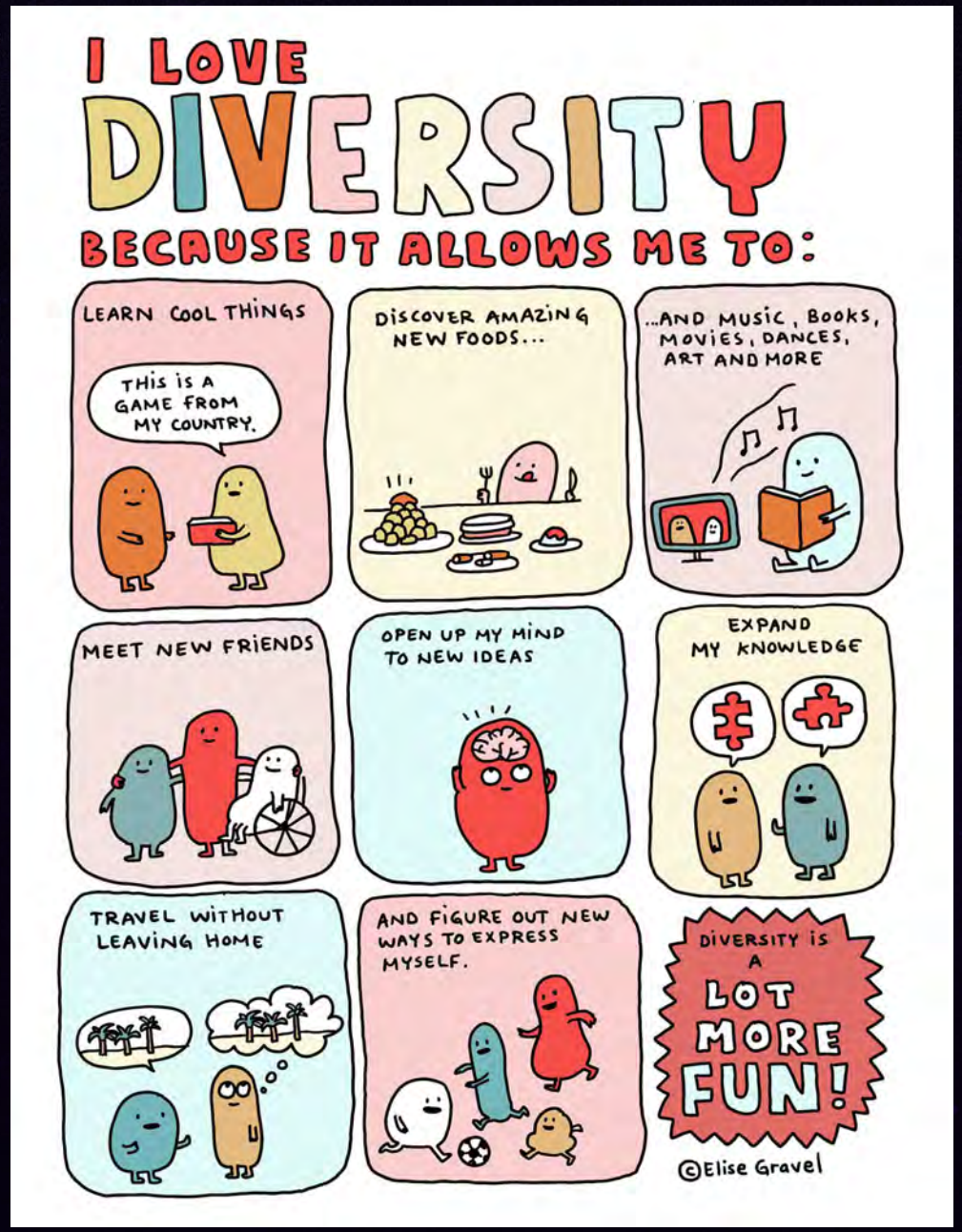
- What prevents men from speaking up for gender equality?
  - *Fear and Apathy*
- **Speak out**, be an advocate, even when it's uncomfortable
- **Share opportunities** (pass up and pass on)
  - *Avoid "Manels"*
- **Acknowledge** the accomplishments of women and people who are underrepresented in HPC at important meetings
- **Volunteer** for non-promotable tasks (taking notes, scheduling meetings)
- **Sponsor someone** – introduce them to the experiences and opportunities that can help them succeed

BE THE PEBBLE IN THE SHOE

- If you feel strongly enough, be the pebble in the shoe. Advocate for diversity programs, modifications to hiring practices, and other activities to improve the workplace.
  - Figure out which angle resonates with management Is it financial? Is it competitive advantage? Planning for the future of the business?
- Be careful – sometimes the pebble has ripple effects, and sometimes it gets thrown away!

READ, RESEARCH, LISTEN

- Newsletter: Ruchika Tulshyan – Inclusion is Leadership newsletter -- https://www.rtulshyan.com/
  - Her book is called *Inclusion on Purpose*
- BCG: How to Advance Gender Diversity in the Workplace
  - https://www.bcg.com/featured-insights/how-to/advance-gender-diversity
- Bizwomen.com
- WomeninHPC.org
- NCWIT – National Center for Women & IT – ncwit.org


I LOVE
DIVERSITY
BECAUSE IT ALLOWS ME TO:
LEARN COOL THINGS
THIS IS A GAME FROM MY COUNTRY.
DISCOVER AMAZING NEW FOODS...
...AND MUSIC, BOOKS, MOVIES, DANCES, ART AND MORE
MEET NEW FRIENDS
OPEN UP MY MIND TO NEW IDEAS
EXPAND MY KNOWLEDGE
TRAVEL WITHOUT LEAVING HOME
AND FIGURE OUT NEW WAYS TO EXPRESS MYSELF.
DIVERSITY IS A LOT MORE FUN!
©Elise Gravel

https://www.txwomeninhpc.org/

https://womeninhpc.org

Thanks!

mfratkin@tacc.utexas.edu

Twitter: @TexasWHPC, @melyssaf

Facebook: https://www.facebook.com/groups/txwhpc

CIUK 2023 Presentations

Dr. Ubaid Ali Qadri (Hartree Centre – STFC)

Industrial Use of High-Performance Computing and Artificial Intelligence: A Hartree Perspective

Abstract: The use of high performance computing can be a great driver for innovation, improved productivity and more generally, economic growth. In this talk, we will focus on the use of HPC by industry to accelerate development, boost innovation, and improve productivity. Firstly, we will discuss the needs that industry have with respect to high performance computing. Secondly, we will discuss how we can address those needs and what we need to take into account when developing software solutions and hardware architectures. Thirdly, we will consider the requirements that AI/ML requirements place on the HPC solutions offered to industry. Finally, we will present examples of use cases where industry has benefitted from the use of HPC/AI solutions through collaboration with the Hartree Centre.

Bio: Dr Ubaid Ali Qadri has a background in computational modelling for fluid mechanics and leads the Multi-Fidelity Design and Twinning team at the Hartree Centre, STFC. He has more than 10 years experience in developing and using adjoint techniques to understand sensitivities and optimal design parameters in multi-physics applications. He has worked across academia and industry on projects related to the aerospace and automotive sectors. At the Hartree Centre, Ubaid leads efforts in the application of novel simulation and design methodologies and high performance computing to industrial challenges.

Science and Technology Facilities Council

Hartree Centre

Welcome

Image ©put image credit here

Science and Technology Facilities Council

Hartree Centre

Industrial Use of High-Performance Computing and Artificial Intelligence

Dr. Ubaid Ali Qadri
Team Lead – Multi-Fidelity Design & Twinning

Agenda

1 Why does it matter?

2 What does ”industry” need?

3 What are we doing?

4 Case studies

Science and Technology Facilities Council

Hartree Centre

Science and Technology Facilities Council

Hartree Centre

Why does it matter?

Why does it matter?

- It is a common belief that improved computing power boosts progress, productivity, and brings about economic benefits.

"Compute has the potential to unlock productivity as sectors across the economy make better and more extensive use of data analysis, simulation and AI technologies.

Future of Compute Review

Why does it matter?

- Is there evidence that more computing power actually produces benefits?

Why does it matter?

- Is there evidence that more computing power actually produces benefits?

Financial ROI Projects

	Average of Revenue $ per HPC $ Invested	Average of Profit or Cost Saving $ per HPC $ Invested
Total	$509.3	$47.2

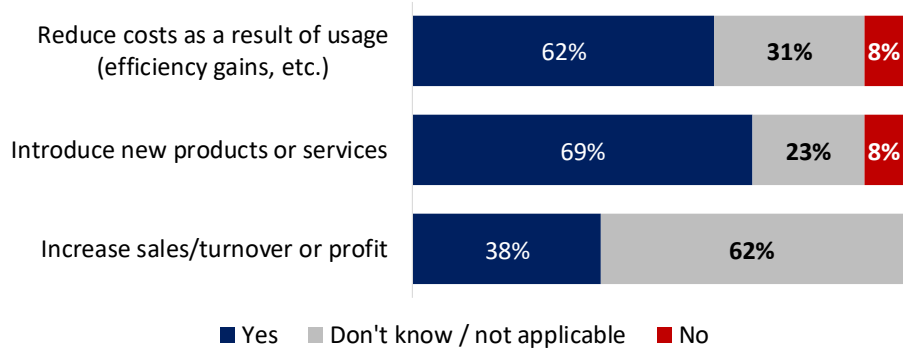
Note: This study analyzed ROI for 26 academic projects, 6 government projects, and 143 industry projects.

Source: Hyperion Research, 2022

The Economic and Societal Benefits of Linux Supercomputers, April 2022, Hyperion Research

The Impact of EPSRC's investments in High Performance Computing infrastructure, Nov 2019, London Economics

Figure 32 Figure 33 'Has access to / usage of EPSRC's HPCs helped your organisation to …?'



Note: Based on 13 responses from **industry**. ***Source: London Economics survey of users of HPC capabilities***

Science and Technology Facilities Council

Hartree Centre

Why does it matter?

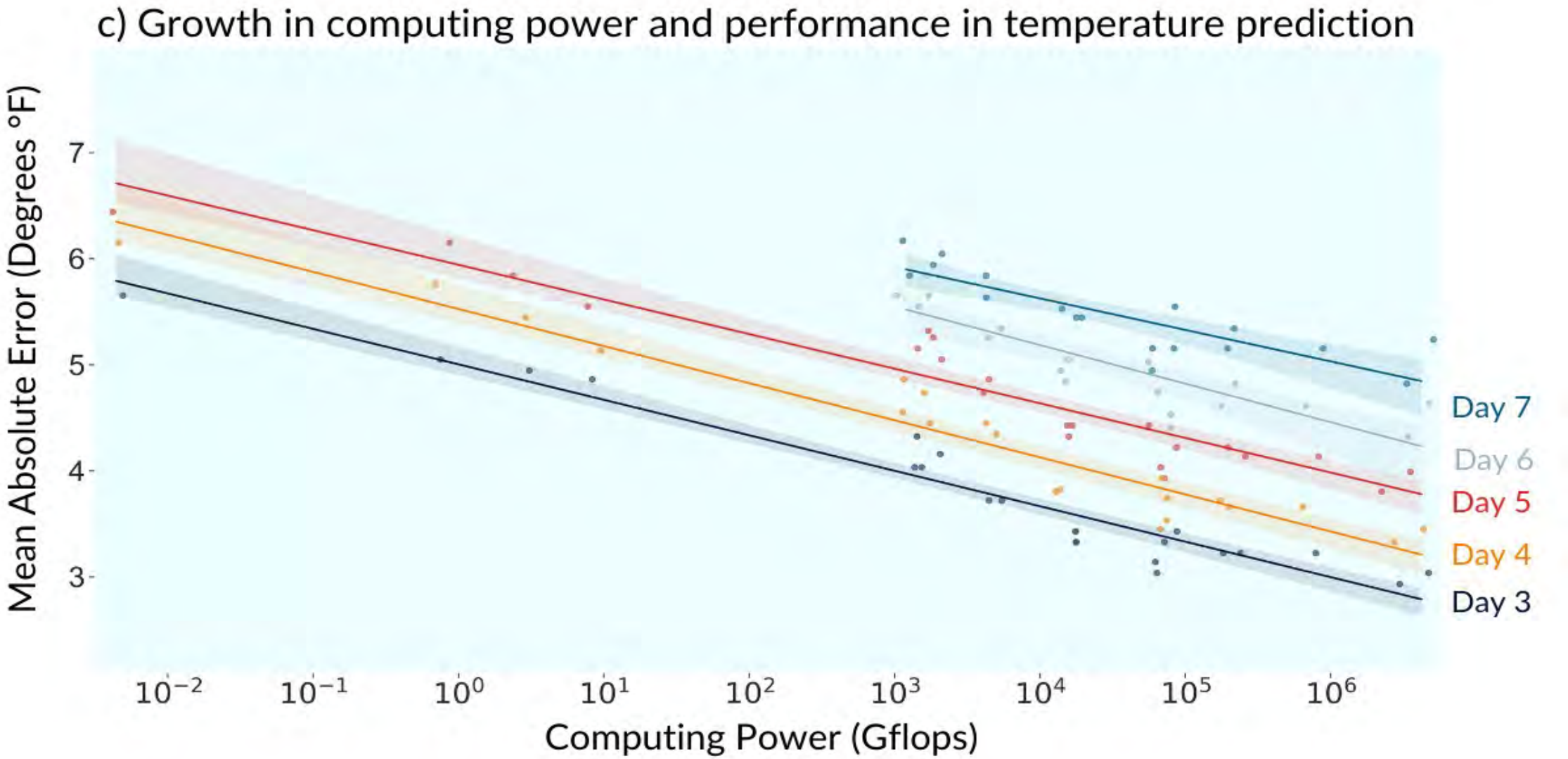
- Is there evidence that more computing power actually produces benefits?

THE IMPORTANCE OF (EXPONENTIALLY MORE) COMPUTING POWER

Neil C. Thompson[1*], **Shuning Ge**[2], **Gabriel F. Manso**[3]

[1]MIT Computer Science and A.I. Lab,
MIT Initiative on the Digital Economy, Cambridge, MA USA
[2]MIT, Cambridge MA, USA
[3]FGA, University of Brasilia, Brasilia, Brazil

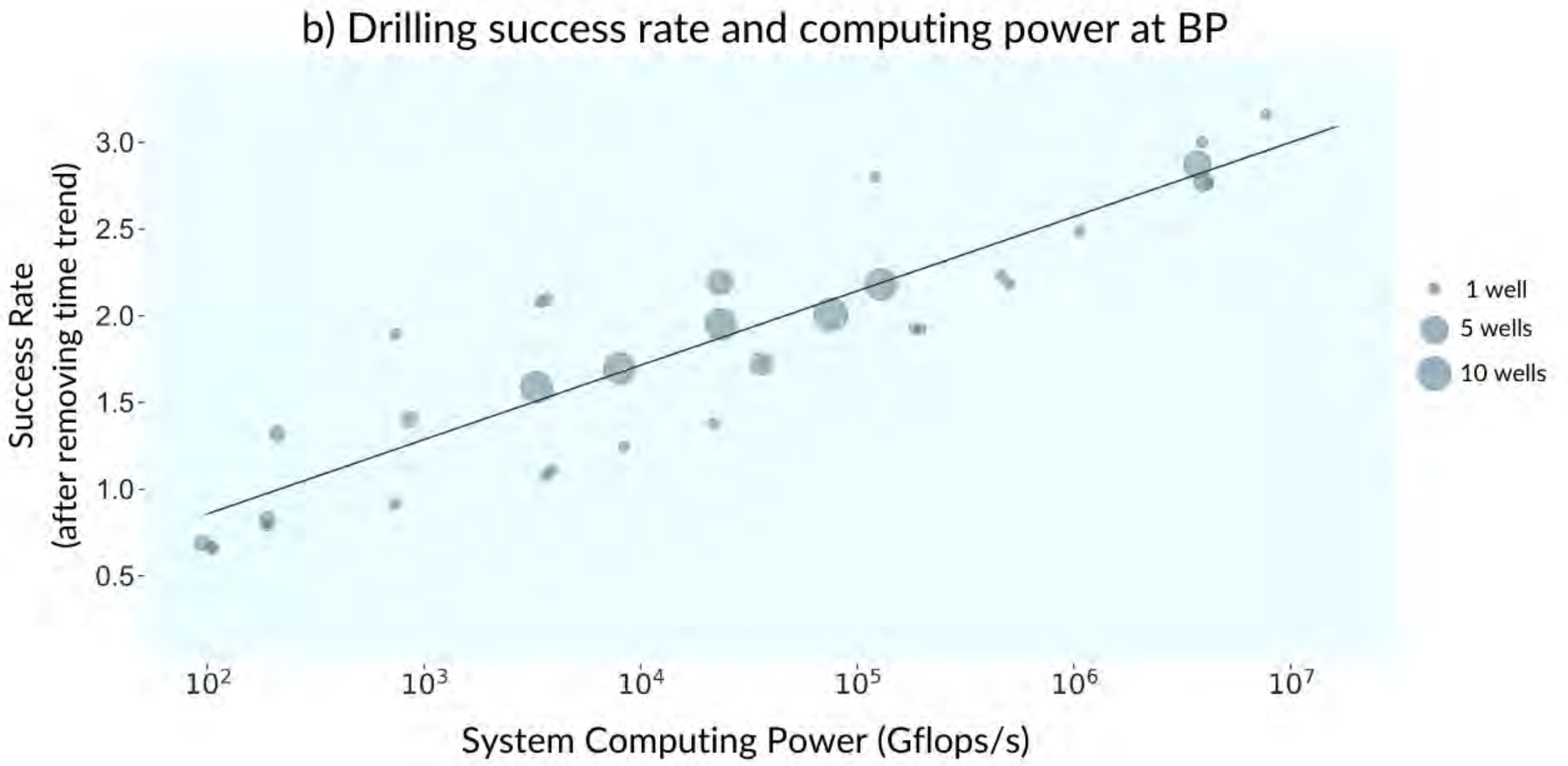
*To whom correspondence should be addressed; E-mail: neil_t@mit.edu.



Science and Technology Facilities Council

Hartree Centre

"The Importance of (Exponentially More) Computing Power", Neil C. Thompson, Shuning Ge, Gabriel F. Manso



Science and Technology Facilities Council

Hartree Centre

"The Importance of (Exponentially More) Computing Power", Neil C. Thompson, Shuning Ge, Gabriel F. Manso

Why does it matter?

- What about AI?

Impact of AI in the innovation process can have more significant effect than impact of AI on final good productions.

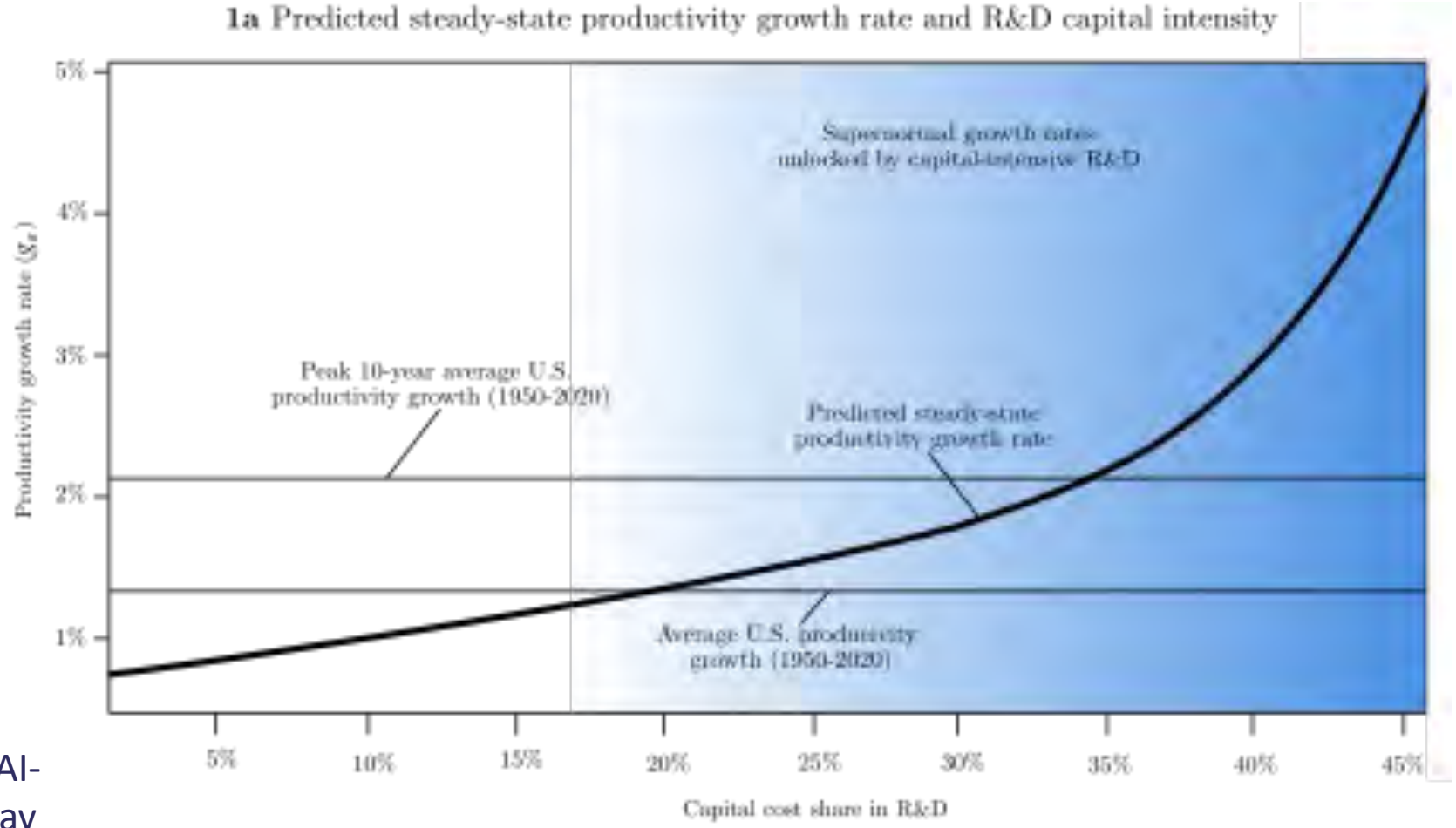
Economic impacts of AI-augmented R&D

Tamay Besiroglu*
MIT FutureTech

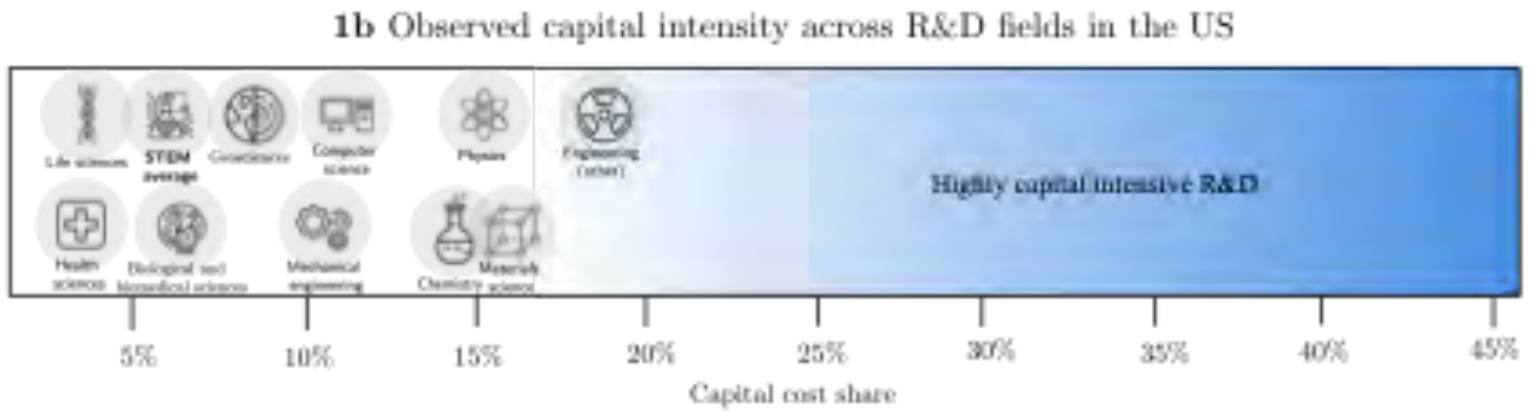
Nicholas Emery-Xu*
UCLA Dept. of Economics, MIT FutureTech

Neil Thompson†
MIT FutureTech

Abstract

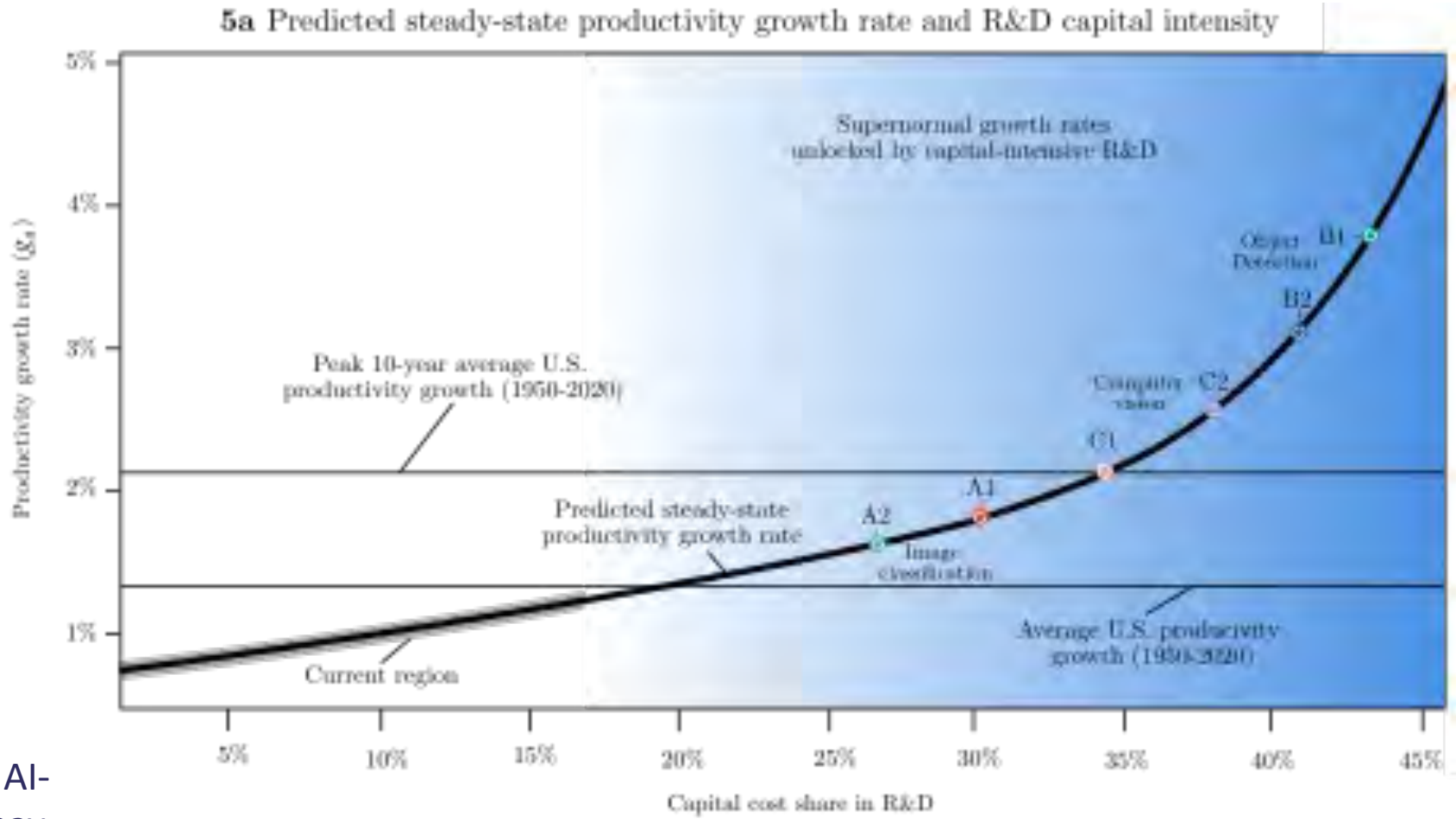


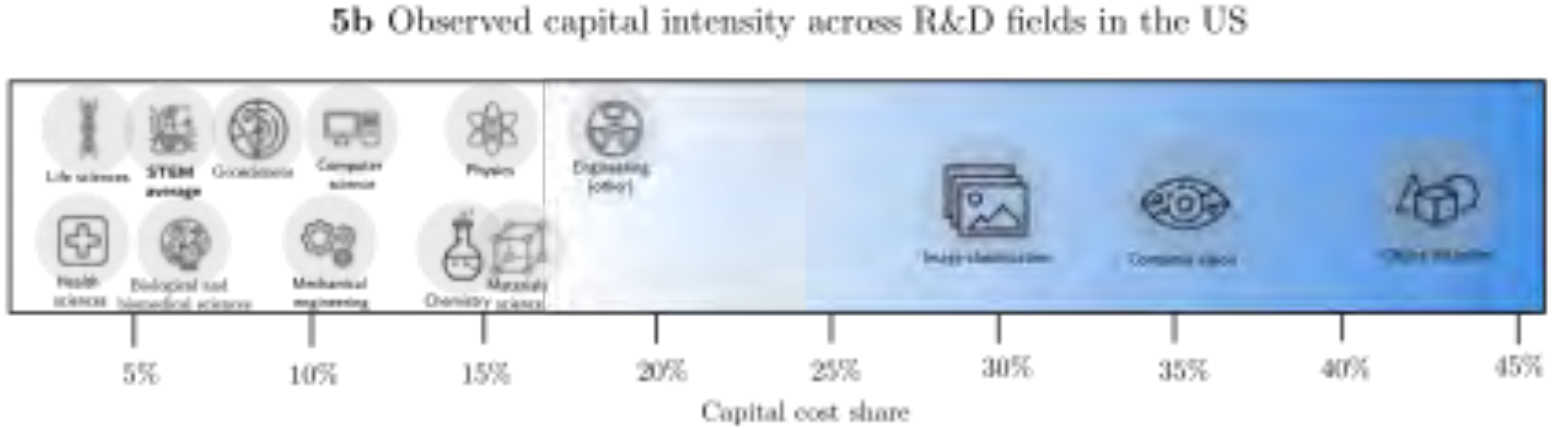
"Economic impacts of AI-augmented R&D" Tamay Besiroglu, Nicholas Emery-Xu, Neil Thompson



Science and Technology Facilities Council

Hartree Centre





"Economic impacts of AI-augmented R&D" Tamay Besiroglu, Nicholas Emery-Xu, Neil Thompson

Science and Technology Facilities Council

Hartree Centre

Why does it matter?

- The use of HPC and AI has demonstrable impact on “good outcomes” in traditional sectors.

- HPC and AI can produce permanent increase in productivity due to their positive impact on the innovation process and idea generation.

Science and Technology Facilities Council

Hartree Centre

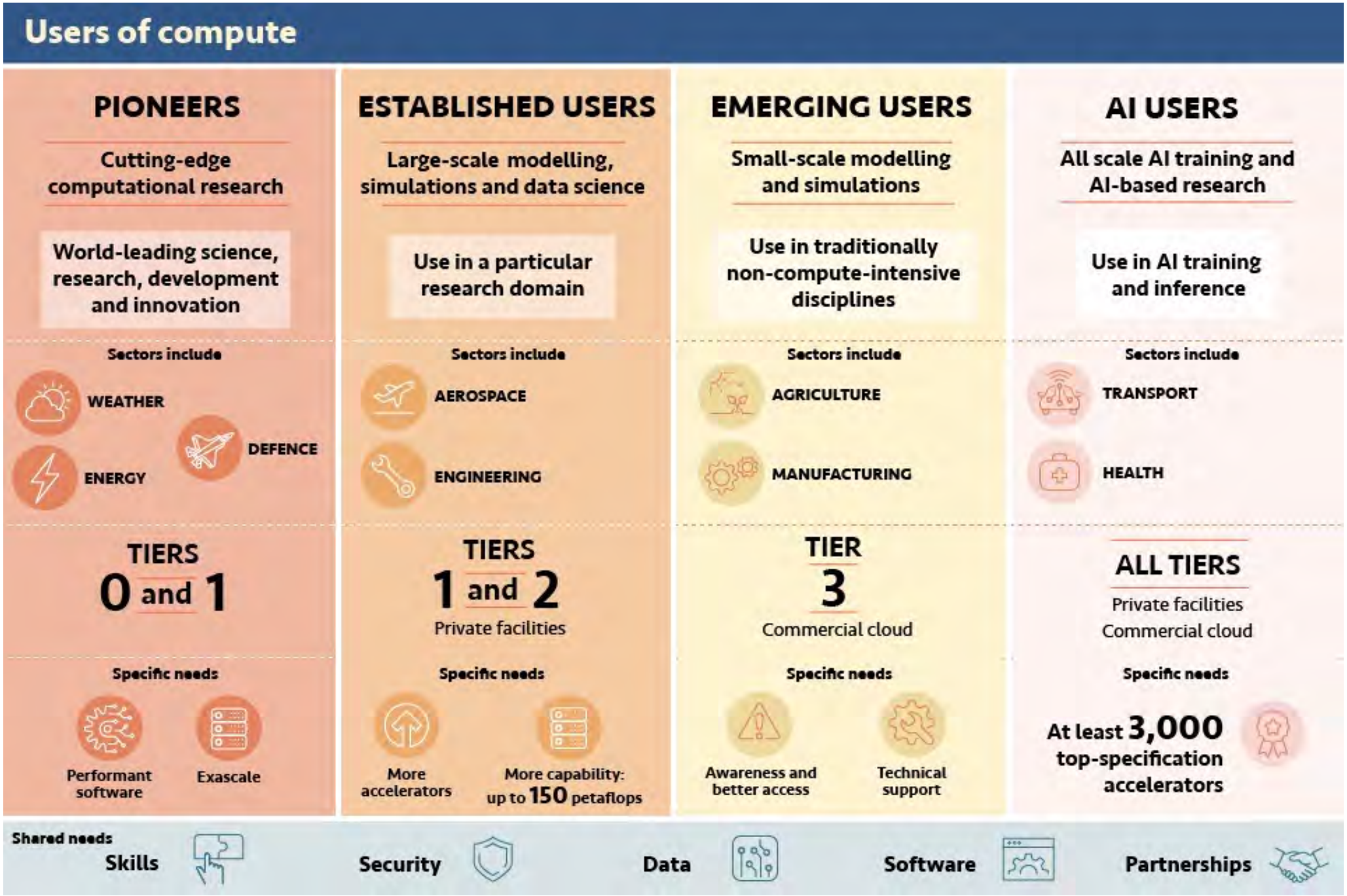
What does industry need?

What is holding us back?

Science and Technology Facilities Council

Hartree Centre

Future of Compute Review, 2023



Science and Technology Facilities Council

Hartree Centre

What are the needs?

Awareness	Accessibility	Applicability	Agility	Accelerators

- We don’t know much about compute or AI, what do we do?
- How will it help us?

What are the needs?

Awareness	Accessibility	Applicability	Agility	Accelerators

- We don’t have the compute power, what do we do?
- We don’t have money to pay for them or people to use them, what do we do?
- We don’t want to be coding, will your solutions work for us?

What are the needs?

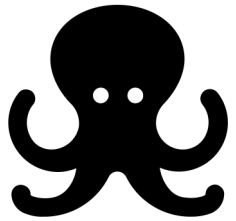
Awareness | Accessibility | **Applicability** | Agility | Accelerators

- Will these technologies work for our problem?
- Will the solutions that we get be trustworthy, reliable and robust?

What are the needs?

Awareness | Accessibility | Applicability | Agility | Accelerators

- We use different platforms for different purposes, would we be able to integrate our workflows?
- We don’t want to be locked in, can we move anytime?



What are the needs?

Awareness	Accessibility	Applicability	Agility	Accelerators

- Our software needs to be able to run on GPUs, can you help us?
- Our ML needs lots of GPUs, can you help us?

What are the needs?

Awareness	Accessibility	Applicability	Agility	Accelerators
• Technology • Potential	• Resources • Skills • Funding • Solutions	• Problem • Solutions	• Platform • Solution	• for AI • for simulations

Science and Technology Facilities Council

Hartree Centre

What are we doing?

Where do we fit in?

Science and Technology Facilities Council

Hartree Centre

What is the Hartree Centre?

- World-leading supercomputing, data analytics, AI and quantum computing technologies
- 120+ scientists, technologists and business professionals in bespoke teams working on challenge-led projects
- UK Government funded to boost productivity and innovation for industry and public sector organisations of all sizes
- Part of the Science and Technology Facilities Council in UK Research and Innovation
- Working with an international network of research communities and technology partners

Science and Technology Facilities Council

Hartree Centre

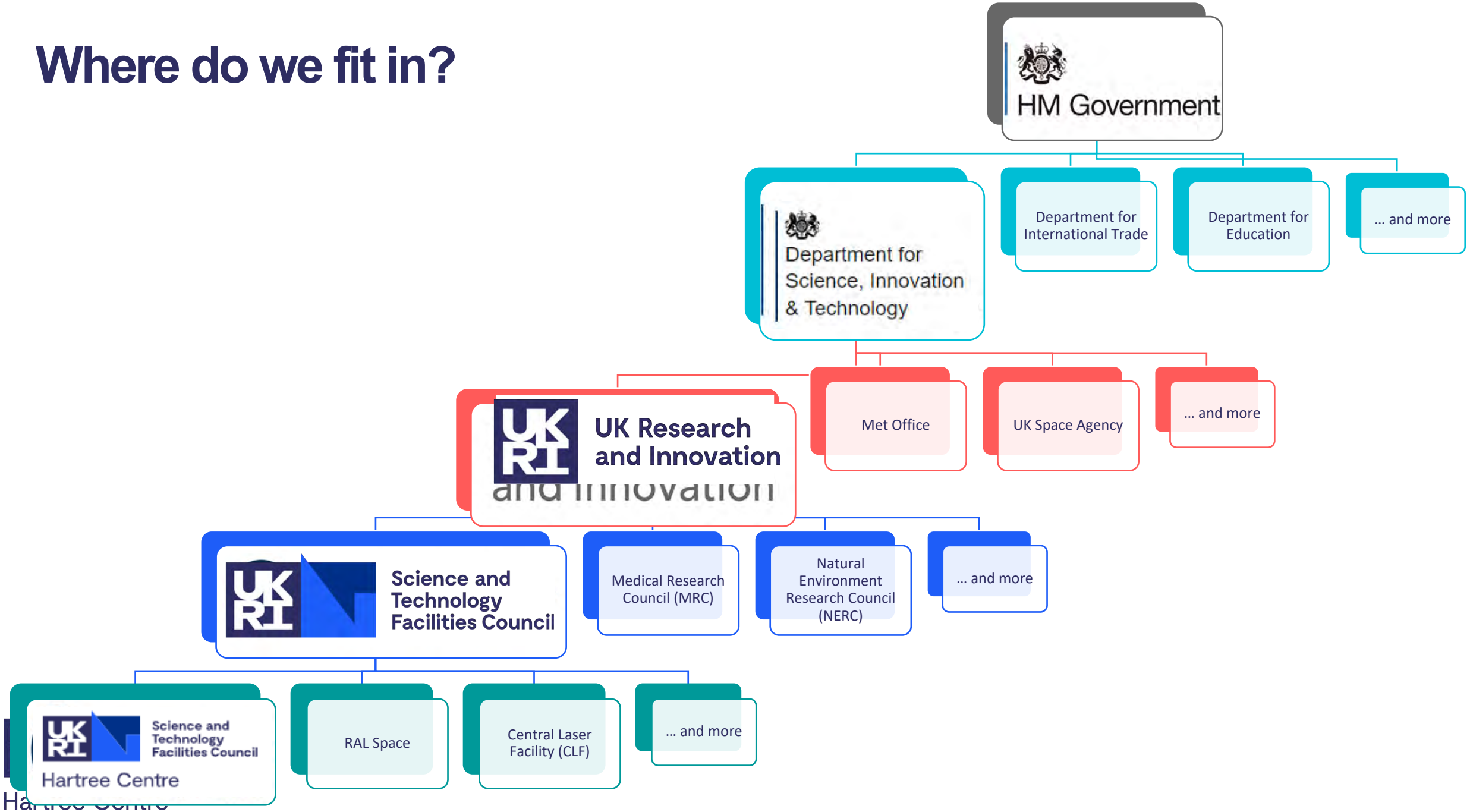
What is our mission?

Transforming UK industry by accelerating the adoption of high performance computing, big data, AI technologies & quantum computing.

Science and Technology Facilities Council

Hartree Centre

Where do we fit in?



Our platforms and facilities

Scafell Pike

Bull Sequana X1000 (4.3PFlop/s, ~80000 cores)

- Normal Compute Nodes – Skylake (Xeon Gold)
- Accelerator Nodes – Knights Landing (Xeon Phi)
- High Memory Nodes – Xeon
- GPU Nodes – nVidia V100

JADE/JADE-2 – Oxford University

Tier 2 Regional Deep Learning Supercomputer
NVIDIA DGX SuperPOD™ architecture
Atos Bull 63x DGX nodes

- 504 NVIDIA V100 Tensor Core GPUs
- 2,580,480 CUDA Cores

Cloud Facilities – On-Premise

- RedHat OpenShift (self-service)
- OpenStack VM provisioned
- AMD CPU/GPU, Nvidia A100, Alveo U200

Cloud Facilities – Public Cloud

We have access to multiple cloud vendor platforms. We are vendor agnostic, so can deploy to a variety of different cloud runtimes during and after projects. Post project, this can simplify handover of solutions into customer production environments.

Visual Computing Suite

Collaborative visual computing technologies enabling exploration of data analytics and computational modelling

Hartree National Centre for Digital Innovation (HNCDI)

- Five year collaborative partnership with IBM Research £172M UK Govt investment + £38M IBM in-kind
- Enabling businesses and public sector organisations to adopt AI and quantum computing
- A dynamic and supportive expert environment for **UK organisations of all sizes** to explore the latest technologies, develop proofs-of-concept and apply them to industry and public sector challenges for **productivity, innovation** and **economic growth**.
- Helping navigate the possibilities, de-risk investment into new technologies and discover the next step

UKRI Science and Technology Facilities Council

Hartree Centre

Tackling industry challenges

Skills
Tackling gaps within your organisation and widening the talent pool

Technical Capability
Exploring and evaluating data-driven AI technologies to help enhance productivity

Application
Developing and implementing practical solutions within your business

Resilience
Knowing how to prepare for and when to invest in the right emerging technologies (e.g. quantum computing)

EXPLAIN

EXPLORE

EXCELERATE

EMERGING TECHNOLOGY

UKRI Science and Technology Facilities Council

Hartree Centre

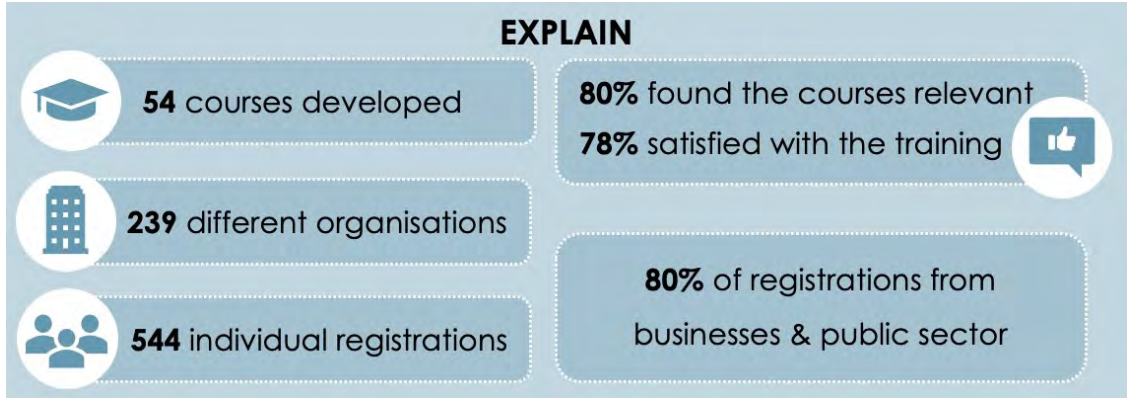
What we do

- **Collaborative R&D**
  Define a challenge in your business and we build a team to deliver a solution in the areas of:
  - Modelling & simulation
  - Code optimisation
  - Data Science and AI
  - Digital product design
- **Platform as a service**
  Give your own experts pay-as-you-go access to our compute power
- **Creating digital assets**
  License the new industry-led software applications we create with IBM Research
- **Training and skills**
  Drop in on our comprehensive programme of specialist training events or design a bespoke course for your team



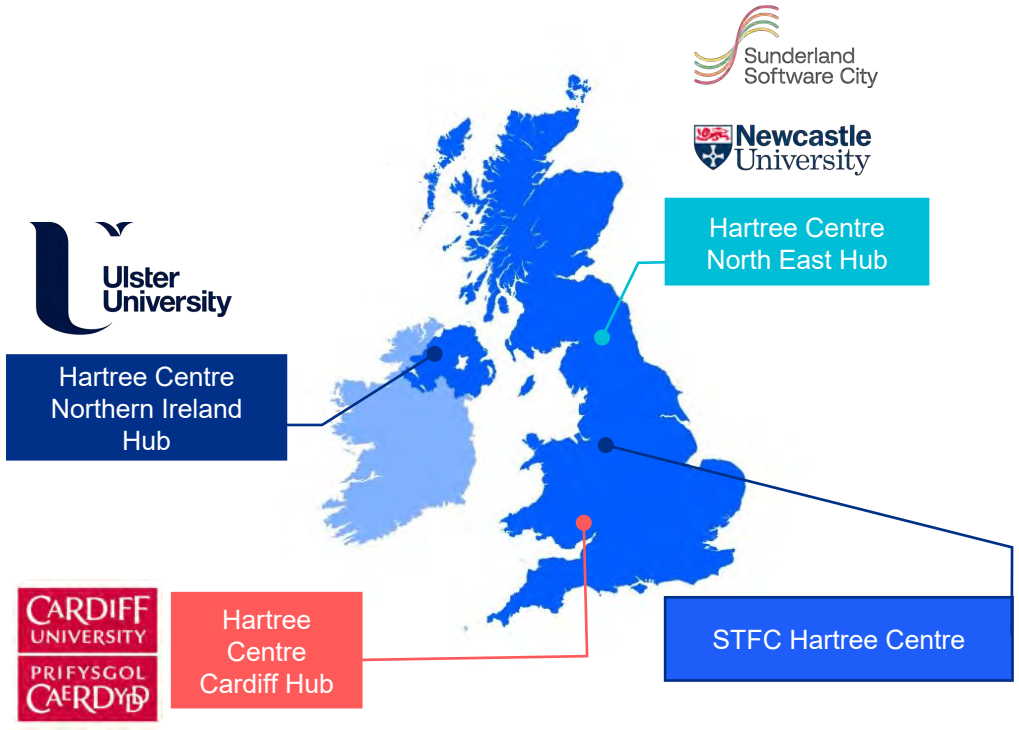

Training and engagement

HNCDI Progress Report June 2021-Dec 2022



HNCDI Explain

- Free at the point of access courses in HPC, data science, AI, full stack, cloud computing.
- Access anytime with scheduled access to technical experts.



HNCDI SME Engagement Hubs

- Engaging with local networks to increase the adoption of technologies
- Upskills SMEs locally through short projects and training.

Science and Technology Facilities Council

Hartree Centre

Case studies

How has Hartree contributed?

Science and Technology Facilities Council

Hartree Centre

Case study

Valve design for hydrogen transport

For a company with decades of experience developing valves, using simulations and virtual design enabled refinement and understanding.

“The Hartree Centre has allowed us to use specialist techniques to refine our designs to a level that otherwise would have been beyond our reach. ”

− Nick Howards, Oliver Hydcovalves

UKRI Science and Technology Facilities Council
Hartree Centre

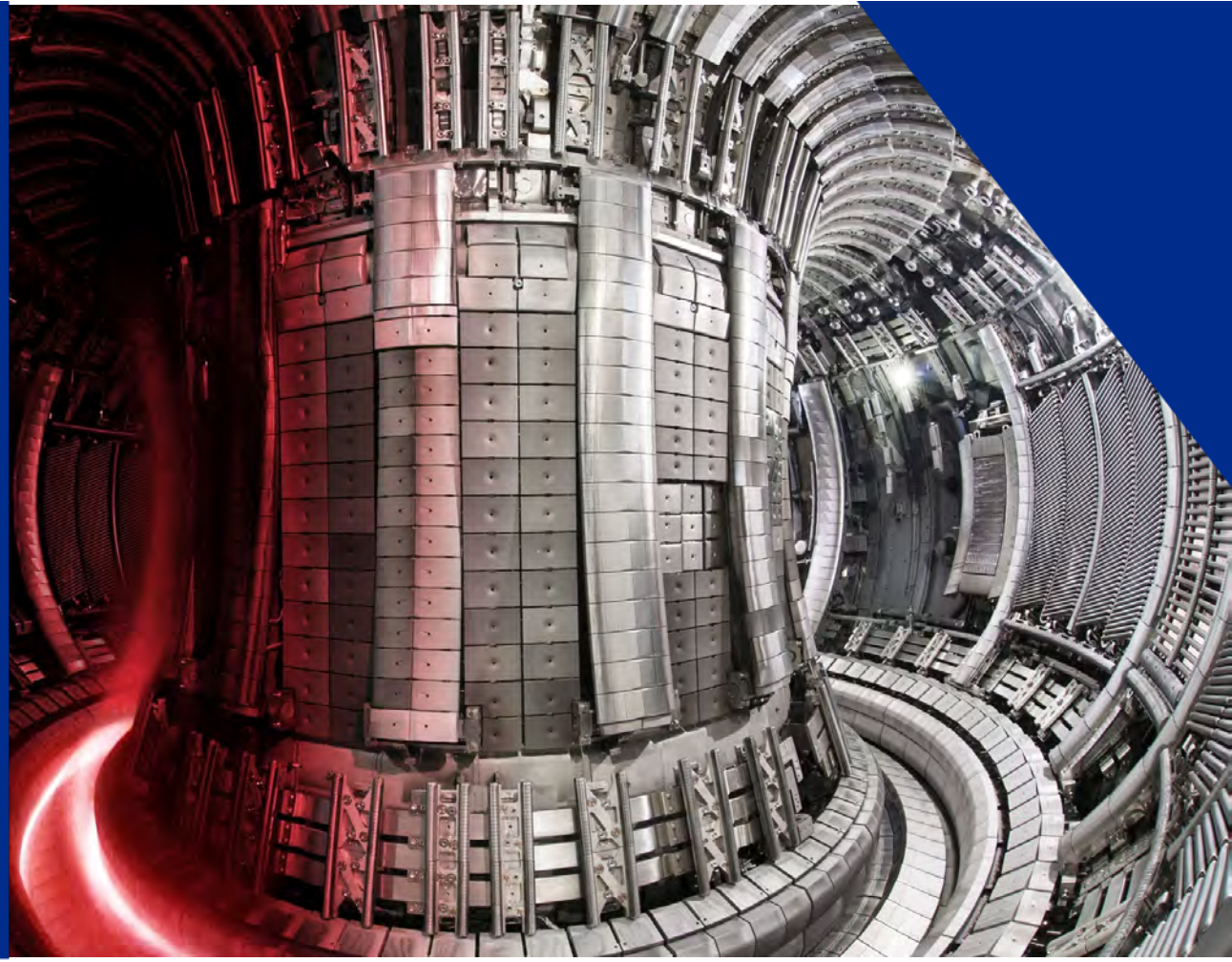
Collaborative R&D

Case study

Accelerating materials discovery

- Computational materials discovery to make high-fidelity predictions of suitable properties is demanding.
- Hartree-MaDE (Materials Discovery Engine) is a tool that simplifies and automates this process

"Working with The Hartree enabled us to efficiently explore an extremely complex area of ceramic material discovery for a niche application where currently available options are far from ideal."
- Richard White, Lucideon



UKRI Science and Technology Facilities Council
Hartree Centre

LUCIDEON

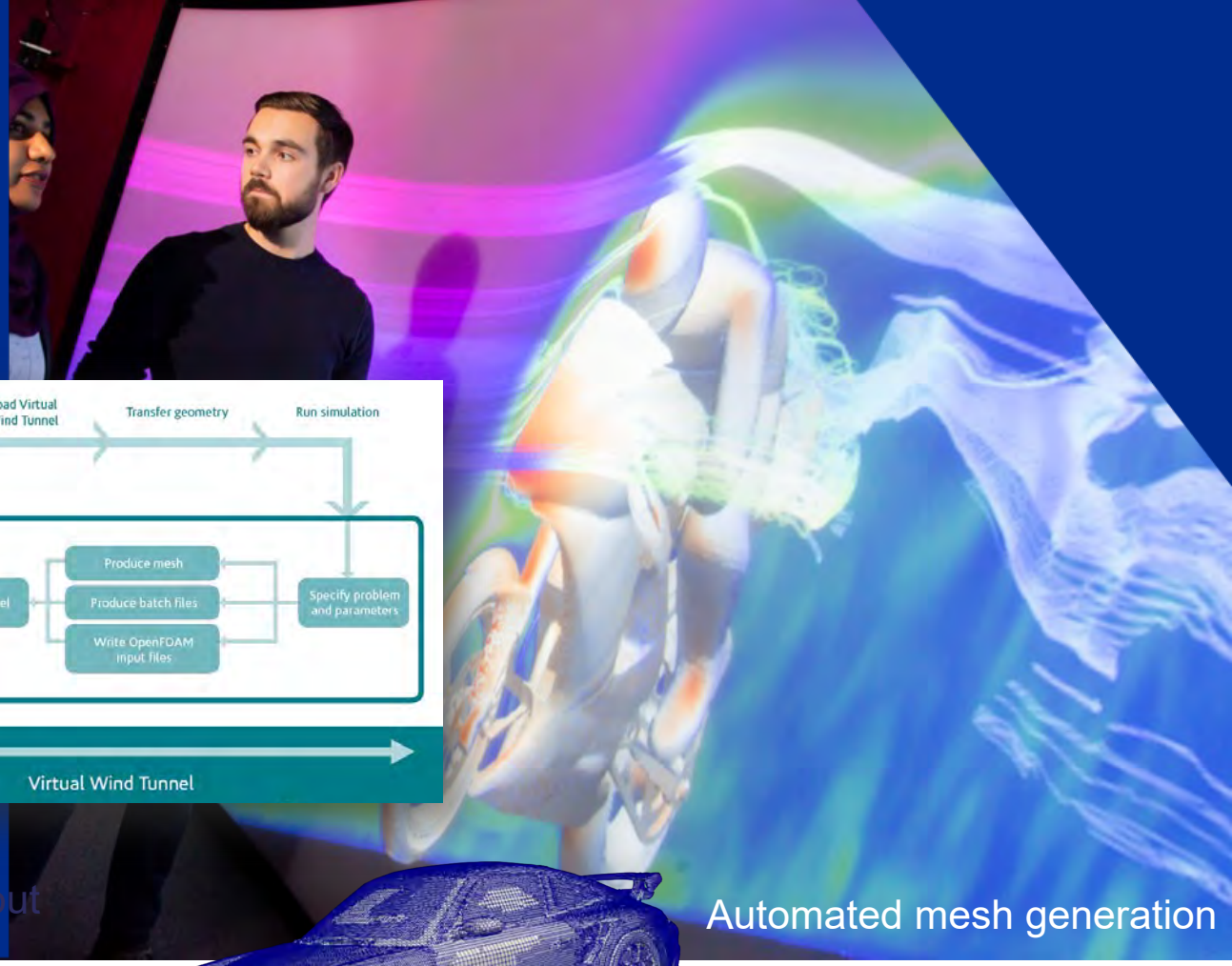
Collaborative R&D

Case study

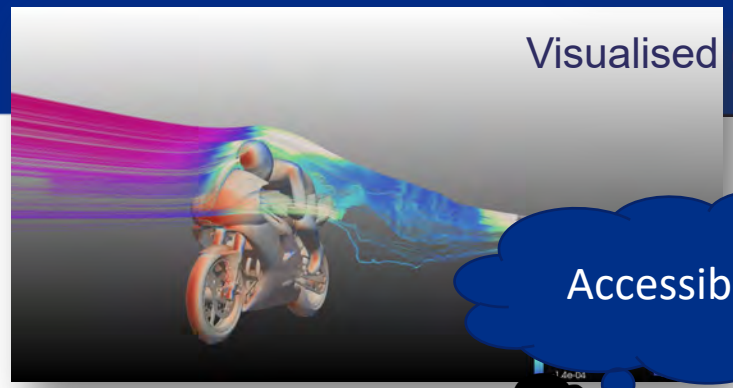
Virtual Wind Tunnel

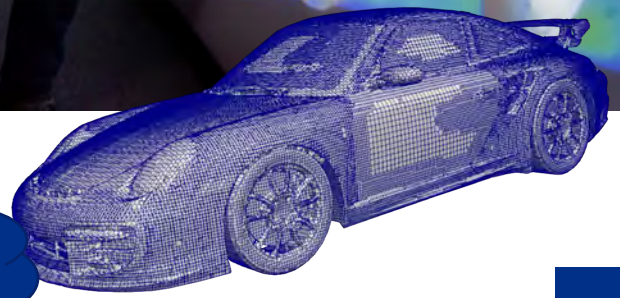
Saving time and money for automotive and aerospace design

- Builds the wind tunnel environment
- Automate the domain decomposition
- Produce an automatic mesh from a 3D model file (.obj / .stl)
- Automatically configure the CFD engine
- Submit the job onto Scafell Pike (Hartree Centre flagship HPC platform)



Automated mesh generation





Accessibility

Collaborative R&D

Science and Technology Facilities Council

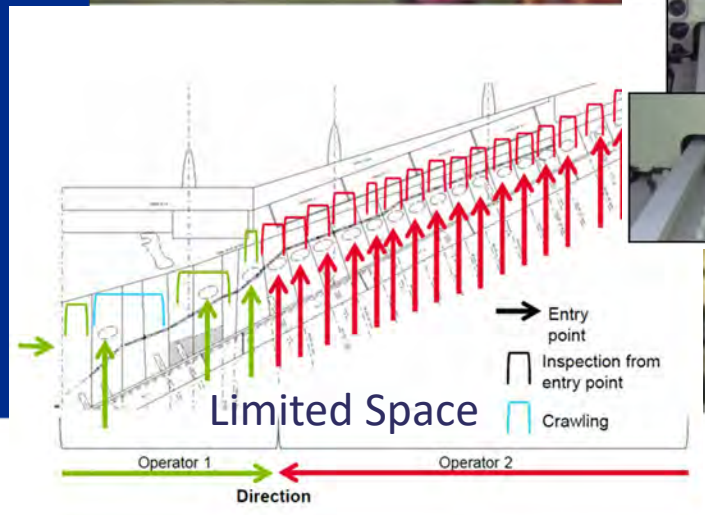
Hartree Centre

Case study

Airbus | Deep Learning for Wing Tank Inspection

Faster quality control process for wing tank inspection (A320 & A321)

- Check correct standard of: sealant, fasteners, paint, adhesion, liquids
- Detection of flaws: Missing/damaged sealant, excess paint, scratches and foreign objects (nuts, bolts, misc tools, swarf, safety glasses, wire, etc.)



Model Creation with >1000 Images

Train/Validate 80:20

Test Set >50 Images

UKRI Science and Technology Facilities Council

Hartree Centre

AIRBUS

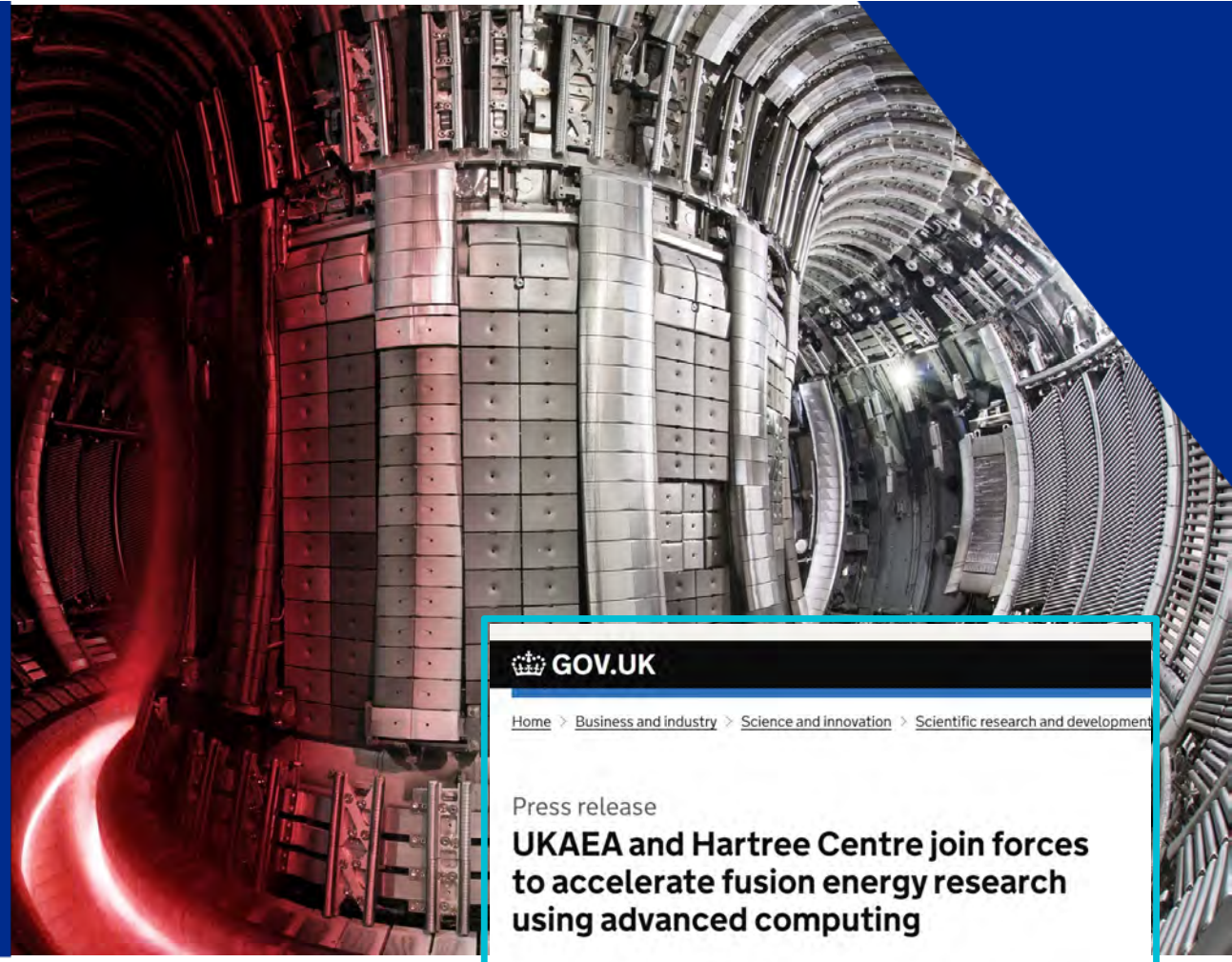
Collaborative R&D

Case study

Virtual design of fusion reactors

Develop framework for harnessing the power of HPC for design of future fusion reactors

- Reduced order modelling
- Identifying libraries and algorithms for multi-physics coupling and exascale computing
- Dealing with large data and efficient data management processes
- New hardware and visualisation capabilities



GOV.UK

Home > Business and industry > Science and innovation > Scientific research and development

Press release

UKAEA and Hartree Centre join forces to accelerate fusion energy research using advanced computing

New Centre of Excellence in Extreme Scale Computing in Fusion to be located at STFC's Hartree Centre

UKRI Science and Technology Facilities Council
Hartree Centre

UK Atomic Energy Authority

Collaborative R&D

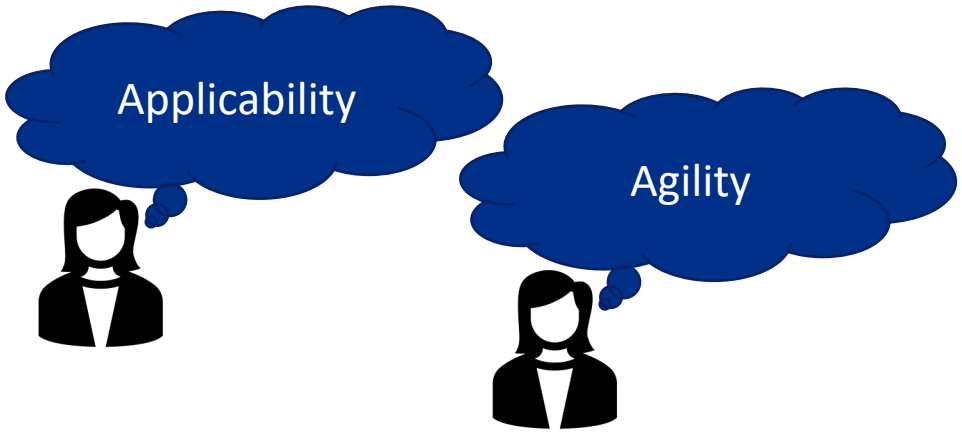
Case study

Enabling separation of concerns in next generation weather models.

The Hartree Centre created PSyclone, a tool which auto-generates the code needed for the Met Office's next-generation weather model to run on different HPC architectures.

- Frees scientific developer from worrying about parallelism and optimisation
- Allows the HPC expert to optimise an entire scientific code for a particular architecture using Python scripting

PSyclone



UKRI Science and Technology Facilities Council
Hartree Centre

Met Office

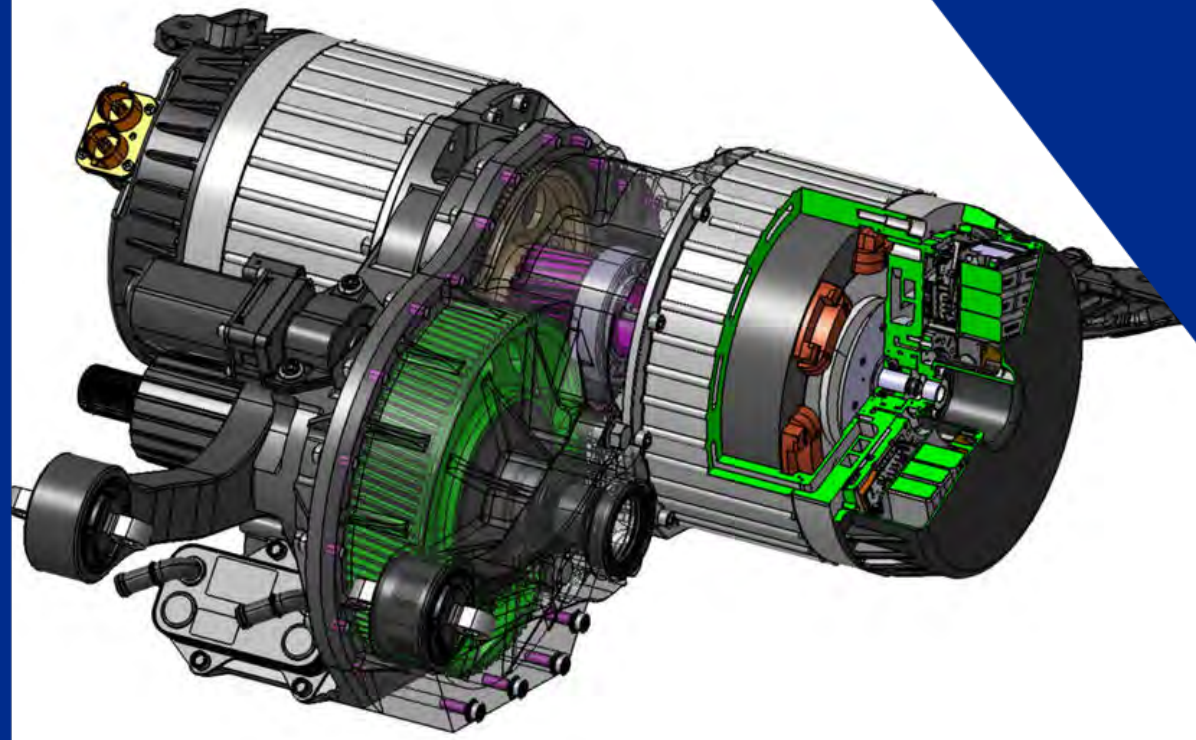
NEMO

Case study

OCTOPUS – Towards a digital twin of an electric vehicle powertrain

Develop an efficient digital solution for virtual design of the gearbox.

- Identify what we want the model to be used for
- Identify what physics needs to be captured to get the answers we want
- Develop a multi-physics GPU-accelerated simulation toolkit based on smoothed particle hydrodynamics



UKRI Science and Technology Facilities Council
Hartree Centre

BENTLEY

Advanced Electric Machines

UKRI Innovate UK

Collaborative R&D

Case study

Computer aided formulation

Faster development process for products like shampoo, reducing testing

"The Hartree Centre's high performance computing capabilities help us achieve better design solutions for our consumers, delivered by more efficient, cost-effective and sustainable processes."

− Paul Howells, Unilever

UKRI Science and Technology Facilities Council

Hartree Centre

Unilever

Collaborative R&D

Science and Technology Facilities Council

Hartree Centre

Thank you

ubaid.qadri@stfc.ac.uk

hartree.stfc.ac.uk @HartreeCentre STFC Hartree Centre

CIUK 2023 Presentations

Ben Rixon & Neil Martin (The Manufacturing Technology Centre)

Automated Deployment of Manufacturing Use-Cases through OpenStack HPC

Abstract: At the MTC we have worked with StackHPC to deploy a private cloud computing platform utilising OpenStack Kayobe, hosted in the Daresbury Science Park. This environment combines DevOps methodologies with the deployment of manufacturing software, providing a rapidly configurable SaaS environment for the testing and valuation of manufacturing use-cases. The OpenStack Kayobe platform also enables the digital thread, by hosting multi-tenant project environments containing a variety of simulation and manufacturing software.

Bio: Ben Rixon is a senior research engineer at The Manufacturing Technology Centre, specialising in infrastructure and cyber security. Graduating from Liverpool John Moores university in 2017 with a degree in cyber security, his career has been focused on designing and delivering infrastructure solutions for projects in a variety of sectors. He has previously led the digital infrastructure design and implementation for an aerospace project, delivering a connected manufacturing cell to enable the digital thread from design to build.

Neil Martin is a research engineer at the Manufacturing Technology Centre with a focus on cloud computing and infrastructure. Neil's previous roles in both infrastructure and software engineering has allowed him to gain insight into methodologies from both disciplines. He is currently working across a variety of technologies areas, from infrastructure and automation to machine learning and analytics, to delivery customer driven projects in the manufacturing sector.

mtc
Manufacturing
Technology Centre

CATAPULT
High Value Manufacturing

Automated Deployment of Manufacturing Use-Cases through OpenStack HPC

Ben Rixon – Senior Research Engineer

Neil Martin – Research Engineer

Background

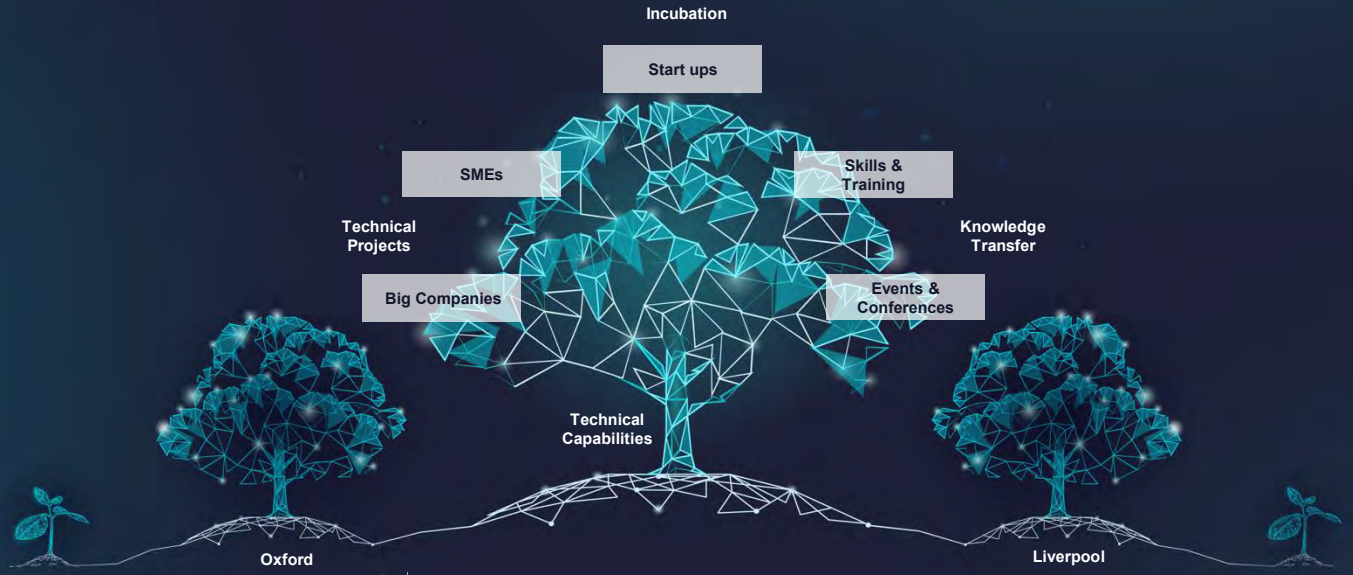
- Opened in 2011
- Independent RTO
- To bridge the valley of death
- Prove innovative manufacturing ideas
- Manufacturing system solutions
- Training & Skills

UNIVERSITY OF BIRMINGHAM

Loughborough University

The University of Nottingham

What do we do?
Building Ecosystems - National Capability - Local Impact



Utilising key technologies

TRANSFORMATION TEAM	DIGITAL ENGINEERING	ADVANCED PRODUCTION SYSTEMS	COMPONENT MANUFACTURING TECHNOLOGY	FUTURE SKILLS
BUSINESS TRANSFORMATION	MODELLING & SIMULATION	AUTOMATION, MECHATRONICS AND ELECTRONICS MANUFACTURING	ADDITIVE MANUFACTURING	
DIGITAL TRANSFORMATION	METROLOGY & NDT	DESIGN & BUILD	LASER PROCESSING	
TECHNOLOGY TRANSFORMATION	INFORMATICS		ADVANCED MATERIALS PROCESSING	

What lead to the project?

What we had	What we needed
Isolated Project Infrastructure	Secure Multi-Tenant Project Environment
Long Software Deployment Times	Flexible Digital Sandpit
Lack of Engineer Autonomy	Re-deployable Manufacturing Software
Microsoft Hyper-V 2016	Reconfigurable Digitally Enabled Shop-Floor
	Modern cloud-based technologies

The Digital Manufacturing Accelerator Programme Overview

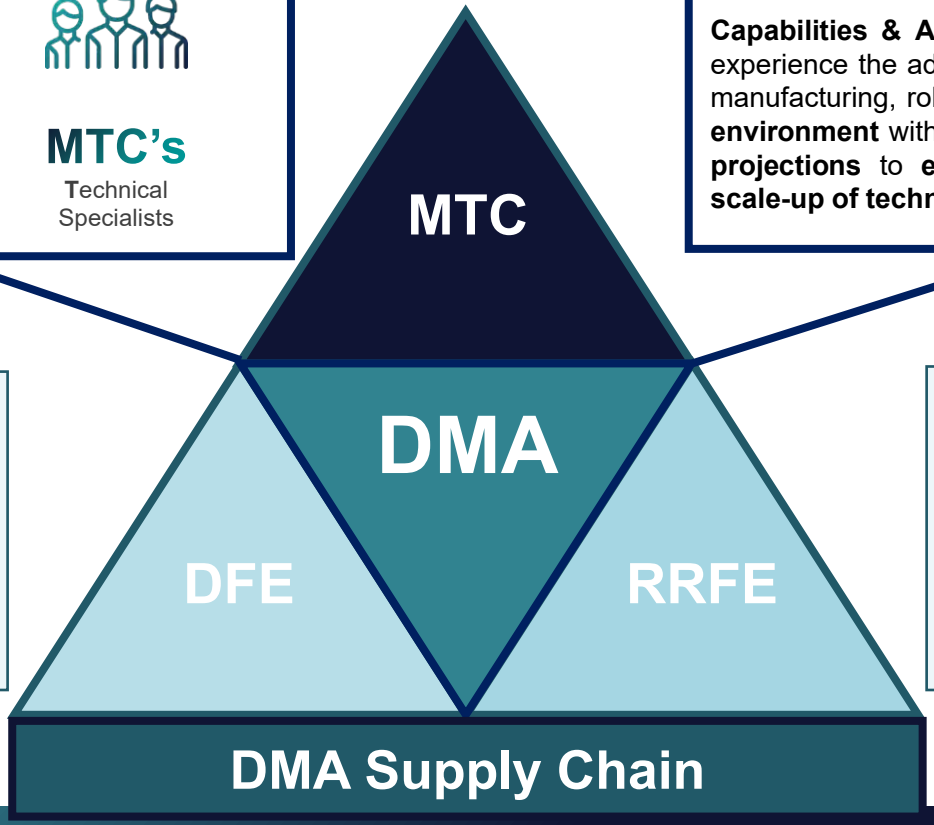
LIVERPOOL CITY REGION COMBINED AUTHORITY · METRO MAYOR LIVERPOOL CITY REGION · NORTHERN POWERHOUSE · Liverpool City Region Local Enterprise Partnership

STRATEGIC INVESTMENT FUND

£17m
Investment in the assets

LCRA
Funding

MTC's
Technical Specialists



Capabilities & Assets which allow manufacturers to experience the advantages of **Industry 4.0**, IoT, smart manufacturing, robotics, and automation in a **risk-free environment** with accurate **financial and operational projections** to **enable the deployment and rapid scale-up of technological capabilities.**

Digital Factory Environment

A **digital sandpit** which provides a **digital infrastructure** & **catalogue of pre-configured solutions** to develop and demonstrate new digital manufacturing solutions and technologies.

Rapidly-Reconfigurable Factory Environment

- A **physical test-bed facility** which provides the **space, utilities and reconfigurable production lines** for developing and demonstrating new manufacturing solutions and technologies.

The Digital Manufacturing Accelerator Infrastructure

Data and Computing Power on Demand

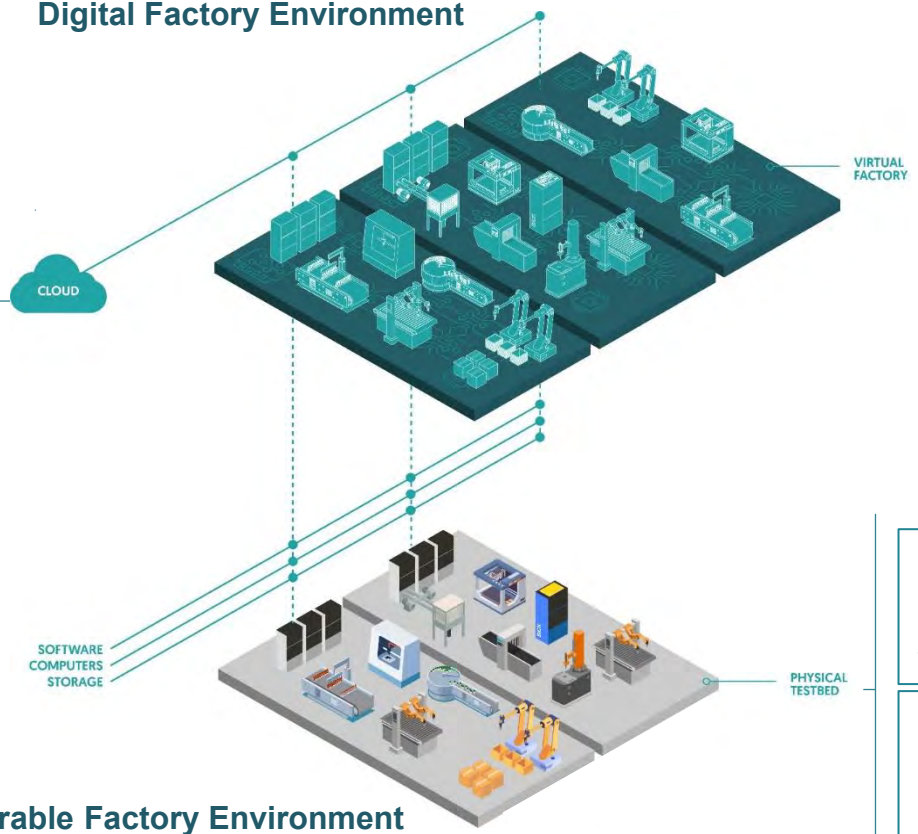
Control and data acquisition modules enabling testing of different digital integration configurations.

A digital infrastructure which can be quickly configured and deployed for client projects in a safe unconstrained environment.
(You can mess about without messing up!)

Client portal providing a single interface to project data and manufacturing systems.

Library of manufacturing ICT systems and simulations from a range of vendors which can be quickly configured and deployed.

Digital Factory Environment



Rapidly-Reconfigurable Factory Environment

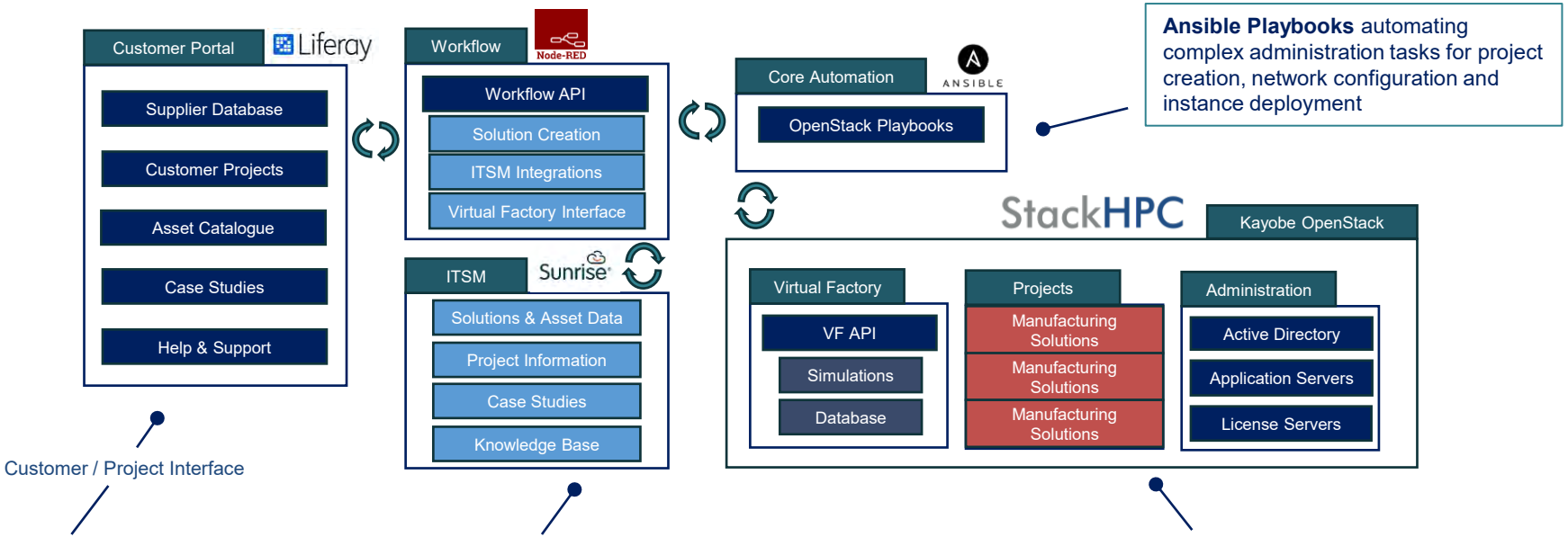
Virtual factory technologies allowing simulation of processes, assets and architectures before deploying physical manufacturing systems, leading to a digital twin.

"Smart factory" control, Digital twinning, monitoring and optimisation of the physical and digital environments ...

Factory space with flexible services and utilities allowing multiple manufacturing systems to be deployed at any one time.

Standardised process interfaces allowing modular reconfiguration of manufacturing systems.

How does the deployment look?



Portal providing information on previous projects, associated **case studies** and **suppliers**

Ability to view and provision **manufacturing software assets** within the Digital Factory Environment

Node-RED API interfacing with OpenStack through Ansible, **Virtual Factory** applications and the ITSM Platform

Information contained in ITSM is queried from the customer portal via the **Workflow API**

Pre-configured manufacturing applications packaged as solutions to be re-deployed when required, reducing installation and configuration time for complex platforms.

Virtual Factory API hosting simulations and modelling applications to be redeployed across projects

What is the Virtual Factory?

- Workflow driven **modelling & simulation platform** combining multiple software platforms to generate analysis results.
- Can be redeployed across multiple projects, **to minimise reconfiguration time** for complicated software required to perform analysis.
- Separates simulation capabilities from individual user devices by having an easy to access **centralised platform**.
- Queries' data from project **Manufacturing Execution Systems (MES) and Historians** hosted on virtual machines within OpenStack.

What have we learnt from this?

- Modern private cloud platforms and automation tools **significantly cut down on deployment time** for all our requirements.
- Everything cloud platforms can provide, can be **engineered for manufacturing use-cases**!
- Manufacturing doesn't need to be pushing the boundary of compute, they need to **understand what we already know** and how best to **utilise it**.
- By creating **pre-configured images**, it provides a significant **reduction in setup and configuration time** for manufacturing, simulation and modelling workloads.
    - Simulation and modelling deployment time reduced **from two weeks, to four hours**!
- Security is still hard! The infrastructure operates in a whitelisted environment, retrieving software packages from multiple repositories with dependences proves challenging.
- A customer portal provides an **easy interface to select manufacturing solutions**, pre-configured with required software, without having to understand the working components of a cloud platform.

What are the Digital Factory Environment teams' next steps?

- The DMA team are continuing to bring new projects onto the platform, creating **new pre-configured manufacturing solutions** and providing a **reconfigurable environment to evaluate software and services**.
- **More Ansible, more automation**. We've only just started our journey in automation, can we automate more than just the core infrastructure?
- Now that we've created this environment to address a project specific purpose, how do we transition this into a fully **operational asset**, maintained in collaboration with IT and our Engineering teams.
- How can we streamline our platform, are there alternatives to consider for the key components which **simplify the deployment process**?

THANK YOU

Ben Rixon – ben.rixon@the-mtc.org

Neil Martin – neil.martin@the-mtc.org

Learn more about the DMA: https://digitalmanufacturingaccelerator.com/

Learn more about the MTC: https://www.the-mtc.org/

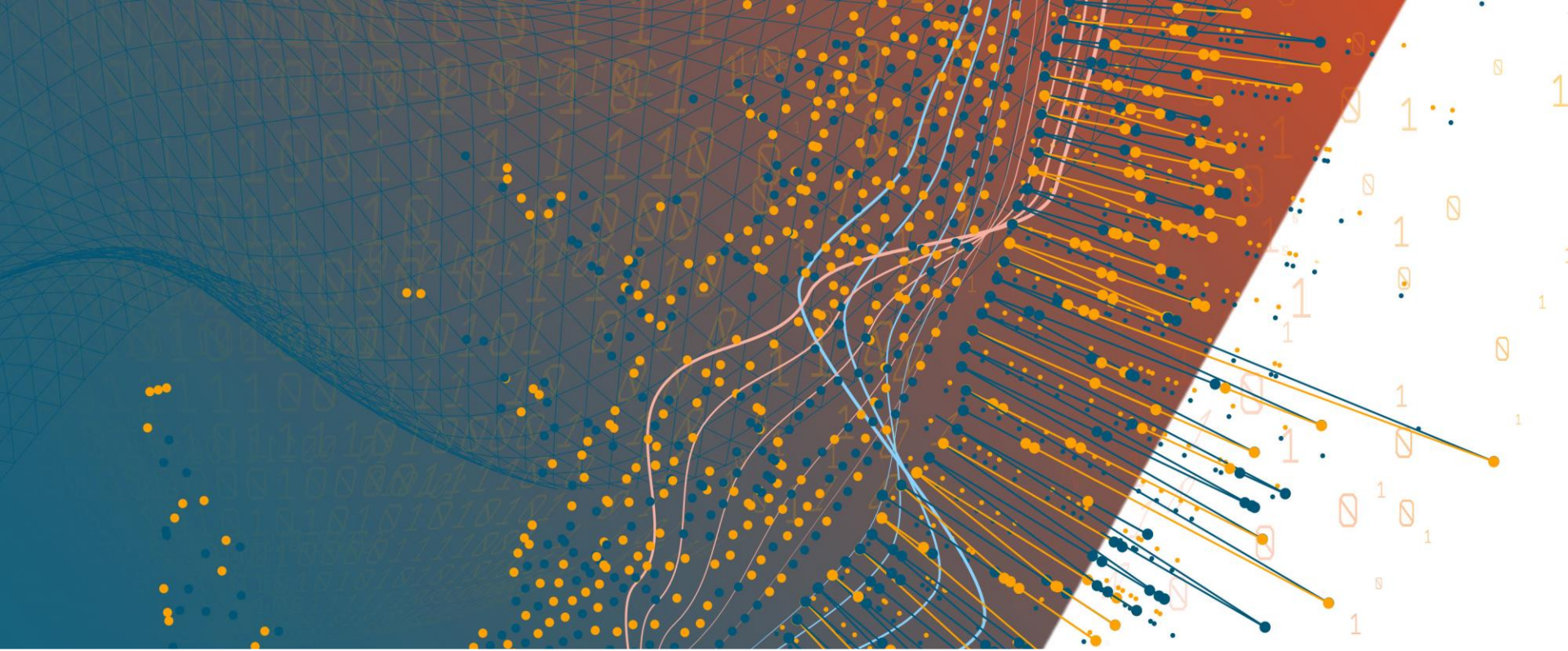
CIUK 2023 Presentations

Dr Rosemary Francis (Altair)

Presenting the biggest change to HPC in 20 years

Abstract: HPC is rapidly becoming more complex. We must support a wide range of new workloads that mix AI/ML with HPC and pull data from IoT, edge compute, cloud and new streaming data sources. There is further diversification of the compute environment as we move away from the default x86 machines to a mix of CPU architectures, GPU devices, and new accelerators like the inference machines or even quantum. How do we cope with this complexity as an industry while delivering ever-better scalability and efficiency? How do we continue to align resource use with value added in a transparent way for the administrator and the user? In this presentation Dr Rosemary Francis, Chief Scientist at Altair will present how Altair plans to address these challenges now, and in the future, introducing the biggest change to HPC in 20 years.

Bio: Dr Rosemary Francis FReng founded Ellexus, the I/O profiling company, in 2010. She grew the company and global customer base for ten years before it was acquired by Altair in 2020. Rosemary obtained her PhD in computer architecture from the University of Cambridge and worked in the semiconductor industry before founding Ellexus. She is now Chief Scientist for HPC at Altair, responsible for the future roadmap of workload managers Altair® PBS Professional® and Altair® Grid Engine®. She also continues to manage I/O profiling tools, Altair® Breeze™ and Altair® Mistral™ and other analytics and reporting solutions across Altair's HPC portfolio. Rosemary is a fellow of the Royal Academy of Engineering and a Chartered Engineer with the Institution of Engineering and Technology.

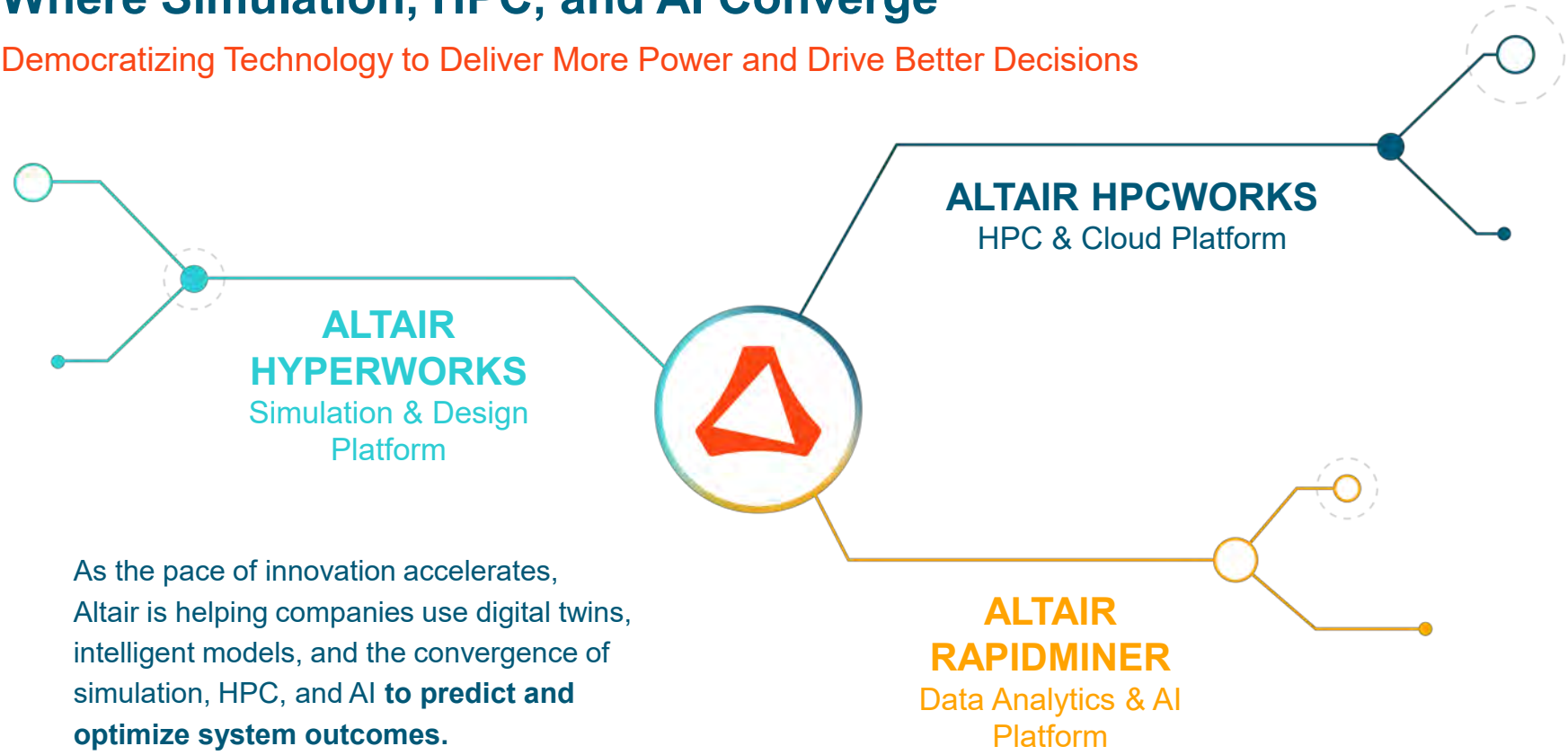


ALTAIR® LIQUID SCHEDULING®: THE BIGGEST CHANGE TO HPC IN 30 YEARS

Dr. Rosemary Francis, Chief Scientist HPC

Where Simulation, HPC, and AI Converge

Democratizing Technology to Deliver More Power and Drive Better Decisions



As the pace of innovation accelerates, Altair is helping companies use digital twins, intelligent models, and the convergence of simulation, HPC, and AI **to predict and optimize system outcomes.**

ALTAIR

AI workloads are taking over

Data analytics

ML on place of solvers

Large language models

Self driving cars

Edge, IoT, streaming data

Hybrid HPC - AI workflows

ALTAIR

How are AI jobs different from HPC

Training vs inference

Run at the edge

Often High Throughput, not High Performance

Often Bursty

Often have tight deadlines for reasonable ROI

Platforms are becoming more diverse

Cloud

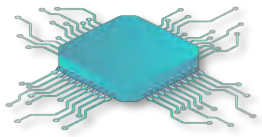
Hybrid Cloud

Kubernetes

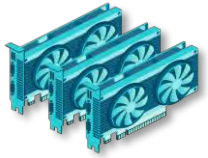
Edge

ALTAIR

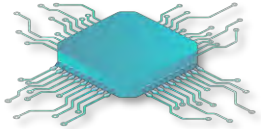
Hardware is getting more complicated



CPU



GPU



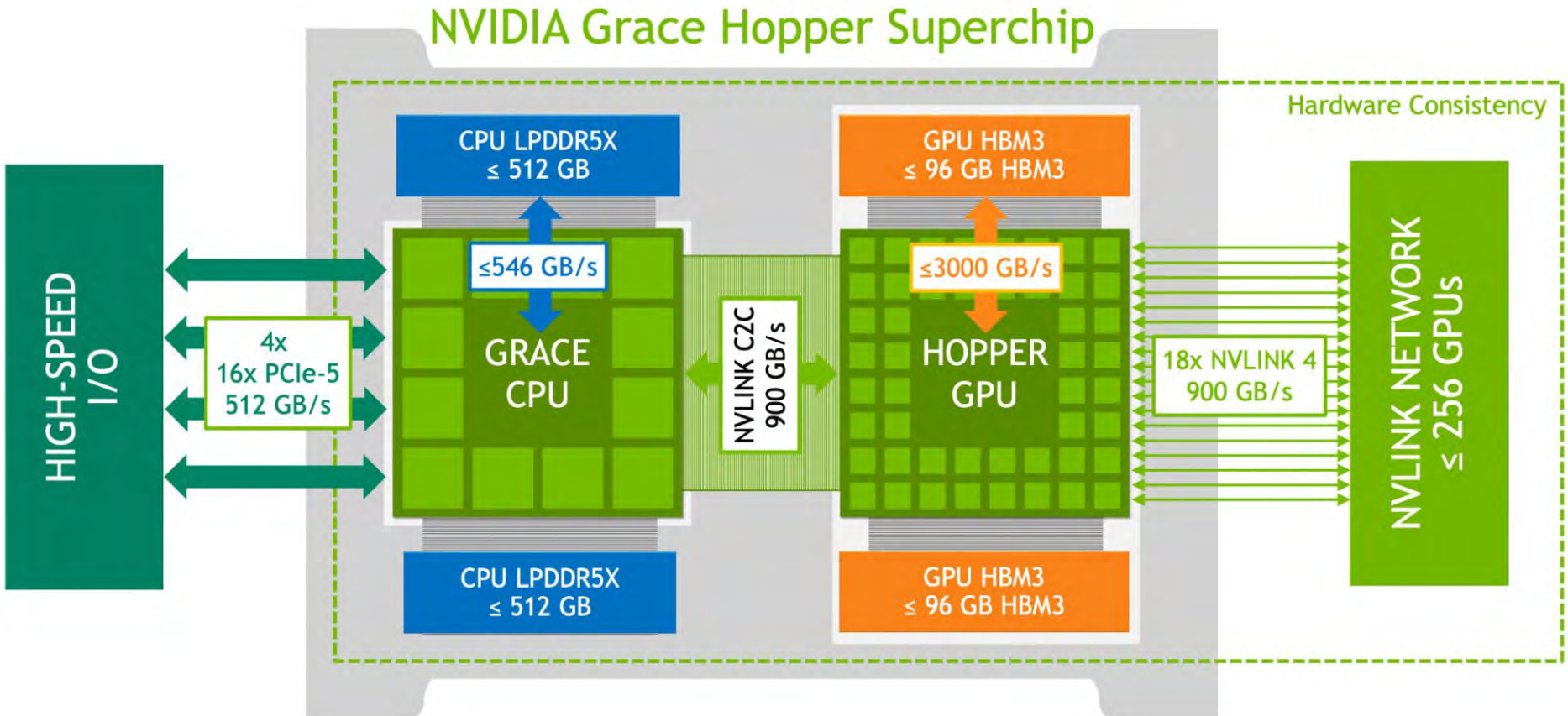
DPU

Networking

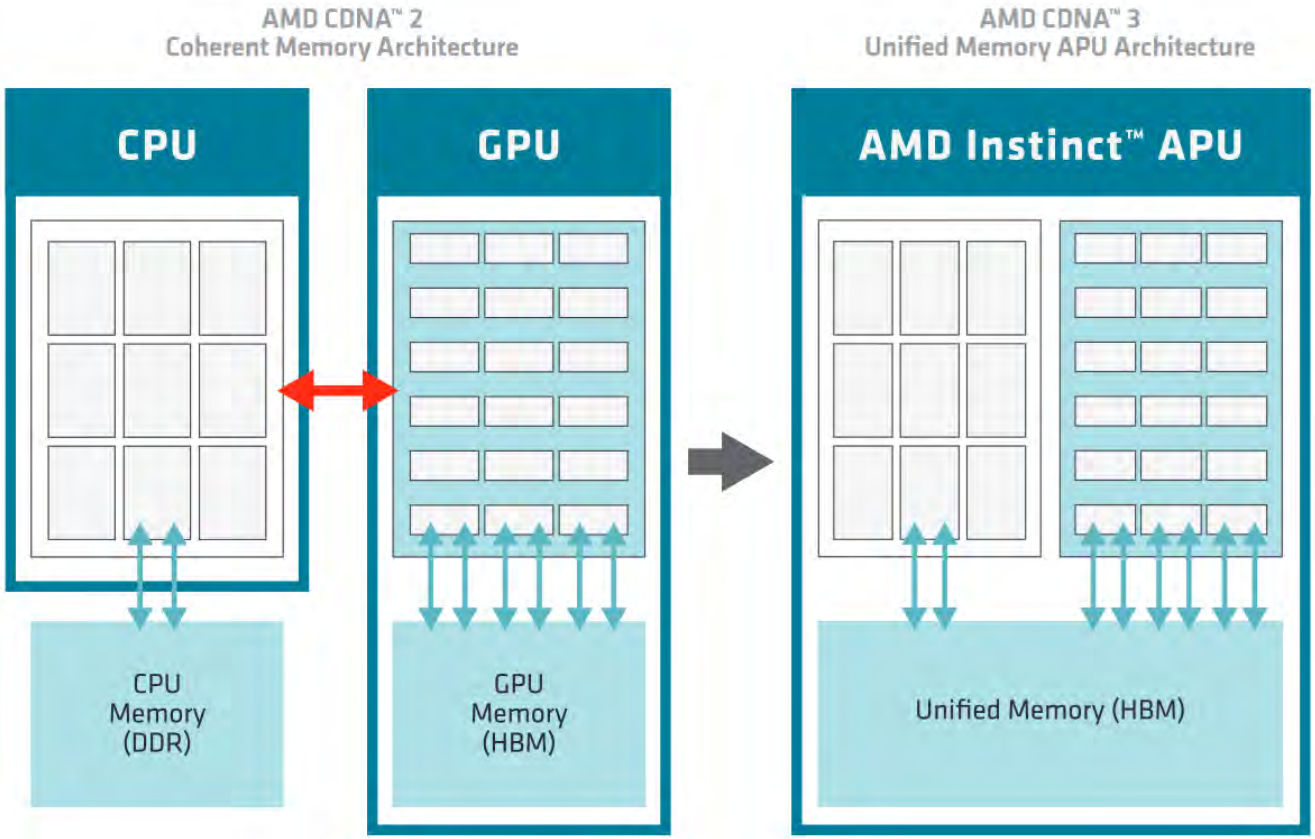
Storage

ALTAIR

GPU SoC mega chips



GPU SoC mega chips



ALTAIR

Inference and AI chips

Groq

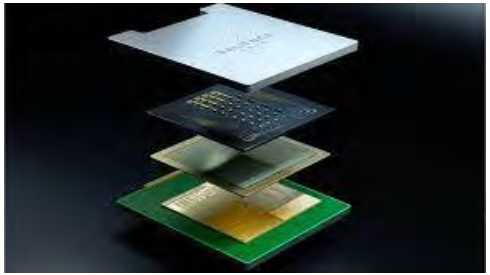
Graphcore

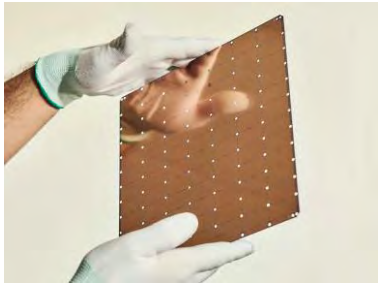
Salience

Cerebras

SambaNova

Habana





ALTAIR

Scaling challenges

End to Moore’s law

Cores increase but node count does not

Exascale

Distributed HPC

Cloud is not infinite

Challenges as HPC increases in complexity

HPC today covers a diverse range of applications, technologies, and hardware

- High overheads in integration and support

Users have to choose where to run their codes

- Leading to low utilization

Administrators have low visibility and control over resources

- Causing poor return on investment (ROI)

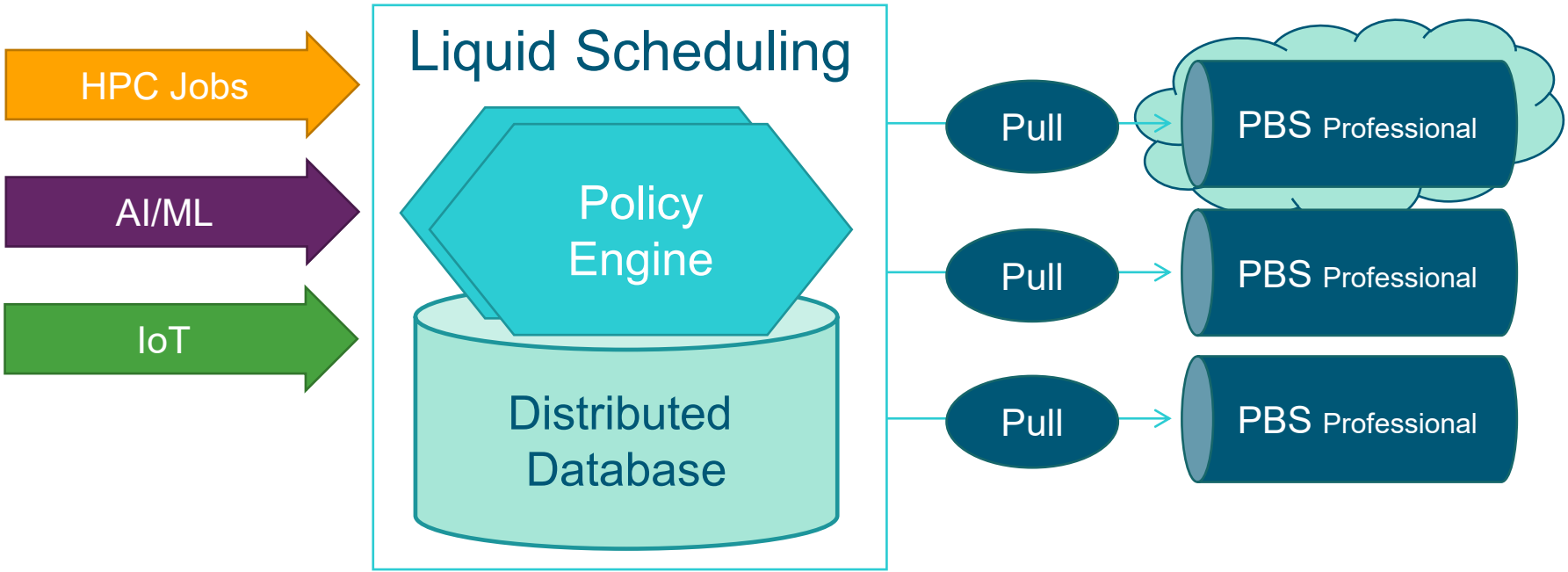
Liquid Scheduling

Liquid Scheduling creates an HPC platform that delivers scalability for the next generation of supercomputers.

- **Leap in performance** and scalability
- **Support diverse new workflows** and exotic compute
- **Break down silos** by connecting multiple HPC clusters
    - Improve utilization and access to resources
- **Built using a modern web-scale technology stack**

Liquid Scheduling

Bringing together new workloads and compute platforms



ALTAIR

Key architectural concepts

Multiple workload managers queue jobs only when they are about to run them

Jobs execute on the first available resources

Liquid Scheduling applies global policies

Global control over fairshare and prioritization for better utilization

Submission through PBS Professional CLIs

No change for end users or applications

Built for future scalability needs

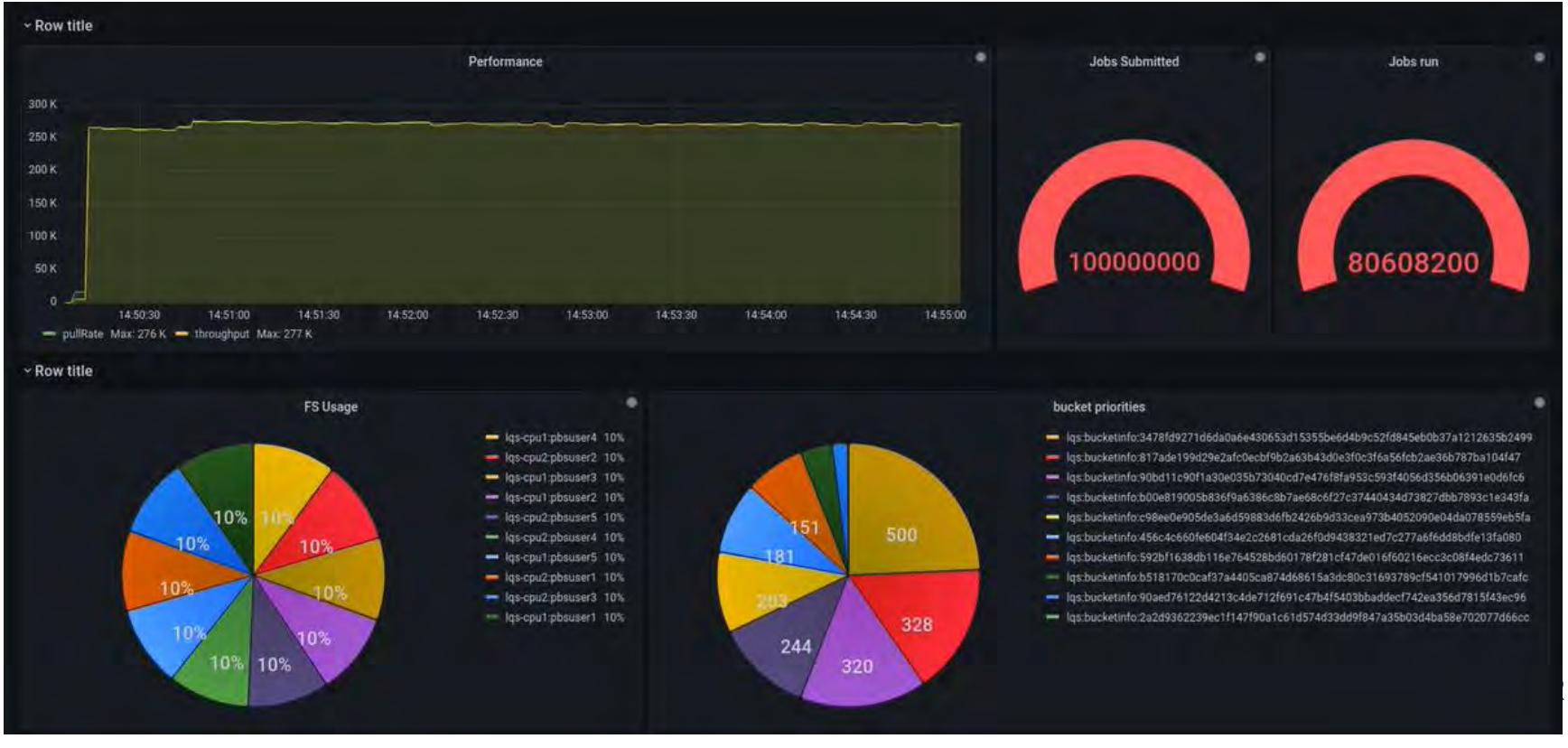
Using modern, web-scale technologies and streaming architectures

What does this mean to PBS Professional customers?

- Schedule workloads automatically across multiple PBS Professional clusters
- Fairshare implemented across multiple PBS Professional clusters
- Resilience against qstat overload
- Better performance and scalability even when using just one PBS Professional cluster

Demonstrating scalability

100M jobs across eight PBS Professional clusters with fairshare delivering >250K jobs per second



Why is Liquid Scheduling so fast?

- Separation of scheduling policies from resource management
  - No bottleneck component that needs to know everything
  - Polices are applied to groups of similar jobs
  - Workload managers only handle small queues of jobs that are ready to run
  - Even with just one PBS Professional cluster, performance is improved

What makes Liquid Scheduling so disruptive?

Altair is the world leader in ensuring workloads run in the right place at the right time.

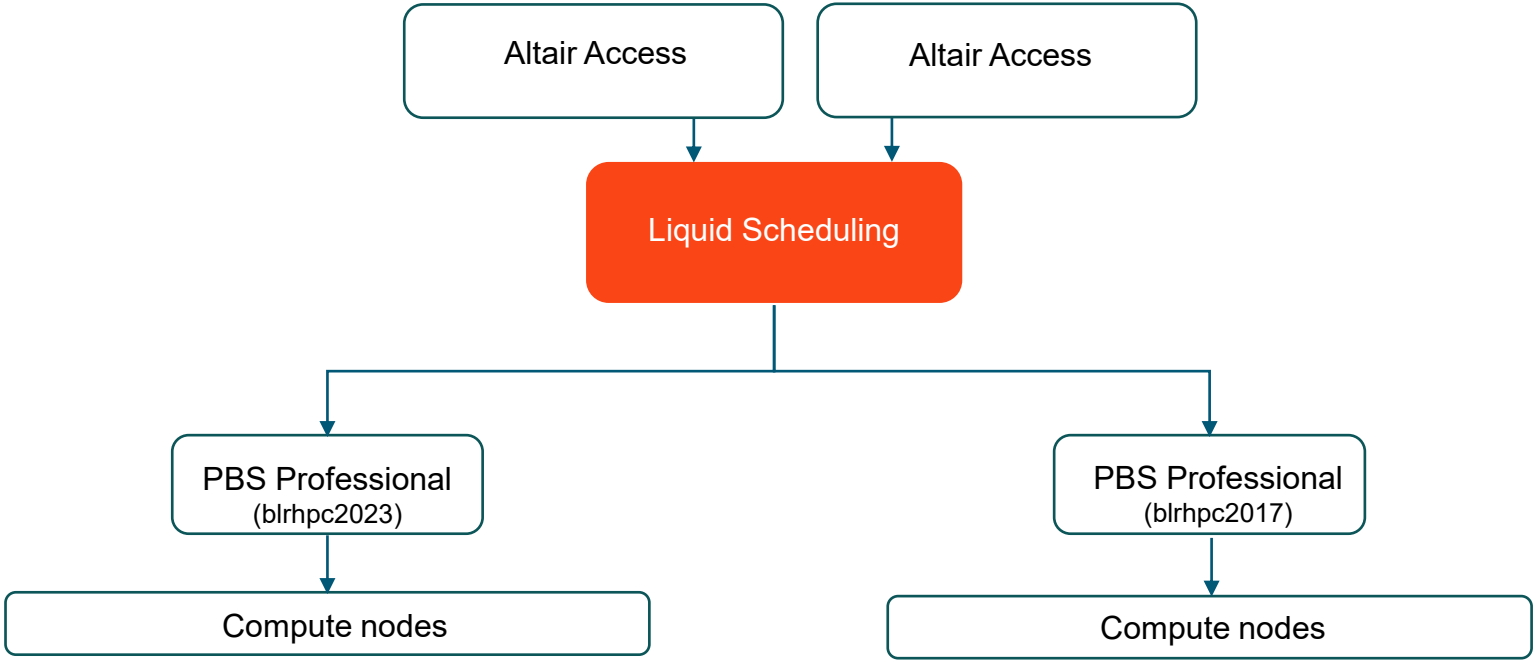
- ✓ Connecting distributed compute resources
- ✓ Users no longer have to choose where to run their jobs
- ✓ Fairshare correctly applied over all compute resources
- ✓ Overall system utilization is improved
- ✓ System monitoring and accounting is now unified

ALTAIR

Liquid Scheduling is already in production at Altair

- Liquid Scheduling has been deployed to unite some of our PDD clusters
- Integration with Altair® Access™ to leverage the whole Altair® HyperWorks® stack
- Users and applications migrated one at a time without disruption
- **Feedback – jobs run sooner, otherwise the system looks the same**

Liquid Scheduling on our PDD clusters

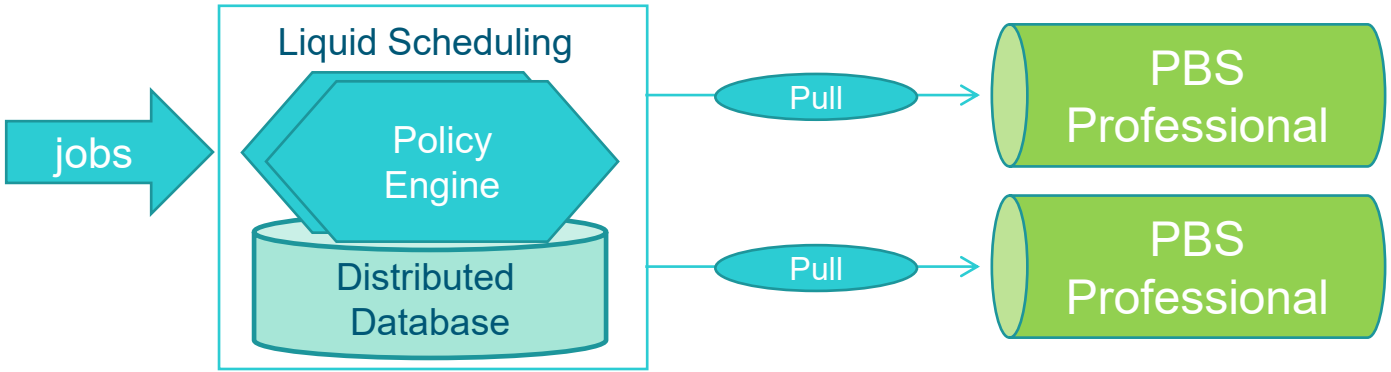


ALTAIR

Solution: Managing multiple workload managers

- ✓ Users no longer have to choose where to run their jobs
- ✓ Fairshare and quotas apply to all compute pools and are centrally managed
- ✓ Overall system utilization is improved
- ✓ System monitoring and accounting is now unified

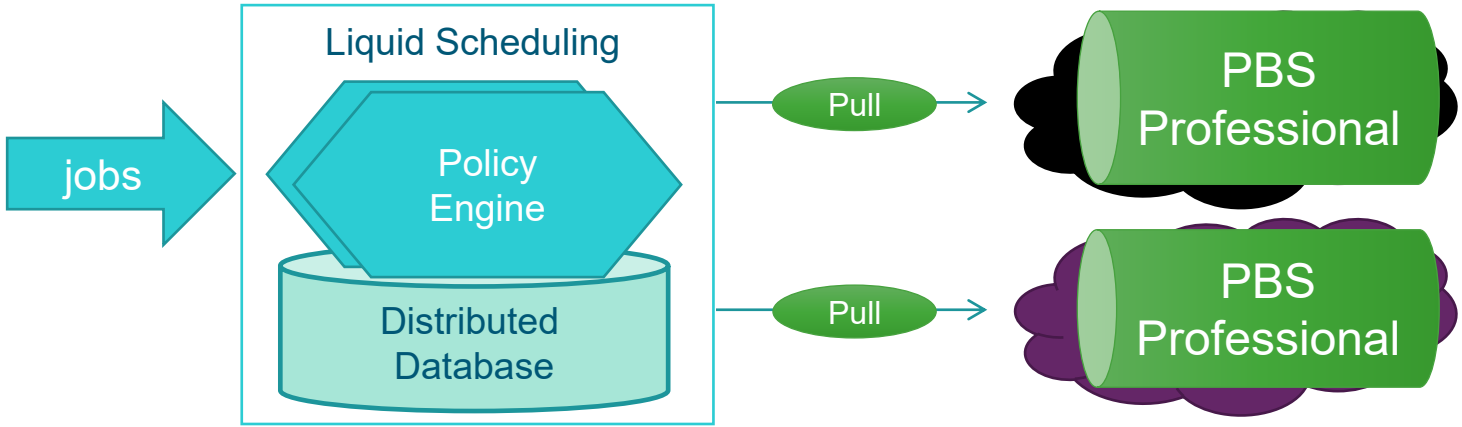
Submission through CLIs and APIs of our WLMs → no change for end users



ALTAIR

Example: Scaling in the cloud

- ❖ Workloads are bursty, very parallel, and business-critical
- ❖ Public cloud can't always deliver the scale needed
- ✓ Deploy Liquid Scheduling to distribute the workload across multiple cloud vendors



ALTAIR

Altair® Liquid Scheduling™

- **Makes Altair® PBS Professional® faster and more resilient,** even with only one cluster
- **Improves PBS Professional scaling** by breaking down silos between multiple clusters
- **Users and applications don’t need to change anything** to have their jobs run sooner#

Architected for the future of HPC

ALTAIR

CIUK 2023 Presentations

Alastair Basden (DiRAC / Durham University)

A foray into composable infrastructure for HPC

Abstract: Composable infrastructure decouples physical compute, storage and network fabrics, enabling software-defined hardware.

The DiRAC Memory Intensive service, COSMA, at Durham University has been running a production test-bed of small amounts of composable hardware for over two years, and here we present our findings, based on composable GPU and RAM systems.

We also present potential future plans and comment on suitability for HPC.

Bio: Alastair is the technical manager for the tier-1 national HPC COSMA service, with a background in astronomical instrumentation.

A foray into composable infrastructure for HPC

Alastair Basden, Peter Draper, Paul Walker, Mark Lovell, Richard Regan,Gokmen Kilic

DiRAC / Durham University
ExCALIBUR H&ES

CIUK 2023

DiRAC

Durham University

COSMA

ExCALIBUR

DiRAC

- UK national HPC service for STFC researchers
  - Tier-1 facility
- 4 sites:
  - Extreme Scaling @ Edinburgh
  - Data Intensive @ Leicester and Cambridge
  - Memory Intensive @ Durham
- Bespoke systems for the associated science
  - More cost effective than a single large system
  - Focus on Capability systems
    - For pushing the boundaries of what can be achieved

DiRAC
High Performance
Computing Facility

COSMA

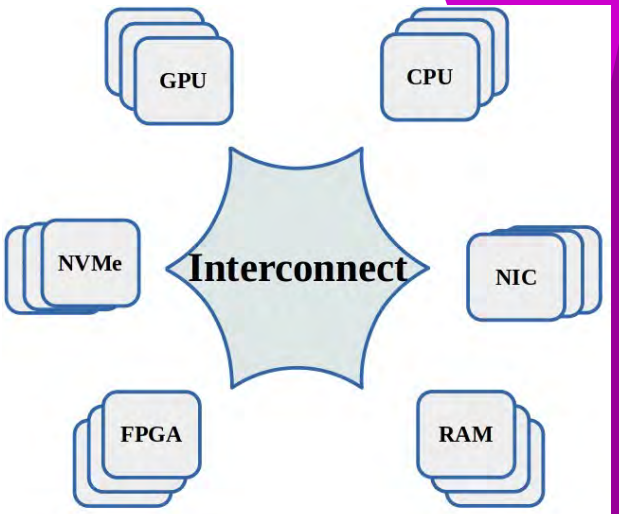
- COSMA7: 452 compute nodes (115kW total)
  - 28 cores, 512GB RAM (~95kW)
  - EDR InfiniBand (100Gb/s) and Rockport 100Gb/s (6kW)
  - Fat tree 2:1 blocking
  - 6PB storage, 420TB fast NVMe (15kW)
- COSMA8: 528 compute nodes (~300kW total)
  - ~70k cores (~250kW)
  - 128 cores, 1TB RAM per node (Rome/Milan)
  - HDR InfiniBand (200Gb/s) (~18kW)
    - Fat tree non-blocking
  - 15PB storage (20kW)
  - 1.2PB fast NVMe storage ~350GB/s (8kW)
  - Cooling distribution units (1kW)

COSMA science

- Primarily cosmology:
  - Simulation of the universe
- Also nuclear physics, particle physics, black holes, planetary collisions, galaxy formation

Composability

- Separation of device resources
  - Compute, accelerators, storage, network, RAM
  - Treated as services
- Physical components no longer in a server
  - Assigned to the server upon demand
    - Building compute capability as required
    - Clusters can be better matched to typical use cases
- Dynamically provision bare metal via software



Compostability

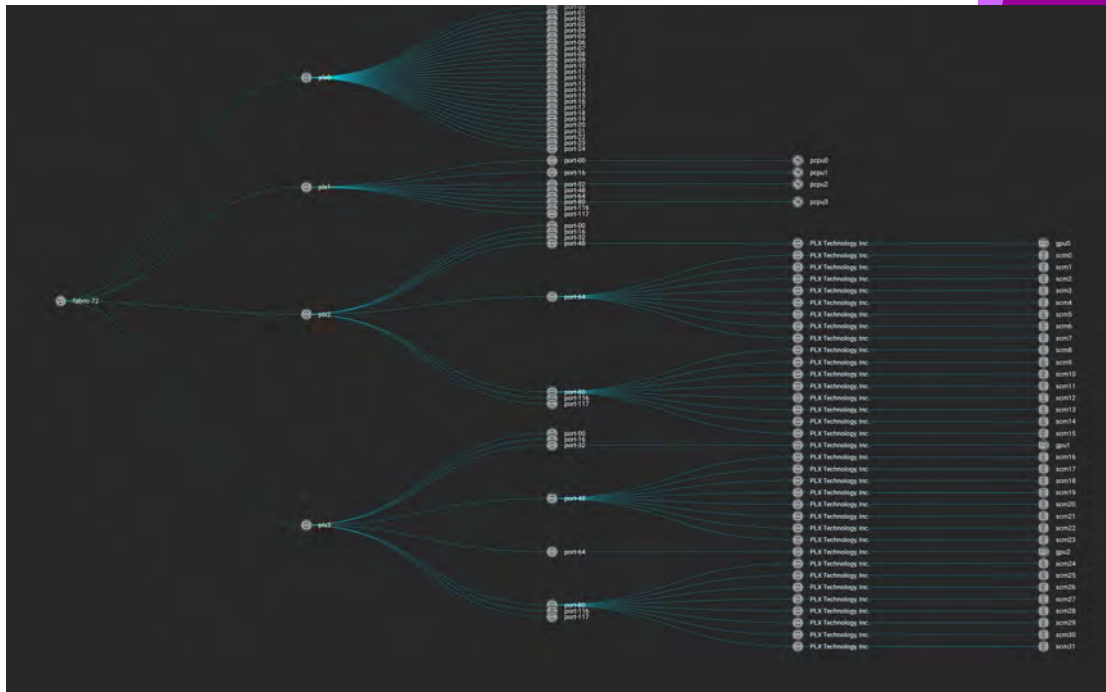
- Compostable infrastructure used for CIUK student cluster competition
- Not to be confused with composability
  - Similar aims (lower emboddied CO2, better resource use, etc)

Uses for composability in HPC/AI

- Massive GPU systems
  - 10s of GPUs per server
- Scarce resource sharing
  - Move GPUs as required to assigned servers
- Memory bursting
  - Adding RAM to servers as required
- Networking?
  - Composing multiple BlueField+GPU cards

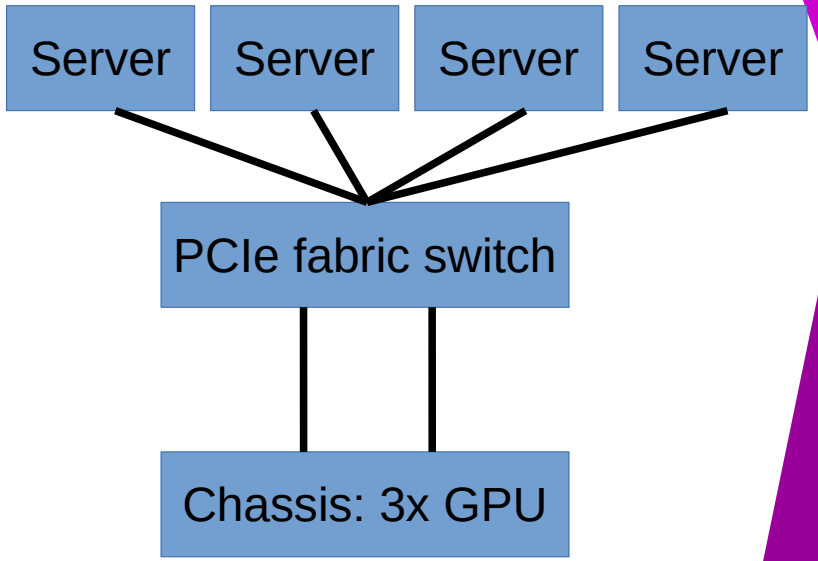
Composability on COSMA

- Liqid fabric:
  - One OSS chassis
    - 3 A100 GPUs (2021)
    - 4 RAM cards (3TB each) (2023)
  - PCIe switch and controller
  - 4 servers with fabric cards
    - One login node
    - 3 in a Slurm partition



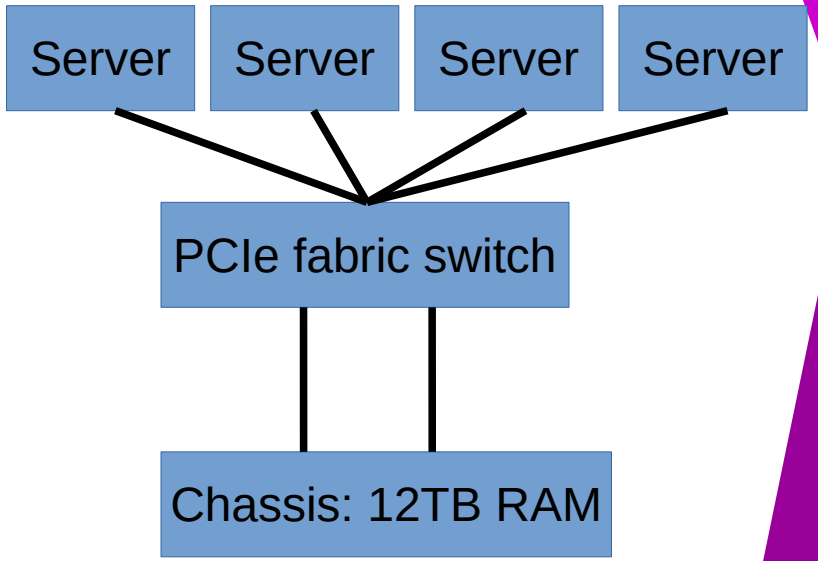
Liqid GPU system

- 3x A100 GPUs assignable to 4 servers
  - Can be hot-swapped
  - Occasionally causes problems
    - 3.10 kernel
- Usually static
  - One is a login node
- Slurm integration possible
  - Automatic provision
  - We use manual approach
- PCIe4 x4 connectors in each server
  - GPU bandwidth limited
  - 2 GPUs share 1 chassis card
- Physical connectivity is a pain
  - 4x SAS-type cables per card



Liqid RAM system

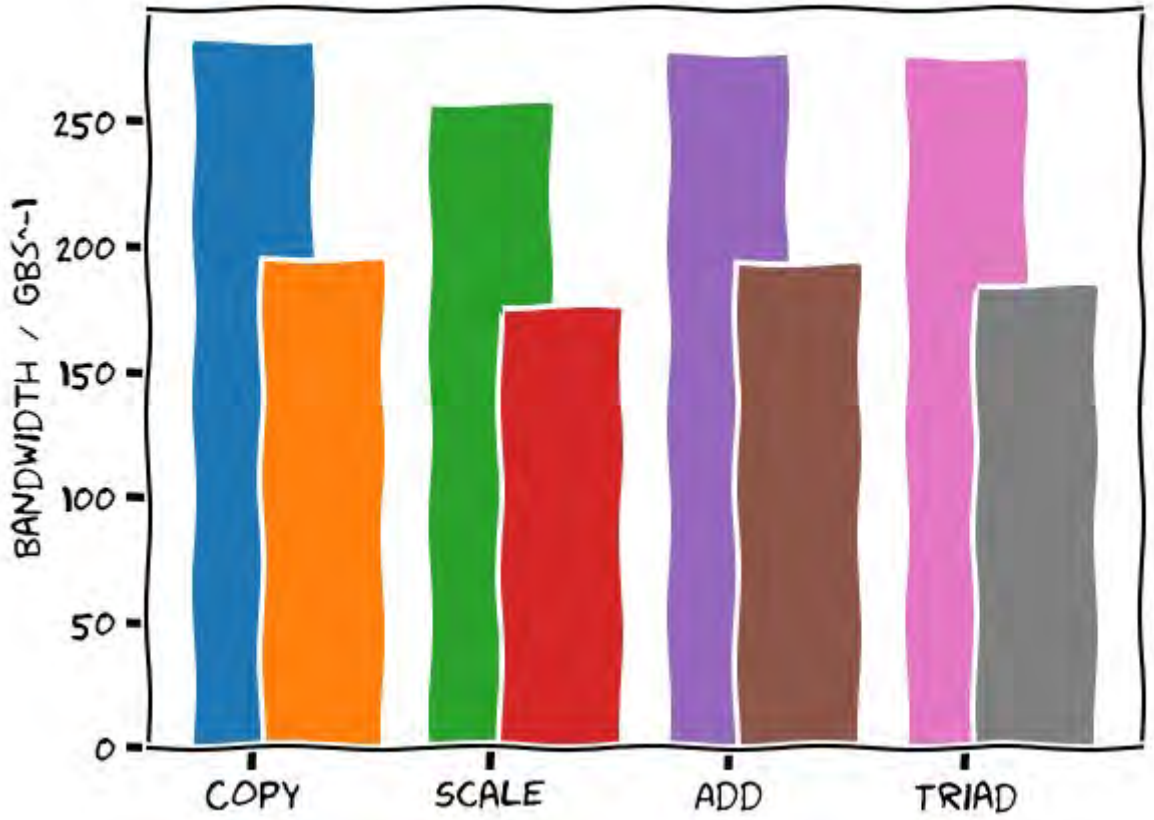
- 4x Honey Badger PCIe cards
  - Each with 8 M.2 NVMe Optane drives
    - 360GB - ~3TB total
  - 11.5TB combined Optane RAM
  - Each drive can be assigned individually to any server
  - Default configuration 8 drives/server (2 per HB)
- Memverge software to map Optane to RAM
  - Tiering of native RAM and Optane
    - Hot data kept in RAM, warm data moved to Optane
- Default setup adds ~2.5TB to each server
  - A bit of a pain to re-compose
- COSMA Jupyter hub can run on it
- Problems:
  - Lack of PCIe device entries in the BIOSs
  - Possible kernel bug
  - Power demands (native)



Memory bandwidth

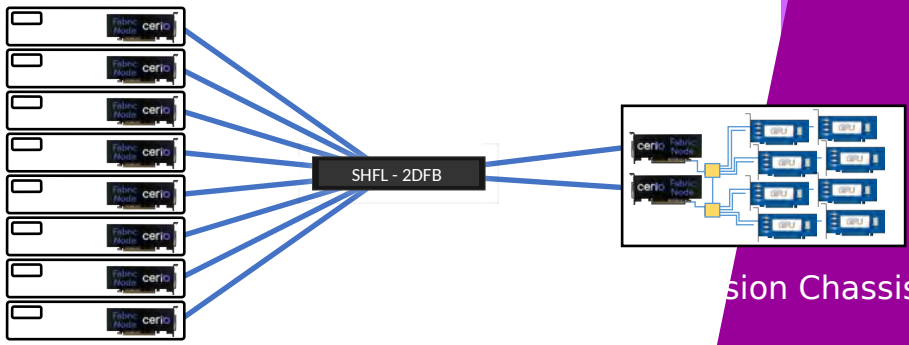
- Limited by the PCIe bandwidth (64Gb/s)
- MemVerge software used
  - tiered memory transfers between local DRAM and composed memory
  - Behind the scenes
- Memory allocations using a LD_PRELOAD
  - Prefix commands with “mm”
    - mm free -h, mm python jupyterlab, mpirun mm …, etc
    - Could be made the default option

STREAM benchmarks tests



Rockport Cerio system

- Moving from in-rack-scale to cluster-scale
  - PCIe-based fabrics don’t scale well
- Cerio flit-based fabric (6D torus) scales to thousands of nodes
  - Experience with the Rockport Ethernet fabric (COSMA7)
- High-density optical cables
  - Standard MTP connections
- Active components solely in server cards
  - “switch” not an active component
- 8-node test system in planning stage
  - 300GBit/s bandwidth, 200Gb/s to nodes
  - PCIe5

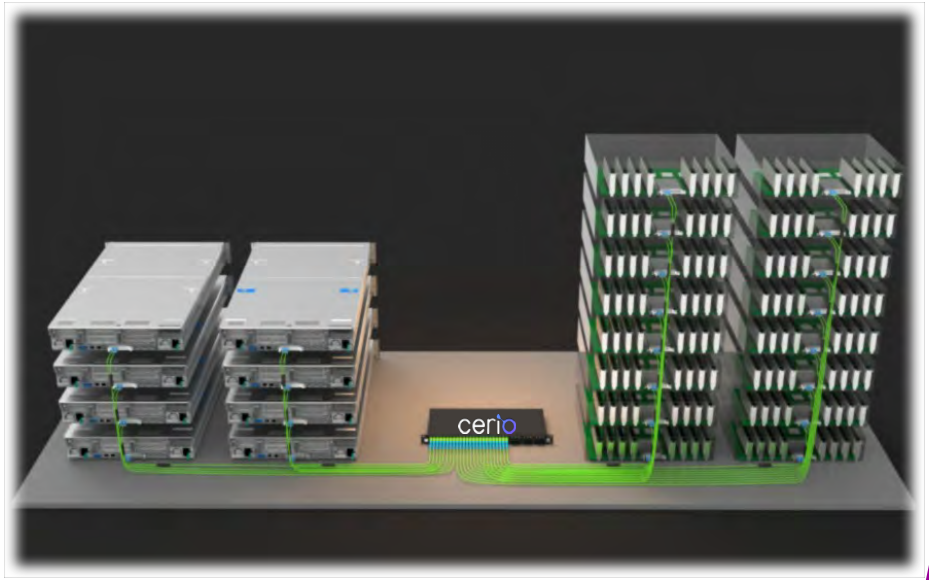


CXL and composability

- Compute eXpress Link
  - Open standard CPU to device and memory connections
- Brings memory coherence
  - Full composability of memory
  - Memory-coherent IO networking
    - Low latency communication - sub-mircosecond
  - /dev/shm spanning multiple servers
    - Multiple hosts work on same data without copying and shuffling

Considerations for composability

- Fabric reboots
  - Single point of failure
  - Cerio system should mitigate this to some extent
- Limited PCIe register slots
  - Determined by the BIOS
- Reliability
  - Particularly for RAM
- Cost: Not necessarily cheaper
- Power usage
  - Server thinks it is powering the cards
  - Custom firmware may be required
- Keeping track of infrastructure
  - Suspected dodgy components
- Not dynamically recoverable
- Bandwidth and latency hits
- Early days!



Net-zero considerations

- Significant potential
  - Lower embodied CO2 (less hardware)
  - Better match supply to demand
    - Less resource sitting idle
  - Expandable upon demand
    - Easy to add GPU resource as required by changing workloads

Conclusions

- Composability works
  - Currently a bit rough at the edges
- Lack of standards and flexibility
  - Vendor lock-in
  - Will hopefully improve
- RAM-based fabric could see significant performance improvement for some codes
  - CXL

CIUK 2023 Presentations

Dr Tim Bellerby (School of Environmental Sciences, University of Hull)

The PM Programming Language : Developing Numerical Models on Distributed Systems

Abstract: PM (Parallel Models) is a new open-source programming language targeting numerical model development on distributed systems. Most HPC software is created using combinations of languages and libraries: Fortran, C/C++, MPI, OpenMP, OpenACC, CUDA, etc.. These combinations can be difficult to program by non-specialists, introducing a barrier to HPC accessibility. They can also make assumptions about the underlying hardware, which can limit portability. The PM language is being developed to enable the programmer to concentrate on developing their model, rather than on making it run on particular hardware. A PM program will naturally scale from laptop to large cluster, taking full advantage of each platform, at the same time enabling full control over the distributed implementation when required. PM version 0.4 (available at www.pm-lang.org) includes both a PM to FORTRAN/MPI cross-compiler and a PM interpreter. A PM to FORTRAN/MPI/OpenMP cross-compiler and GPU support via OpenACC and/or OpenMP are planned for future releases.

Bio: Dr Tim Bellerby has been active in the fields of hydrology and meteorology for over 25 years, conducting research into new models and algorithms and managing the implementation of operational software systems. Ten years ago, as a result of insights gained in both research and operational program development and recognising the challenges that coding for distributed systems posed to non-specialists, he began development of the PM programming language.

Dr Bellerby received a BSc. in Applied Mathematics from the University of Warwick and a Ph.D. in Geophysics from the University of Sheffield. He has previously worked for the University of Bristol, NOAA and UN/FAO.

The PM Programming Language

IMPLEMENTING NUMERICAL MODELS ON DISTRIBUTED HARDWARE

Tim Bellerby
School of Environmental Sciences
University of Hull, UK

HULL VIPER HPC

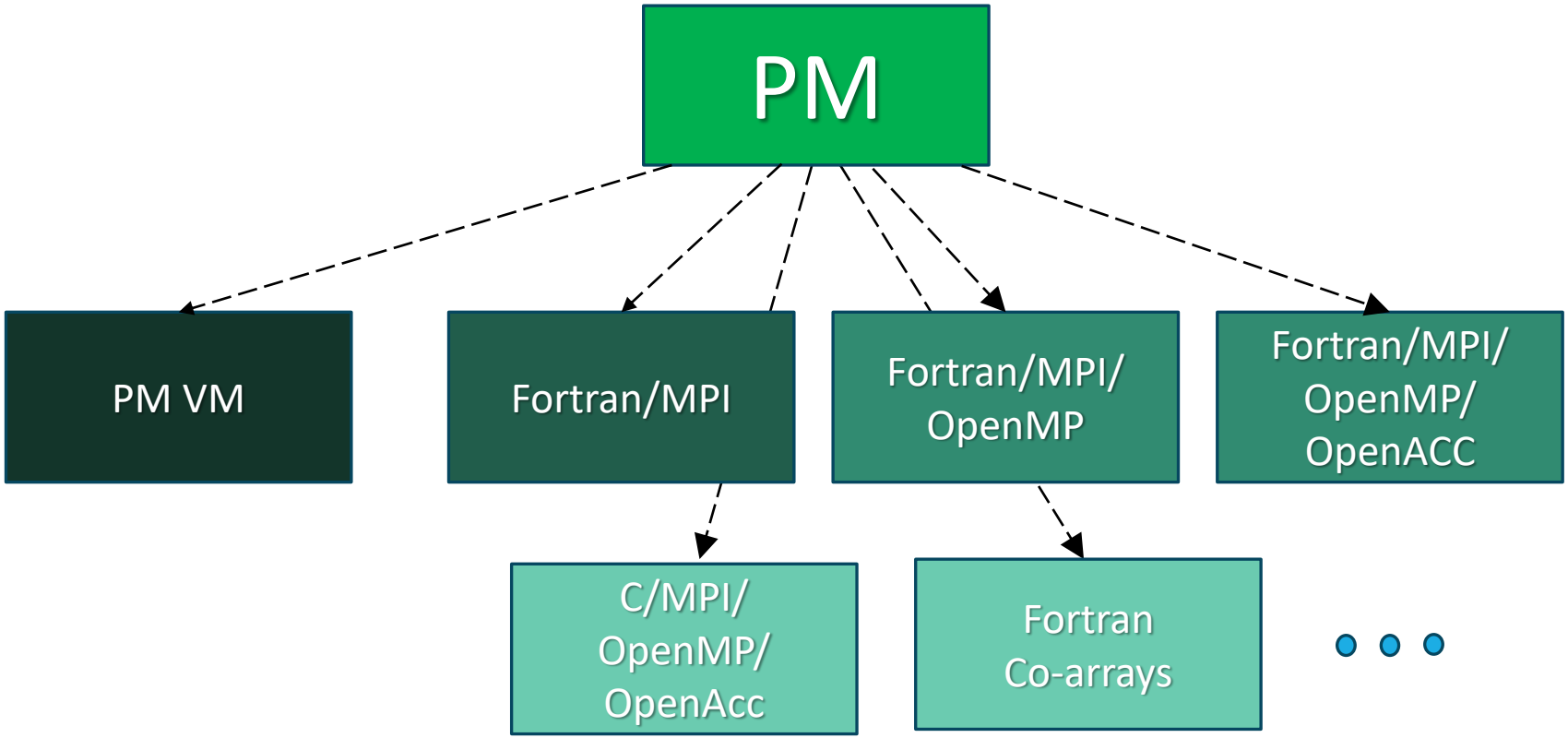
Research/HPC Software Divide

Python / Matlab / …

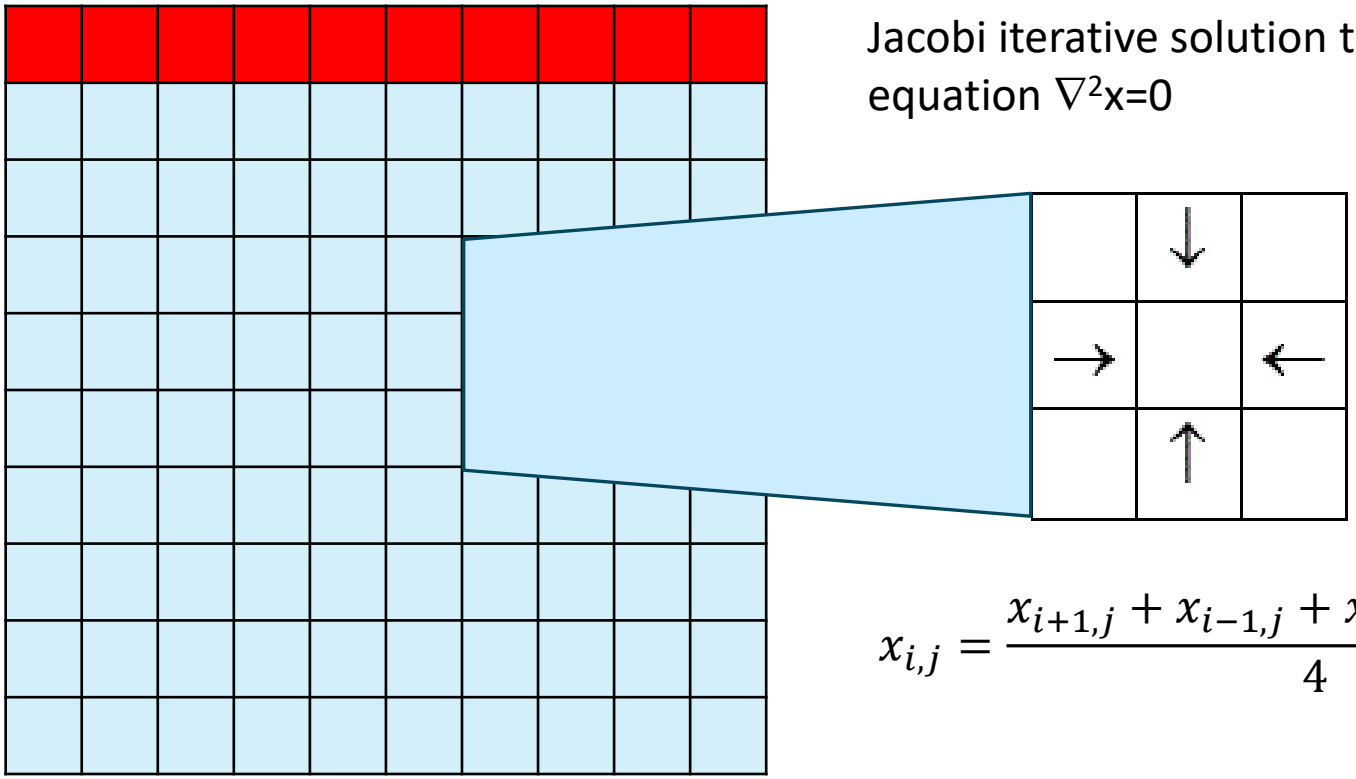
FORTRAN/C/C++
+ MPI + OpenMP
+ OpenAcc/CUDA/…

PM Programming Language

New programming language and programming language implementation designed for numerical modelling on distributed systems



An example model

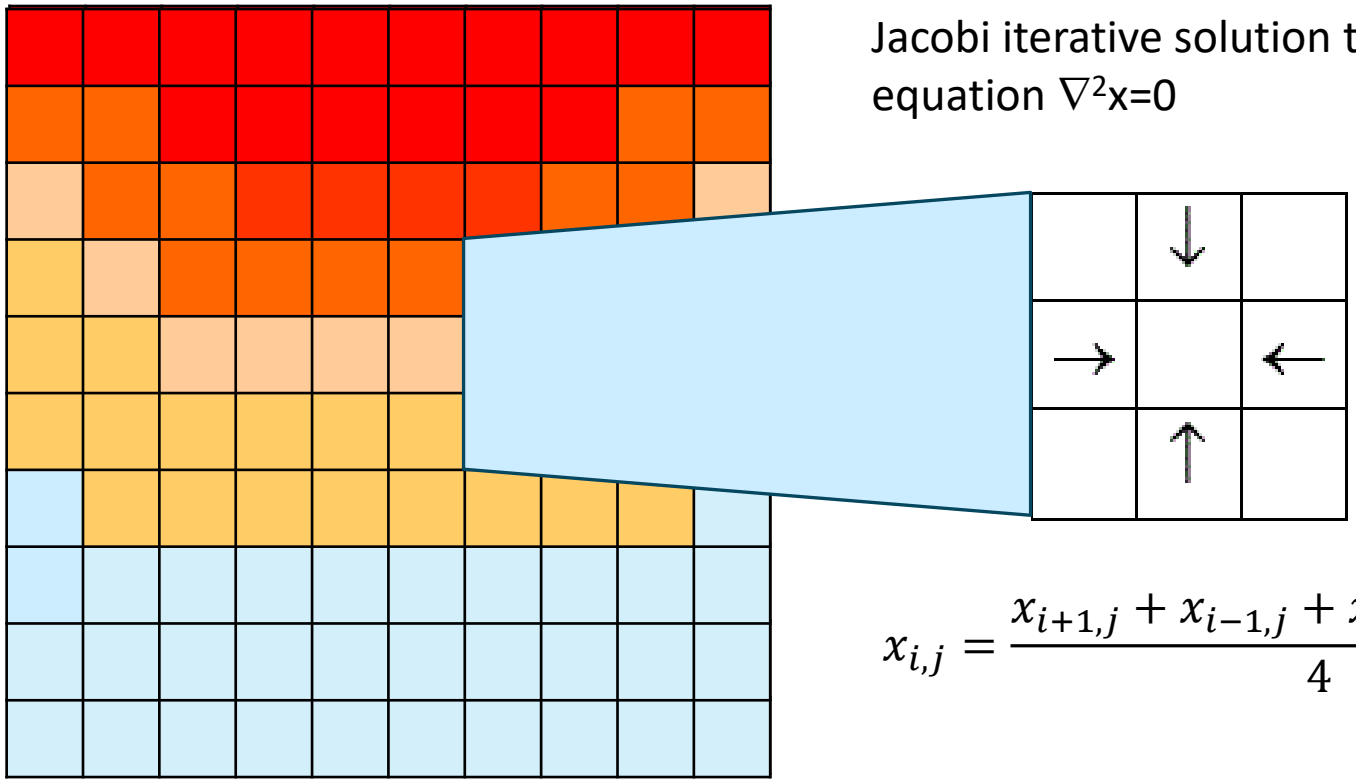


Jacobi iterative solution to 2D heat equation $\nabla^2 x=0$

$$x_{i,j} = \frac{x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1}}{4}$$

Repeat until no further change

An example model

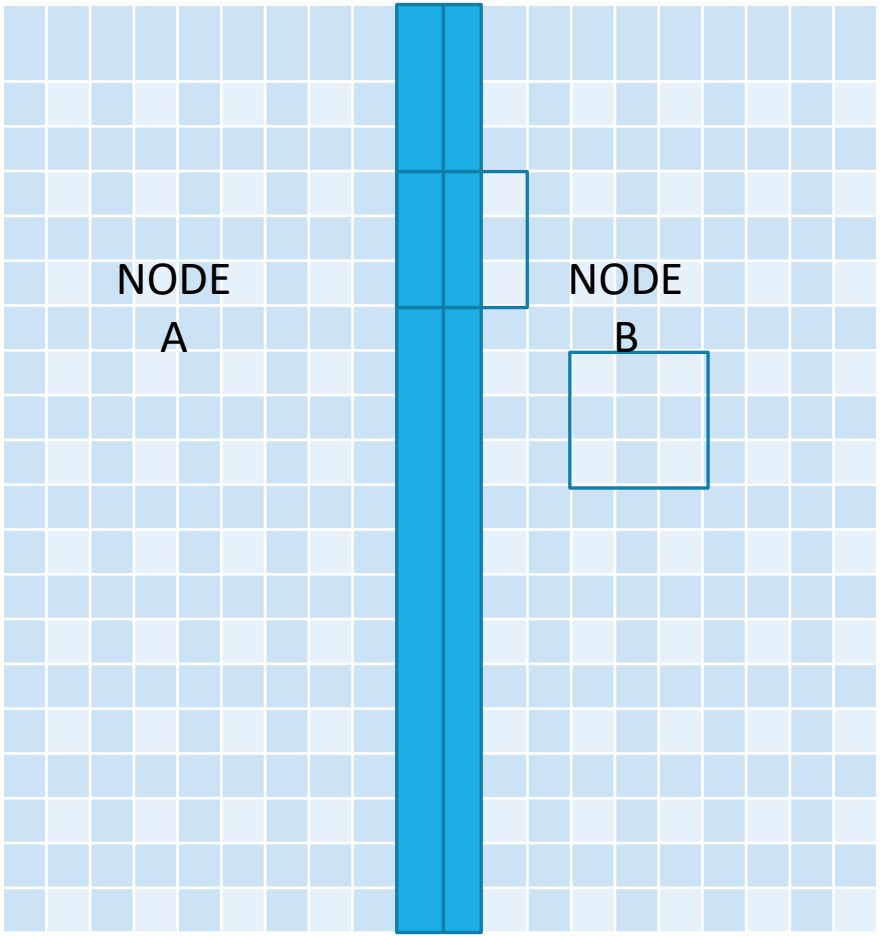


Jacobi iterative solution to 2D heat equation $\nabla^2 x=0$

$$x_{i,j} = \frac{x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1}}{4}$$

Repeat until no further change

Distribution



Coding models in PM

```
var out_array=darray(0.0,[0..MODEL_COLS+1,0..MODEL_ROWS+1])
for x in out_array {
   over [0,]: x=1.0
   until invar totdiff<TOL {
     var diff=0.0
     nhd [-1..1,-1..1] dx of x bounds EXCLUDED {
      cell=(dx[-1,0]+dx[1,0]+dx[0,-1]+dx[0,1])/4.0
      diff=(x-dx[0,0])**2
     }
     totdiff=sqrt(sum%(diff)/size(out_array))
   }
}
```

Coding models in PM

```
var out_array=darray(0.0,[0..MODEL_COLS+1,0..MODEL_ROWS+1])
for x in out_array {
    over [0,]: x=1.0
    until invar totdiff<TOL
      var diff=0.0
      nhd [-1..1,-1..1] dx
       cell=(dx[-1,0]+dx[1,0]+dx[0,-1]+dx[0,1])/4.0
       diff=(x-dx[0,0])**2
      }
      totdiff=sqrt(sum%(diff)/size(out_array))
    }
}
```

Extensive, optional, compile-time type inference

Coding models in PM

```
var out_array=darray(0.0,[0..MODEL_COLS+1,0..MODEL_ROWS+1])
for x in out_array {
  over [0,]: x=1.0
  until invar totdiff<TOL {
    [] dx of x bounds EXCLUDED {
      dx=(dx[1,0]+dx[0,-1]+dx[0,1])/4.0
      diff=(x-dx[0,0])**2
    }
    totdiff=sqrt(sum%(diff)/size(out_array))
  }
}
```

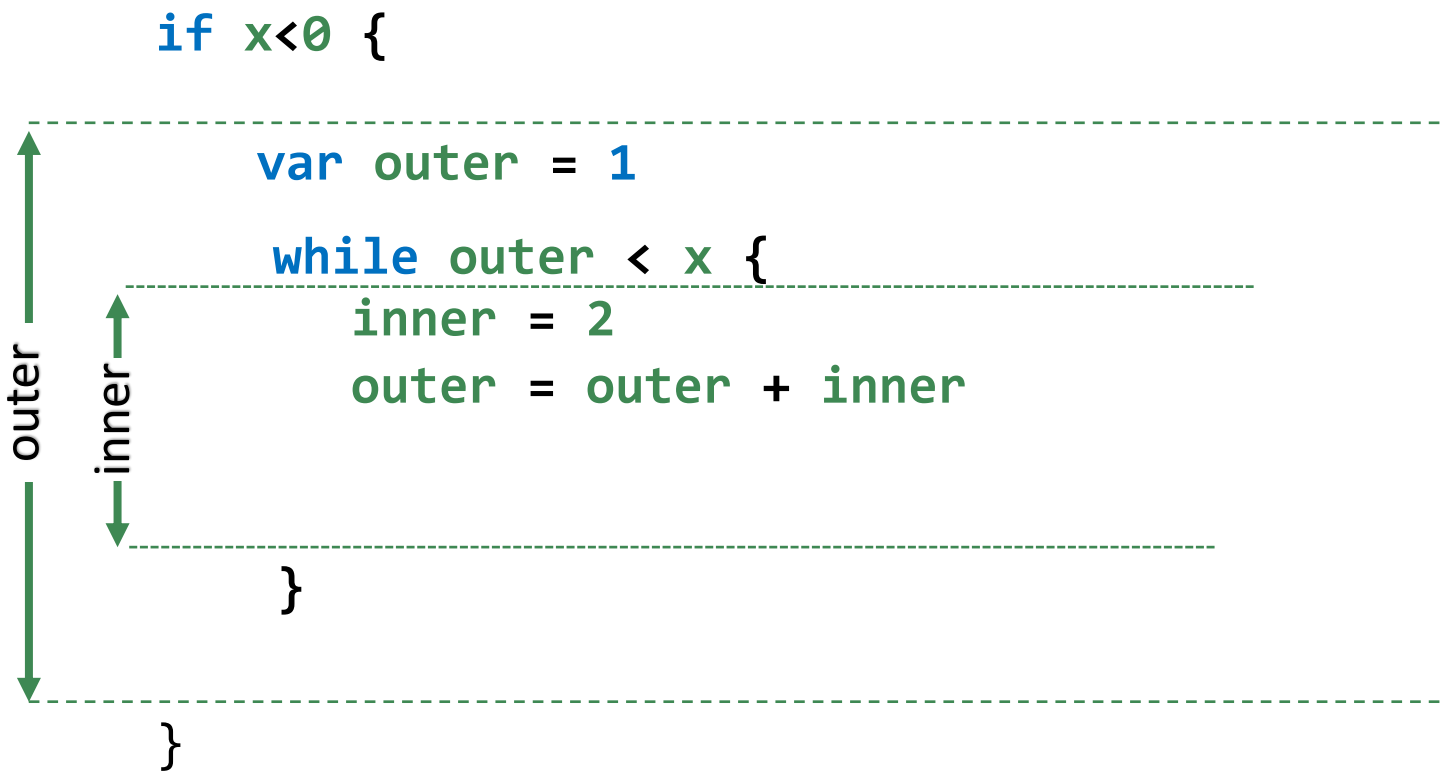
Explicitly parallel statements

Coding models in PM

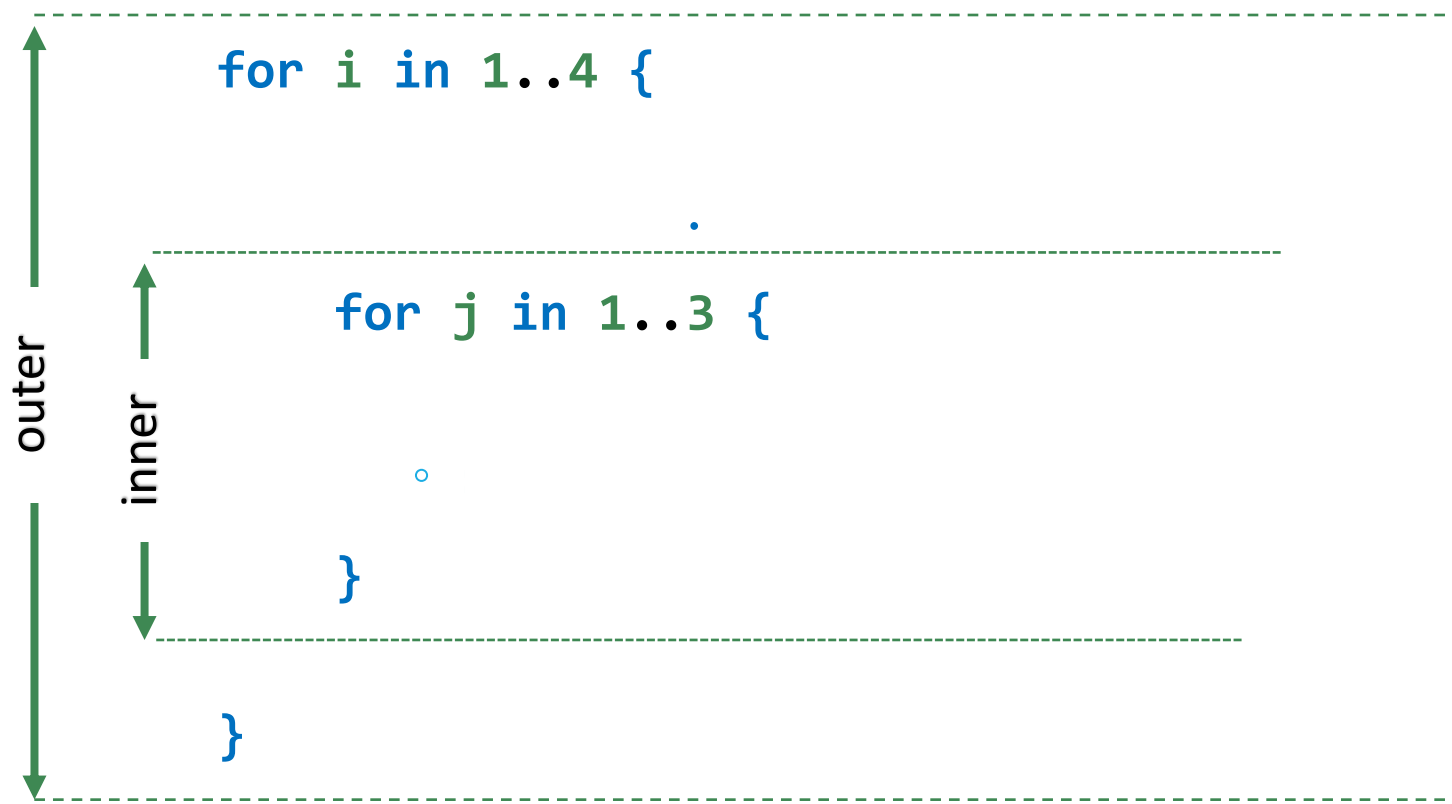
```
var out_array=darray(0.0,[0..MODEL_COLS+1,0..MODEL_ROWS+1])
for x in out_array {
   over [0,]: x=1.0
   until invar totdiff<TOL {
     var diff=0.0
     nhd [-1..1,-1..1] dx of x bounds EXCLUDED {
      cell=(dx[-1,0]+dx[1,0]+dx[0,-1]+dx[0,1])/4.0
      diff=(x-dx[0,0])**2
     }
     totdiff=sqrt(sum%(diff)/size(out_array))
   }
}
```

Operations that communicate between parallel strands

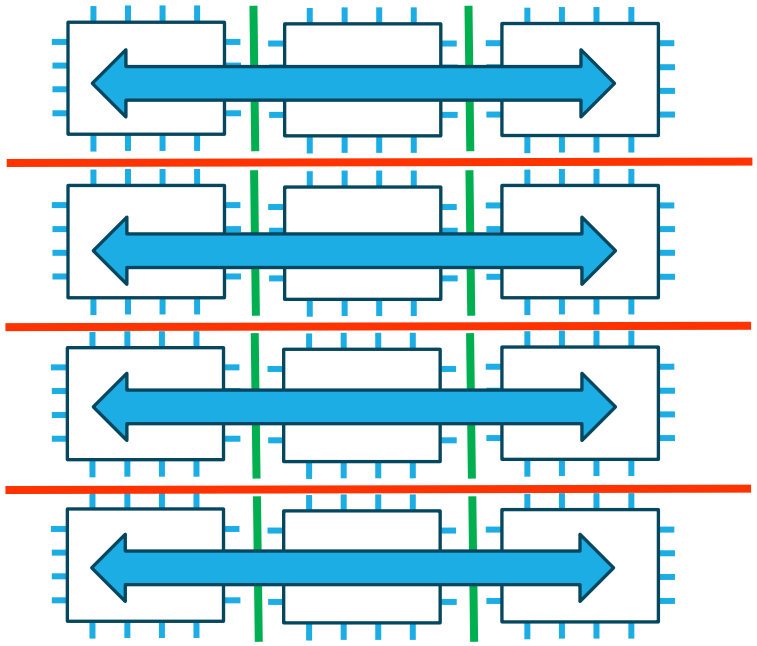
Scope


if x<0 {
var outer = 1
while outer < x {
inner = 2
outer = outer + inner
}
}
outer
inner

Parallel Scope

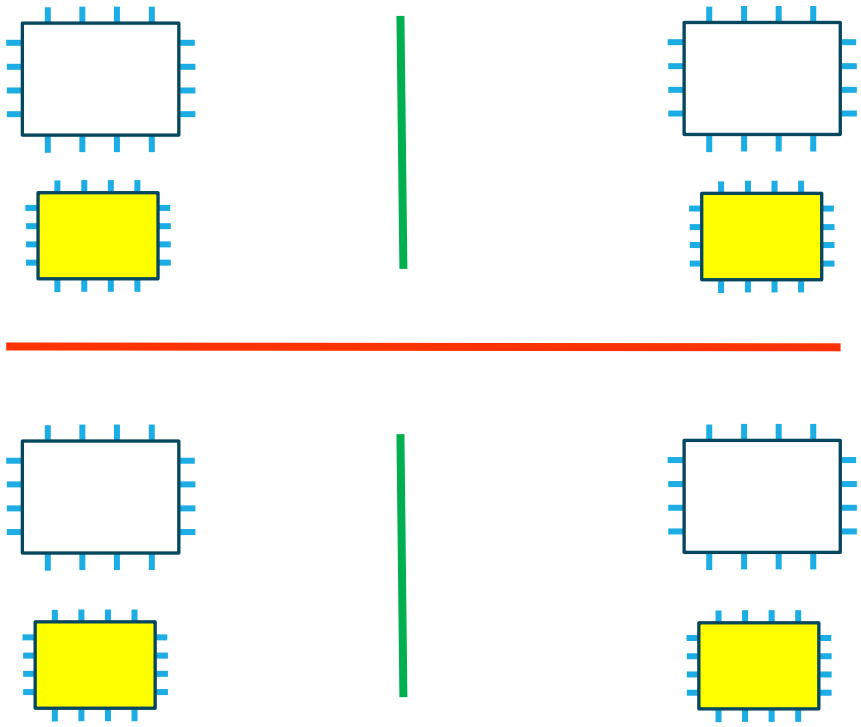


Parallel Scope



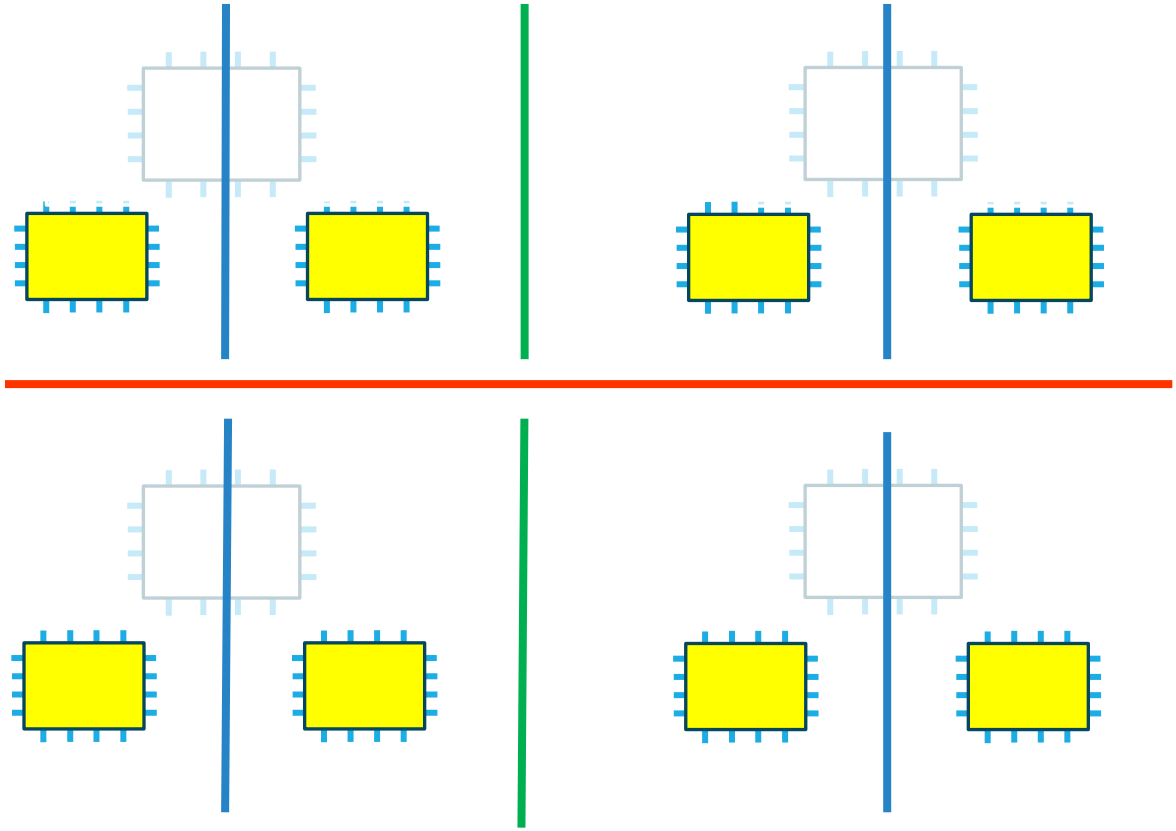
```
for i in 1..4 {

    for j in 1..3 {

    }

}
```

Adding accelerators*

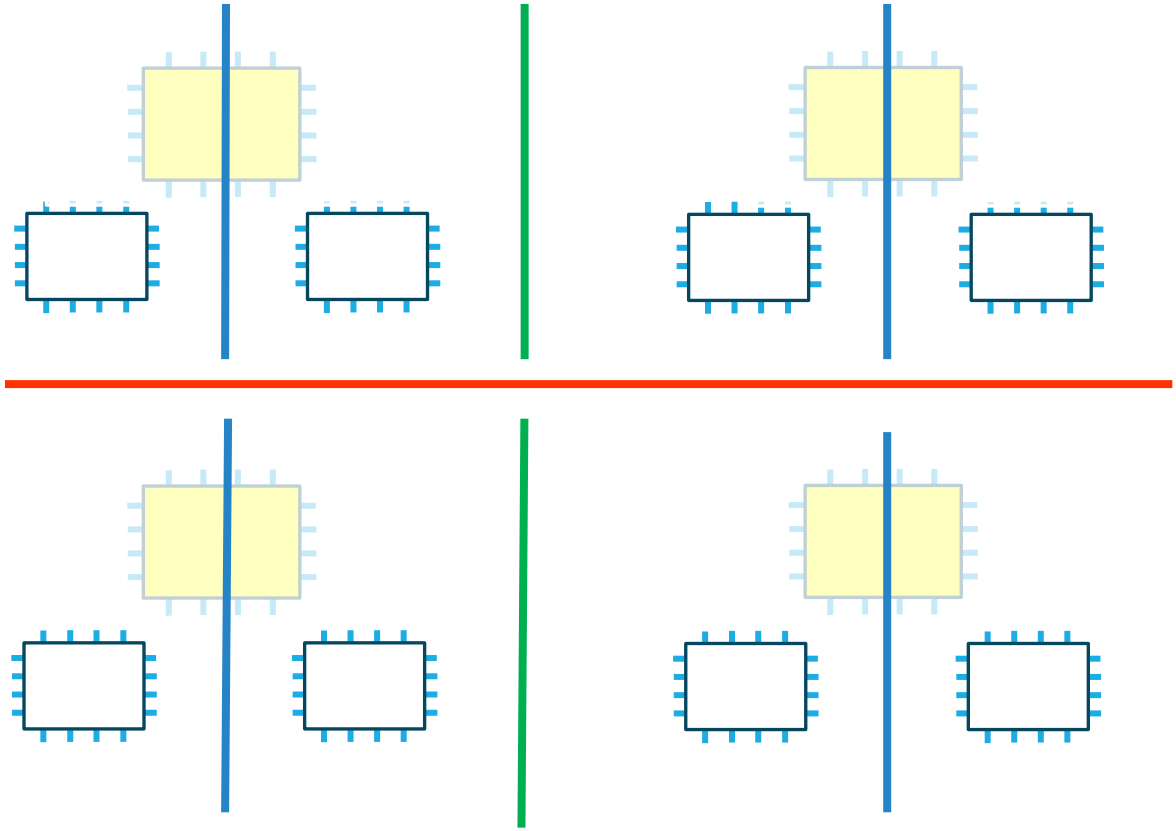


* Planned

Adding accelerators



Adding accelerators



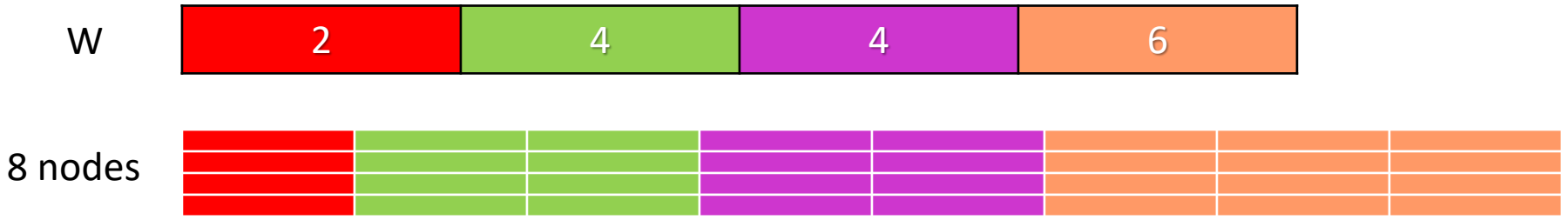
Assigning nodes

```
for i in 1..4 {
    for j = 1..8 {
        ...
    }
}
```

8 nodes

Assigning nodes

```
W={2,4,4,6}
for i in 1..4 <<work=W>> {
    for j in 1..W[i]*2 {
        ...
    }
}
```



Assigning nodes

```
proc process_node(node) {
   …  // Check for and process leaves
   par {
   task LEFT <<work=node.left.nchildren>>:
        process_node(node.left)
   task RIGHT <<work=node.right.nchildren>>:
        process_node(node.right)
   }
}
```

Distributions

10 array elements on 3 processors

Variable Block

Fixed Block (block=4)

Cyclic

Block cyclic (block=2)

Distributions

2D block cyclic execution with 2x2 block size over 7x8 grid

1,1	2,1	3,1	4,1	5,1	6,1	7,1
1,2	2,2	3,2	4,2	5,2	6,2	7,2
1,3	2,3	3,3`	4,3	5,3	6,3	7,3
1,4	2,4	3,4	4,4	5,4	6,4	7,4
1,5	2,5	3,5	4,5	5,5	6,5	7,5
1,6	2,6	3,6	4,6	5,6	6,6	7,6
1,7	2,7	3,7	4,7	5,7	6,7	7,7
1,8	2,8	3,8	4,8	5,8	6,8	7,8

```
for j in [1..7,1..8] <<dist = BLOCK_CYCLIC(2)>> {
    …
}
```

Distributions

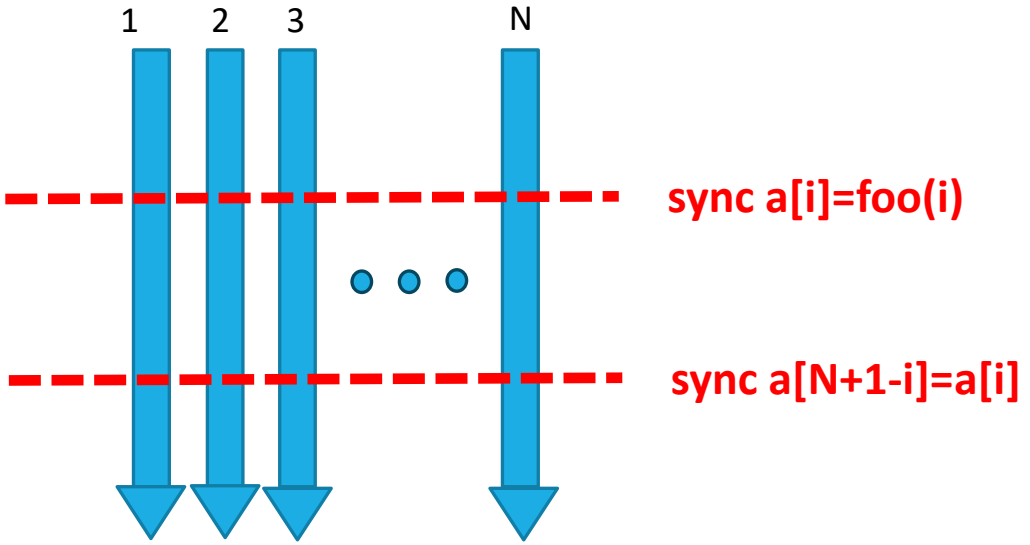
2D block cyclic array with block size of two on first dimension only over 7x8 grid

1,1	2,1	3,1	4,1	5,1	6,1	7,1
1,2	2,2	3,2	4,2	5,2	6,2	7,2
1,3	2,3	3,3`	4,3	5,3	6,3	7,3
1,4	2,4	3,4	4,4	5,4	6,4	7,4
1,5	2,5	3,5	4,5	5,5	6,5	7,5
1,6	2,6	3,6	4,6	5,6	6,6	7,6
1,7	2,7	3,7	4,7	5,7	6,7	7,7
1,8	2,8	3,8	4,8	5,8	6,8	7,8

```
A = array (0.0,[1..7,1..8],[BLOCK_CYCLIC(2),])
```

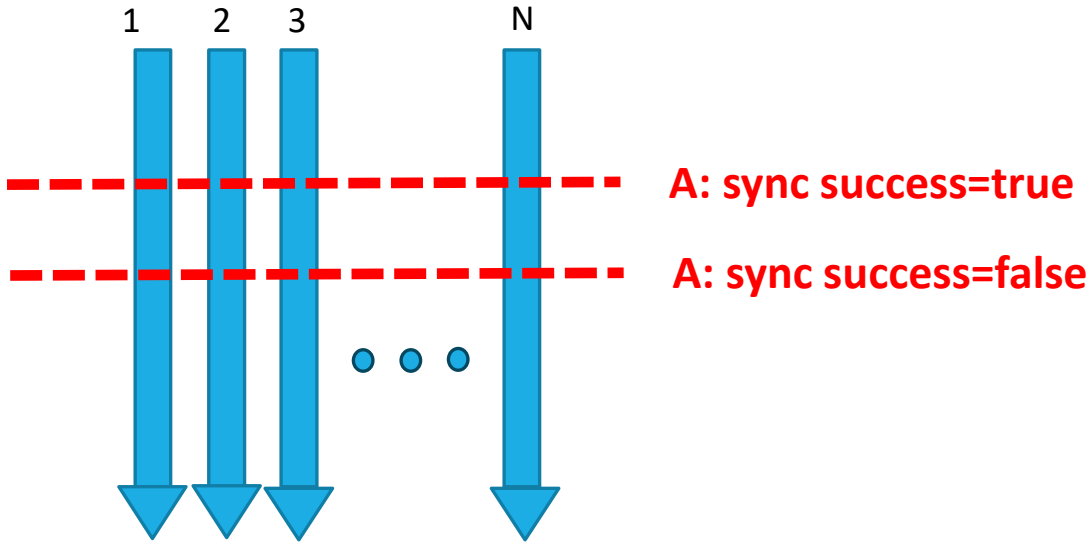
Synchronisation

```
var a = array(0.0,[1..N])
for i in 1..N {
    sync a[i]=foo(i)
     …
    sync a[N+1-i]=a[i]
     …
}
```



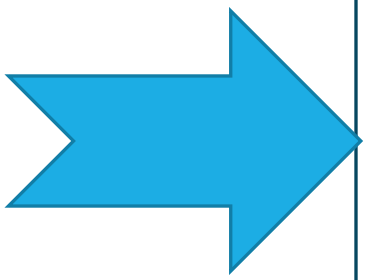
Synchronisation

```
var success = false
for i in 1..N {
   if criterion_met(i) {
        A:  sync success=true
   } elseif blocking_criterion_met(i) {
        A:  sync success=false
    }
}
```



Implementation

```
for ... {
    if E1 {
        S1
        L: sync a= ...
        S2
    } else {
        S3
        L: sync b = ...
        S4
    }
}
```



```
DIMENSION ...
!OMP PARALLEL
DO ...
    IF E1 THEN
        S1
    ELSE
        S2
    ENDIF
ENDDO
!OMP END PARALLEL
 A= -> MPI
 B= -> MPI
!OMP PARALLEL
DO ...
    IF E1 THEN
        S3
    ELSE
        S4
    ENDIF
ENDDO
!OMP END PARALLEL
```

Scheduling

Reorder instructions to:

- Minimise loop start/stops
- Minimise storage requirement
- Interleave computation and communication
- Merge synchronisation points

iSend A
iRecv B
Wait
Process B
iSend C
iRecv D
Wait
Process D
Process E

iRecv B
iRecv D
iSend A
iSend C
Process E
Wait
Process B
Process D

PM Version 0.4

Formal language definition released under Creative Commons Attribution 4.0 International License.

Language Implementation:

- PM to FORTRAN/MPI compiler
- Vector-virtual machine (intended for development/debugging)
- Both available under MIT Licence

PM Version 0.5

A small number of language additions

- Closures
- Sparse Arrays
- Interoperability with C/FORTRAN

Added backends

- FORTRAN/MPI/OpenMP
- FORTRAN/MPI/OpenMP/OpenAcc

Planned summer 2024.

Thank you for your attention

Questions? T.J.Bellerby@hull.ac.uk

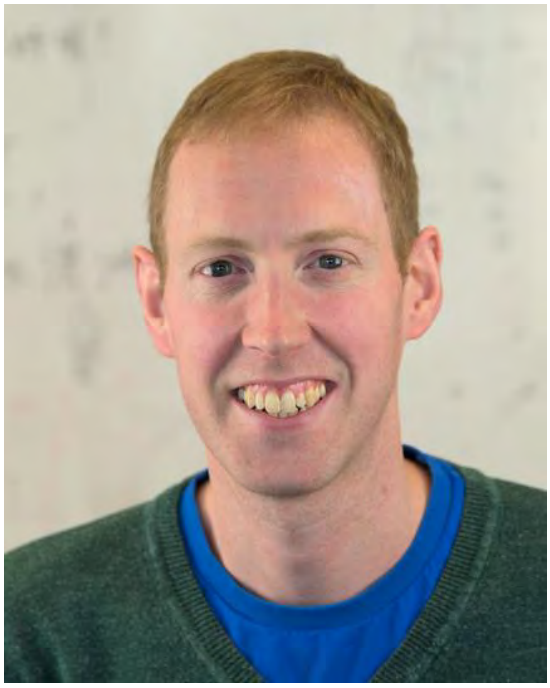
Follow progress: www.pm-lang.org

𝕏 @pmlanguage

CIUK 2023 Presentations

Nick Brown (EPCC at the University of Edinburgh)

Making HPC more accessible: Effective HPC programming via domain specific abstractions



Abstract: Effectively programming modern day supercomputers is very difficult, and this problem is going to become even more acute as we move further into the exascale era and our HPC machines become more complex. A potential solution is that of Domain Specific Languages (DSLs), which enable users to describe their solution at a high level, and then using this rich source of information the compiler can make decisions around the tricky, low level, details of parallelism. However, a major challenge to the adoption of DSLs is the fact that the underlying compilation infrastructure is often supported by only one group, can be immature, and without any guarantee around long term maintenance or support.

In the ExCALIBUR xDSL project we are exposing the ubiquitous MLIR and LLVM frameworks as a shared ecosystem, enabling DSLs to be a thin abstraction layer atop this very well supported and actively developed infrastructure. This talk will share with the audience some very exciting activities that are currently ongoing in the field of compilers and how this momentum will likely deliver a revolution in how we program future exascale supercomputers, ultimately democratising HPC and making it more accessible to a wider audience.

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.

Making HPC more accessible: Effective HPC programming via domain specific abstractions

Nick Brown
n.brown@epcc.ed.ac.uk

Emilien Bauer
emilien.bauer@ed.ac.uk

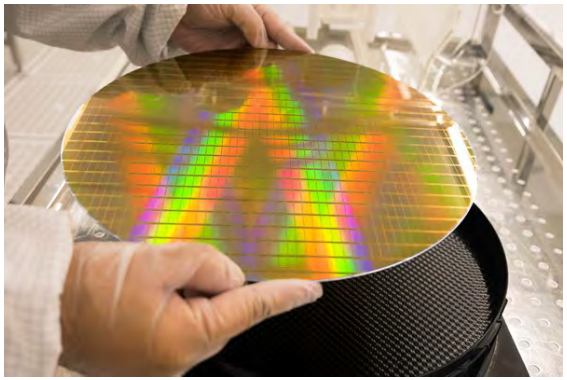
epcc

THE UNIVERSITY of EDINBURGH
informatics

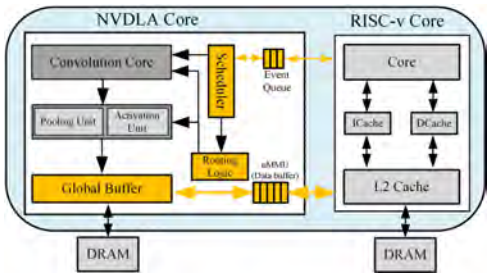
THE UNIVERSITY of EDINBURGH

The challenge

- Writing parallel code that can exploit present day supercomputers is extremely hard and requires highly specialist skills



- But this is going to get even more difficult as we move further into the exascale era
- It is no longer tenable to directly leverage serial languages and add in our own parallelism (e.g. MPI, CUDA, vectorisation etc)



Domain Specific Languages to the rescue!

GridTools

OP-DSL

- Raise the abstraction level so the programmer can provide a high level description of their algorithm that the compiler can then exploit to make tricky, low level decisions around parallelism

Saiph

PSyclone

- *Languages* is a poor term, *abstractions* is far better

ANTAREX 10^18

NMODL
Code Generation Framework

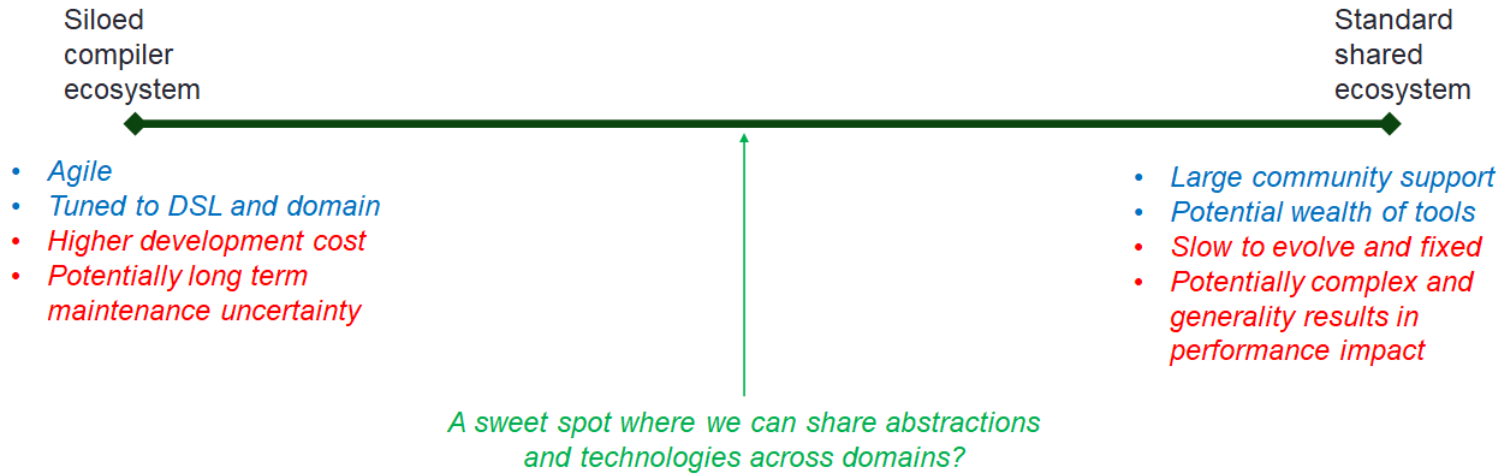
CLAW

Halide
a language for fast, portable computation on images and tensors

ExCALIBUR

Breaking down silos

- The elephant in the room is that these are all heavily siloed and reinvent the wheel
  - Requires significant development effort from the DSL designers
  - Risk for users (e.g. will the DSL be maintained in the future?)
  - Challenges supporting new architectures



There is therefore a sweet spot in the middle, where we gain the best of both worlds

Step in MLIR and LLVM

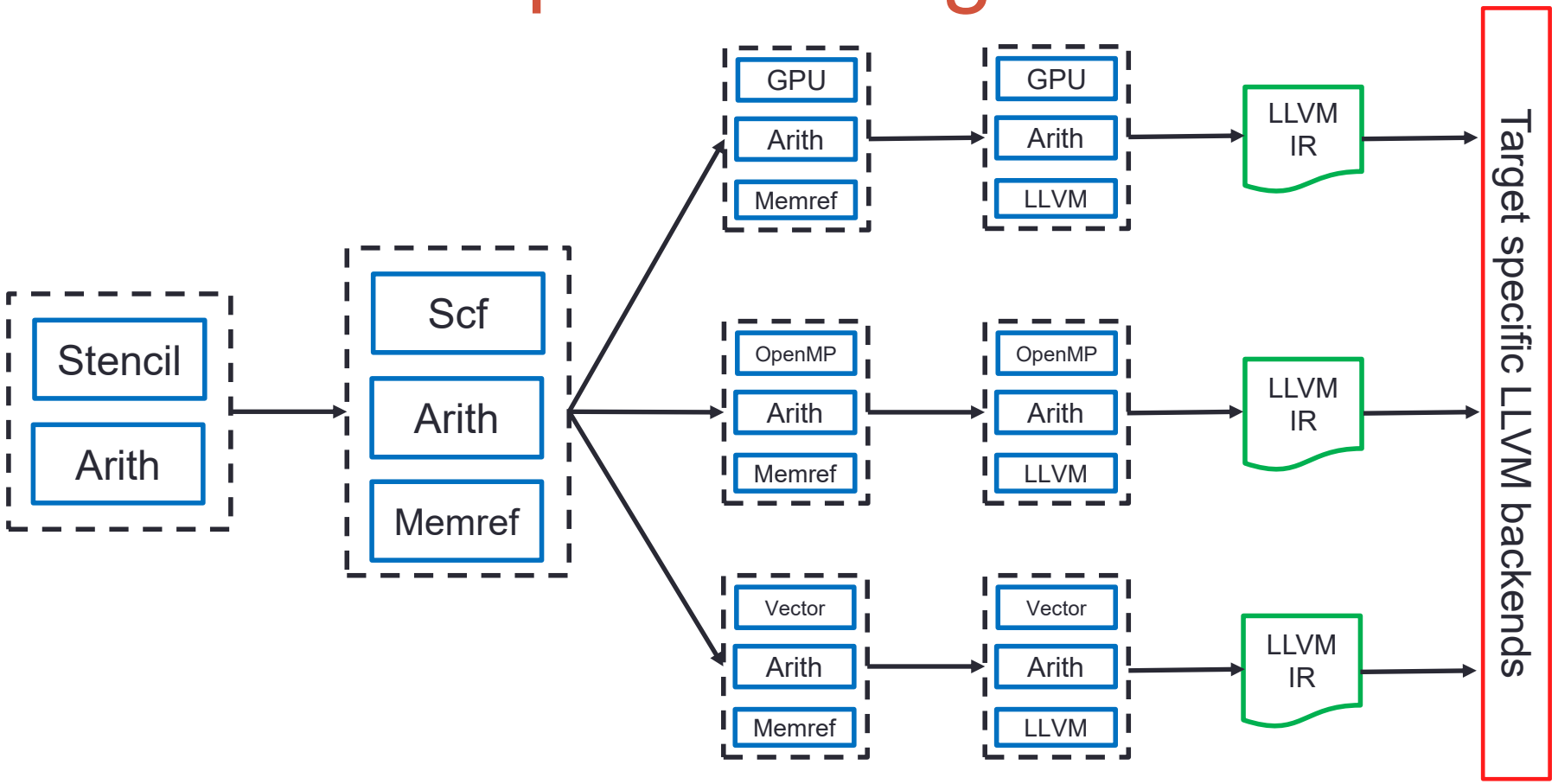
- LLVM is the ubiquitous compiler framework that has been around for over 20 years
  - In addition to providing its own compilers, AMD, Intel and Arm compilers are all built on-top of LLVM, as is the Cray C/C++ compiler and AMD Xilinx's FPGA HLS technology.

- MLIR was developed by Google in 2020 and since 2021 has been part of the main LLVM repository

MLIR

- At its core MLIR is a framework for developing different types of Intermediate Representations (IR) at different levels
- Numerous (IR) dialects and transformations are provided which enables lowering between these
- Can add your own easily
- A big community has grown up

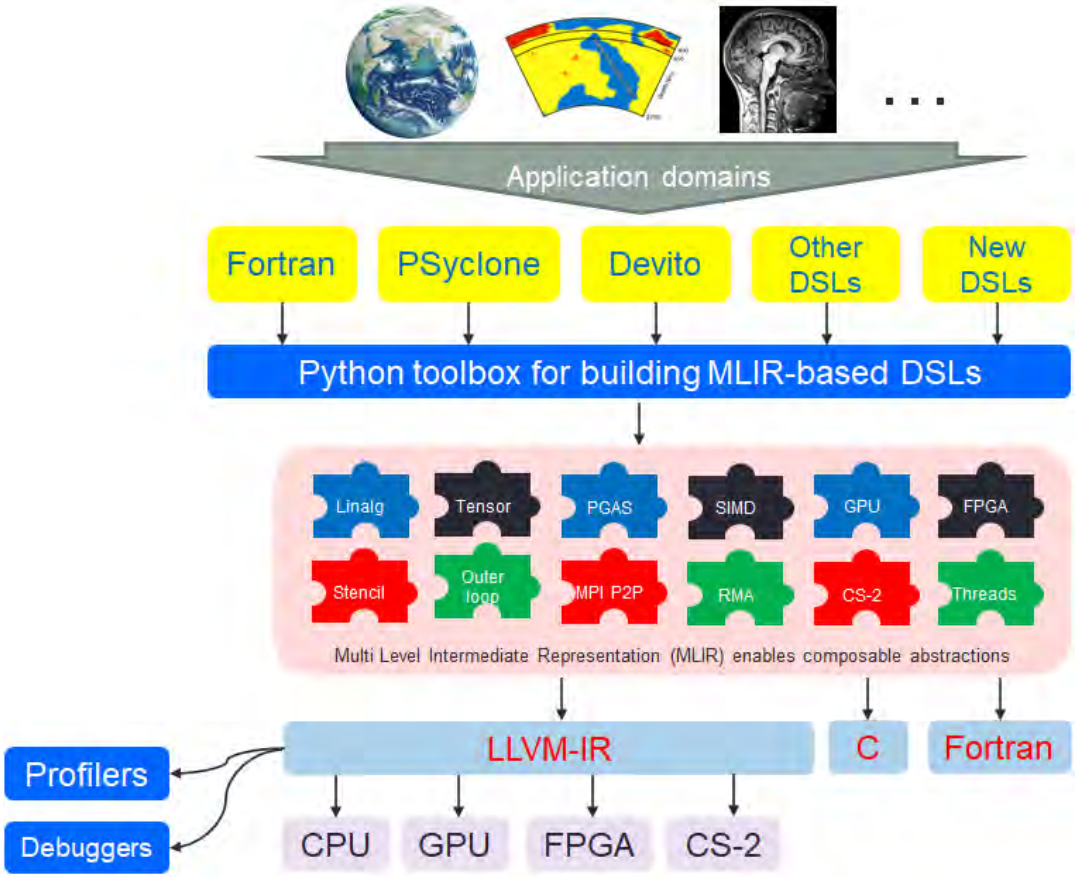
MLIR example lowering



- But MLIR is written in C++ and using specialist Tablegen configuration format for dialects, MLIR is esoteric and requires a steep learning curve

xDSL: A Python toolkit for MLIR

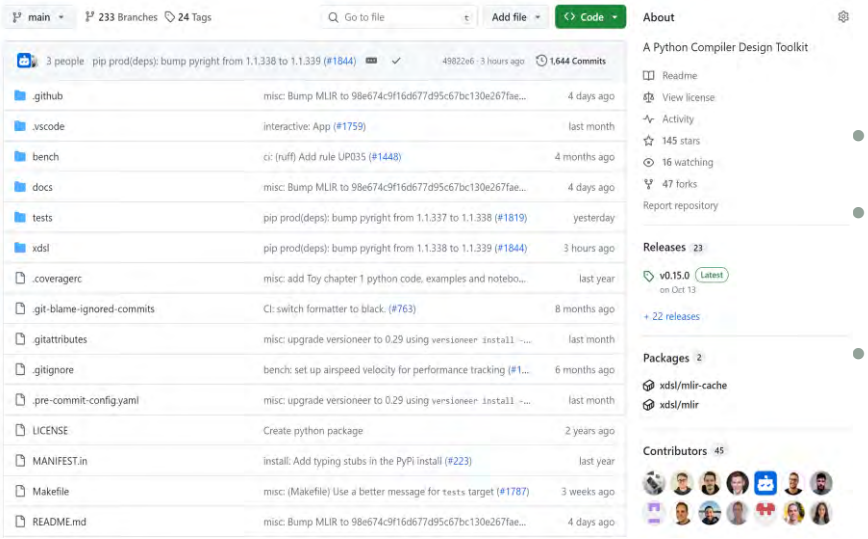
- Python toolkit for MLIR that enables high productivity development of dialects and transformations



- Contains existing MLIR dialects & transformations and we are adding HPC focussed ones too
- Whole load of other things also, such as an MLIR interpreter and Python frontend

ExCALIBUR

xDSL



https://github.com/xdslproject/xdsl

xDSL

- Makes experimenting with MLIR trivial
- Can go between xDSL and MLIR, leveraging transformations in both
- For our DSL purposes also means that a DSL can be a thin abstraction layer on-top of xDSL which provides a wealth of dialects and transformations that will ultimately drive MLIR/LLVM

intel

Google

ETH zürich

CINECA

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA A.D. 1088

Numba

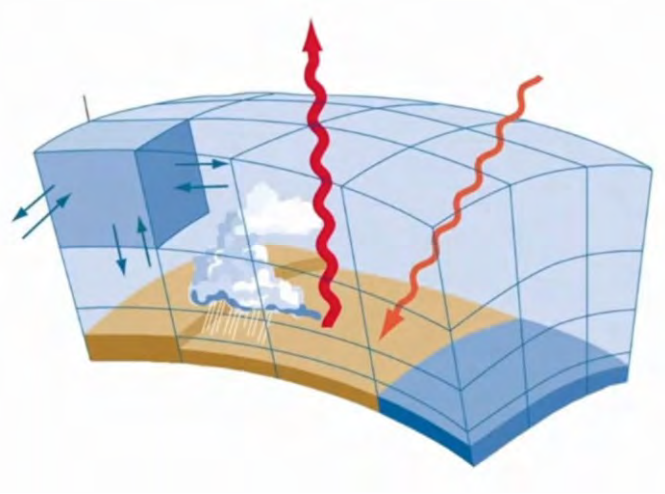
TECHNISCHE UNIVERSITÄT DRESDEN

CIUK Theme: Productive supercomputing

- Making HPC More Accessible
  1. For HPC developers as they can more easily leverage supercomputers by using Domain Specific Languages
  2. For DSL developers as these are now a thin abstraction layer atop a common, well supported, ecosystem

Domain Specific Compilation

- The Open Earth Compiler project from ETH Zurich used MLIR for domain specific compilation of stencil codes
- Successfully leveraged MLIR's qualities to leverage high-level information and reach high throughput on GPUs



ETH zürich

EXCALIBUR

First targets

PSyclone

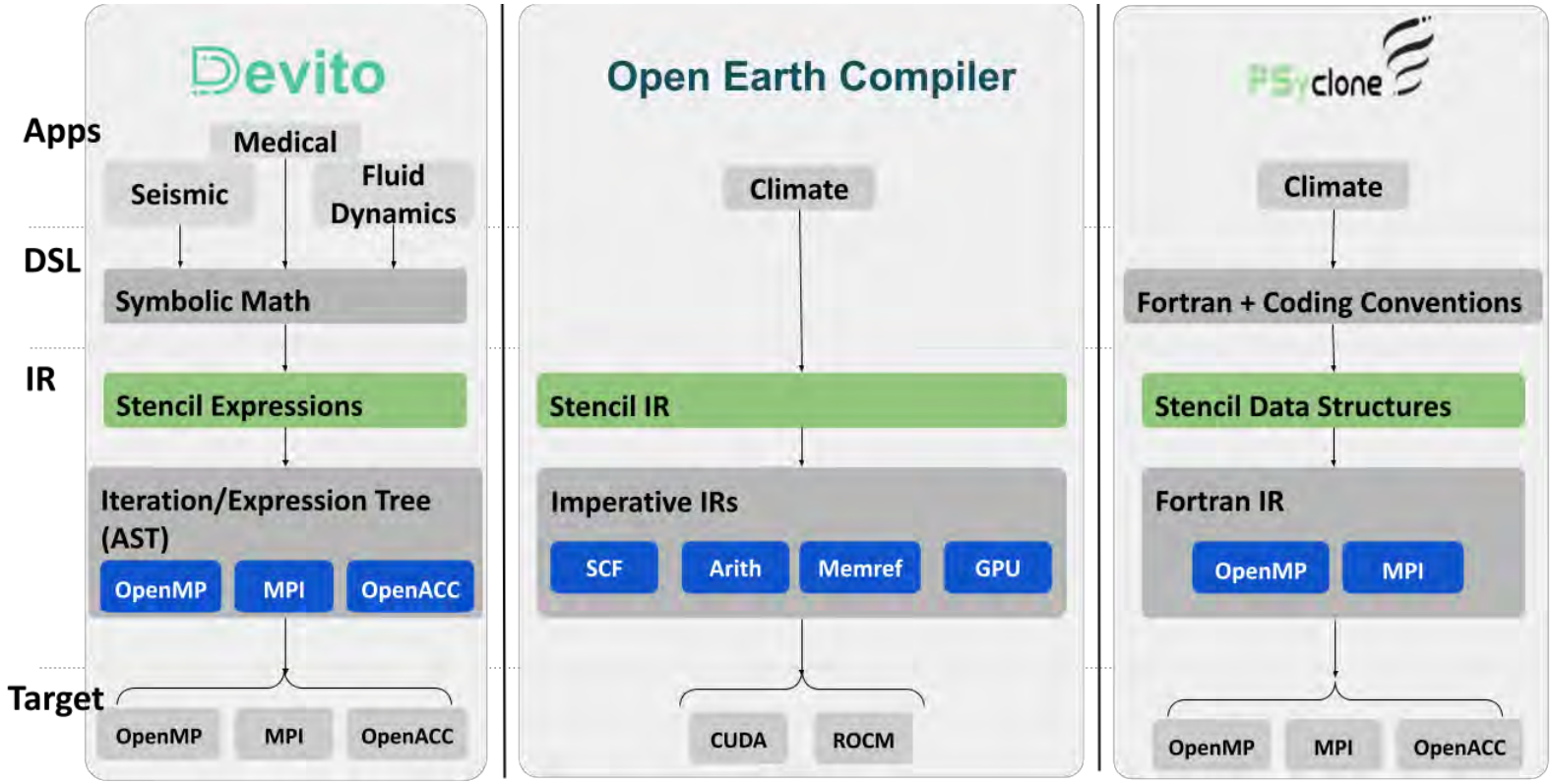
- Climate simulation
- Discovers stencil code in Fortran
- Apply Domain Specific optimizations
- Generates MPI, OpenMP, OpenACC code

Devito

- Seismic and fluid simulation, medical imagery
- Generates stencil code from Python PDEs
- Apply Domain Specific optimizations
- Generates MPI, OpenMP, OpenACC code

ExCALIBUR

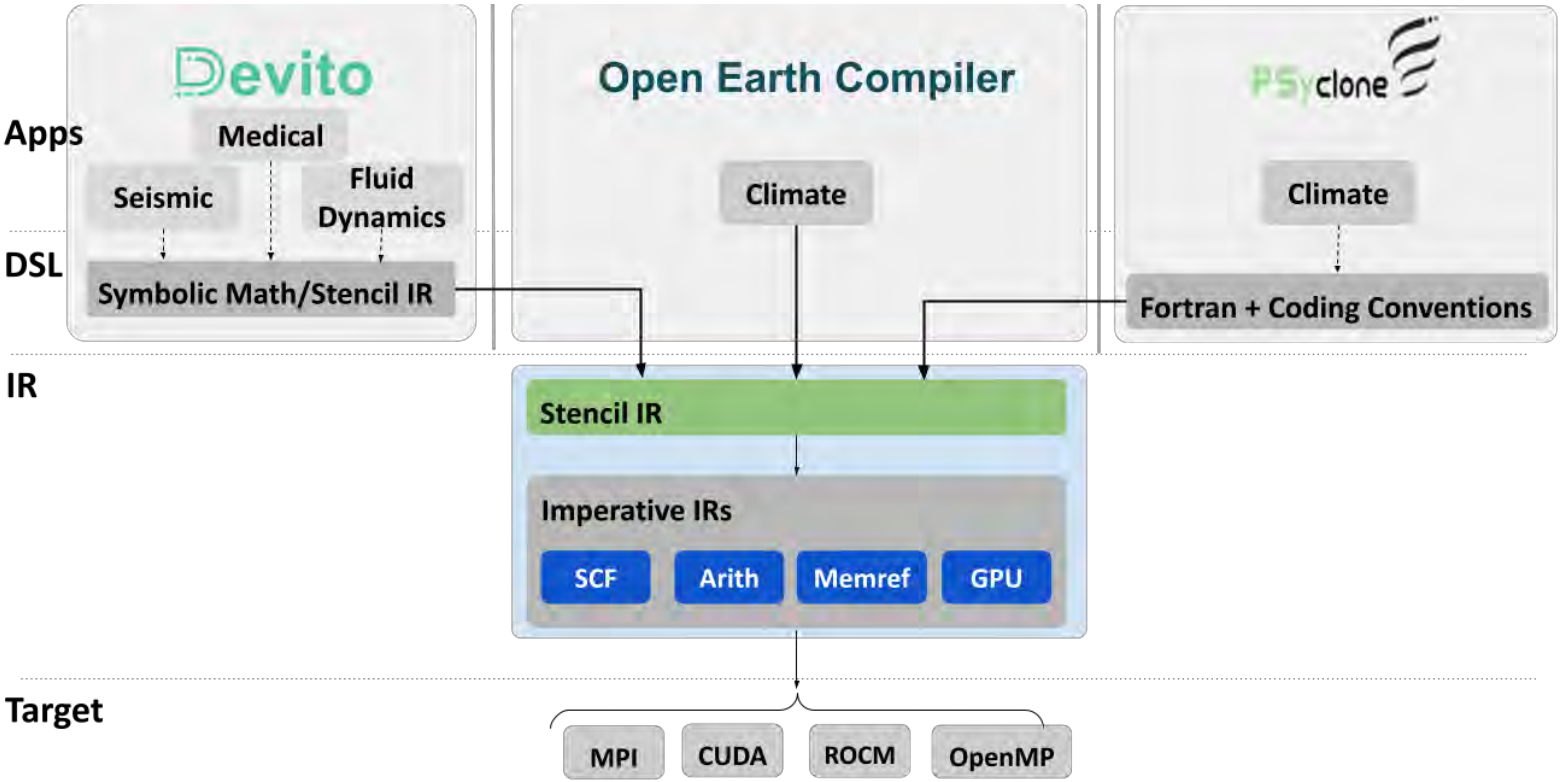
The broken silos



- Everything below the DSL layers is reinvented wheels

The sweet spot

xDSL

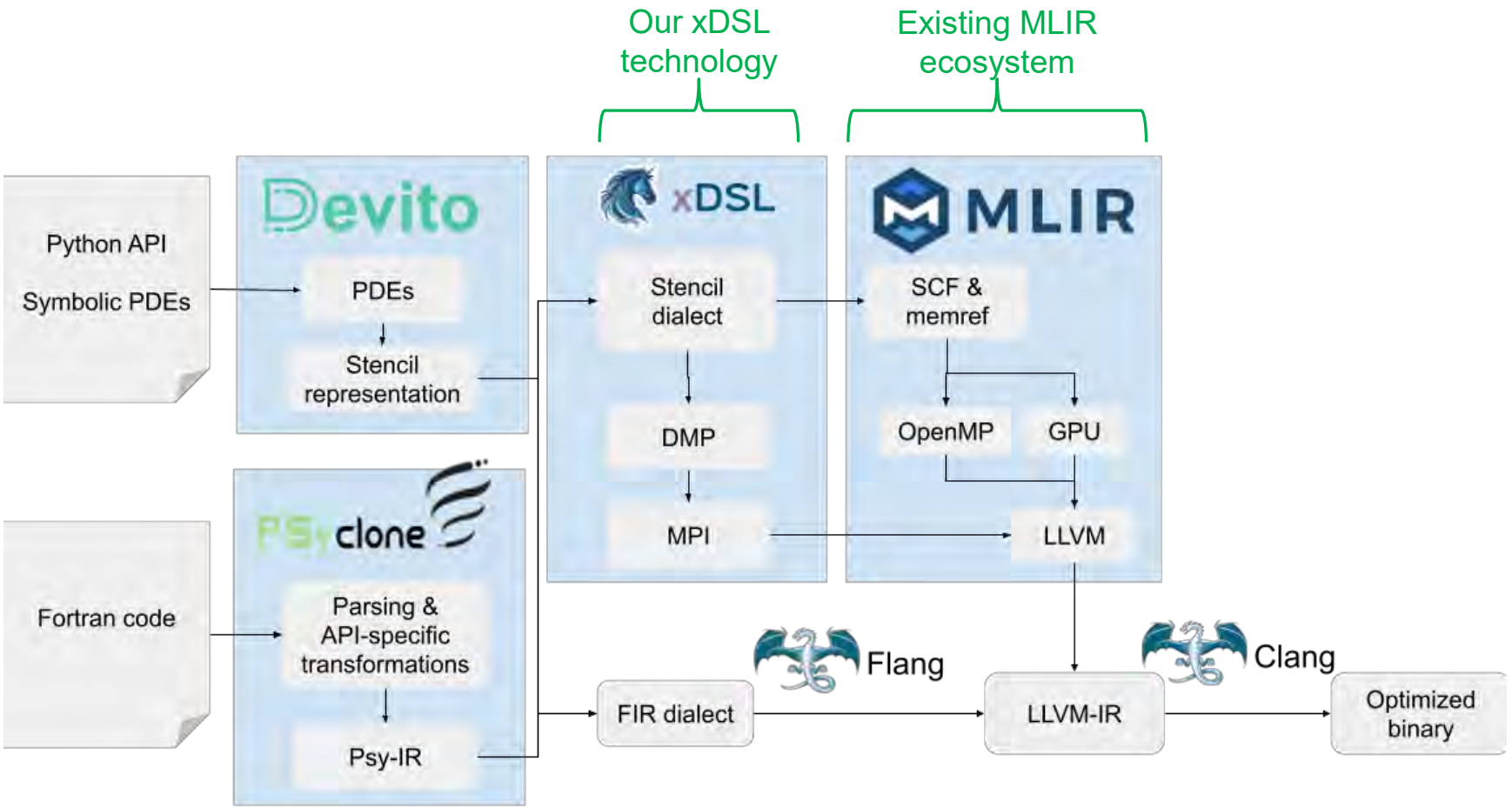


Sharing infrastructure:

- Implementation and Maintenance cost is spread across projects
- Everyone gets all benefits
- Can still be driven by specific needs

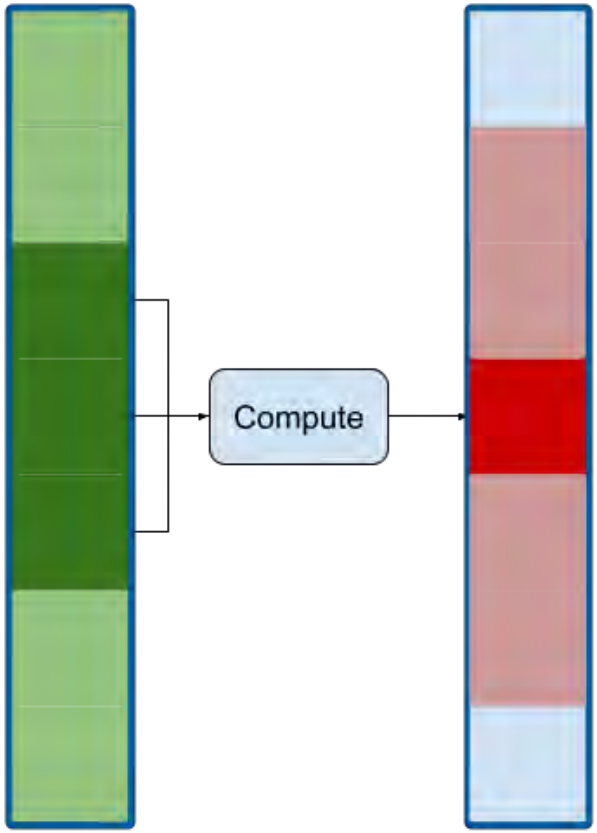
ExCALIBUR

The sweet spot

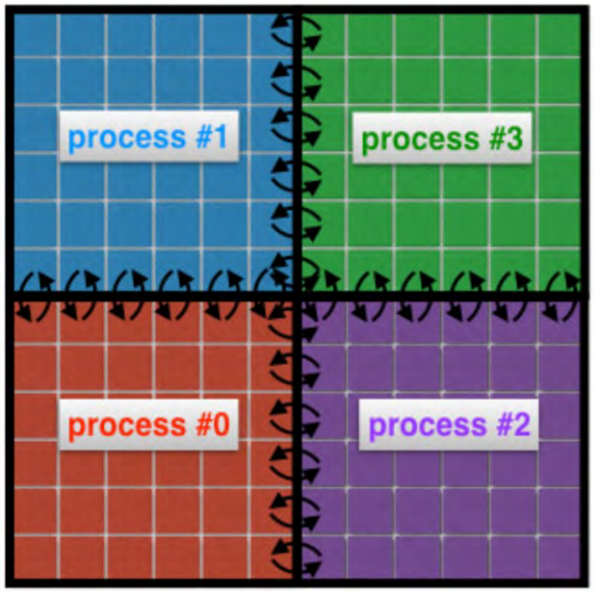


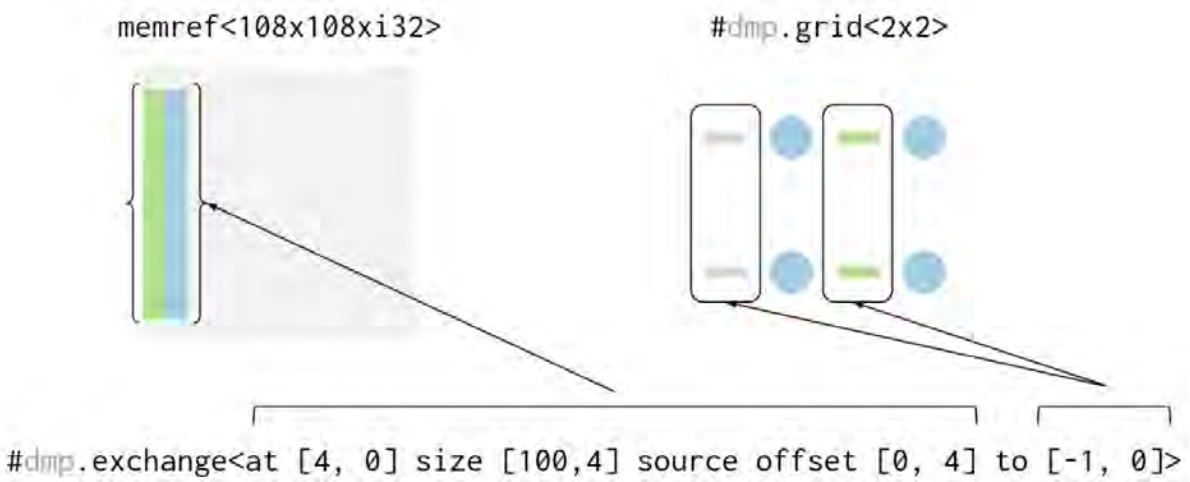
A flexible abstraction

```
  %input = stencil.load(%input_buffer) :
   (!field<[0,7]xf64>)-> !temp<?xf64>
  %out = stencil.apply(%arg = %input : !temp<?xf64>)-
> !temp<?xf64> {
    %l = stencil.access %arg[-1] : f64
    %c = stencil.access %arg[0] : f64
    %r = stencil.access %arg[1] : f64

    // %v = Some arbitrary computation

    stencil.return %v : f64
  }

  stencil.store %out to %target([1]:[6])
       : !temp<?xf64> to !field<[0,128]xf64>
```



High-level distribution



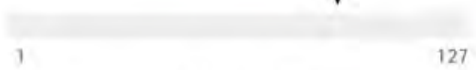


Halo exchange is a simple idea, let's *keep it* simple

High-level distribution

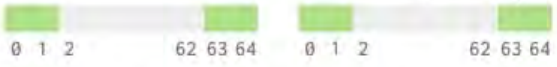
Stencil level IR → Global to Local → DMP level IR → DMP to MPI → MPI level IR

```
%source = stencil.load(%114) : (!field<[0,128]xf64>)
                             -> !temp<?xf64>
%out = stencil.apply(%arg = %source : !temp<?xf64>)
                  -> !temp<?xf64> {
  %l = stencil.access %arg[-1] : f64
  %c = stencil.access %arg[0] : f64
  %r = stencil.access %arg[1] : f64
  // %v = %l + %r - 2.0 * %c
  stencil.return %v : f64
}

stencil.store %out to %target([1]:[127])
```



Global Domain

```
%ref = builtin.unrealized_conversion_cast %114 :
    !field<[0,64]xf64> to memref<64xf62>
dmp.swap(%ref) {
  "grid" = #dmp.grid<2>,
  "swaps" = [
    #dmp.exchange<at [0] size [1]
                  source offset [1] to [-1]>,
    #dmp.exchange<at [64] size [1]
                  source offset [-1] to [1]>
  ]
} : (memref<64xf64>) -> ()
%source = stencil.load(%114) ...
%out = stencil.apply(%source) ...
stencil.store %out to %target([1]:[64])
```



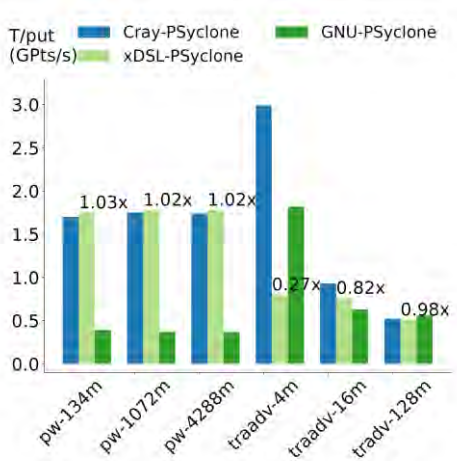
Local Domains with halo exchanges highlighted

```
%rank = mpi.comm_rank : i32
// First swap communication calls
%dest = arith.add %rank, %minus_one : i32
%is_in_bounds = arith.cmpi sge, %dest, %zero
scf.if %is_in_bounds {
  %view = memref.subview %ref[0][1][1] : memref<64xf64>
                                      to memref<1xf64>
  // copy data into send buffer and set up communication
  // (omitted for clarity)
  mpi.isend %sptr, %count, %dtype, %dest, %tag, %send_req
  mpi.irecv %rptr, %count, %dtype, %dest, %tag, %recv_req
}
// Second swap
// ...
mpi.waitall %requests, %four  // synchronization barrier
// First swap copy back
scf.if %is_in_bounds {
  %view = memref.subview %ref[1][1][1] : memref<64xf64>
                                      to memref<1xf64>
  memref.copy %recv_buffer_1 to %view
}
// Second swap copy back
// Lowered stencil comes here
```

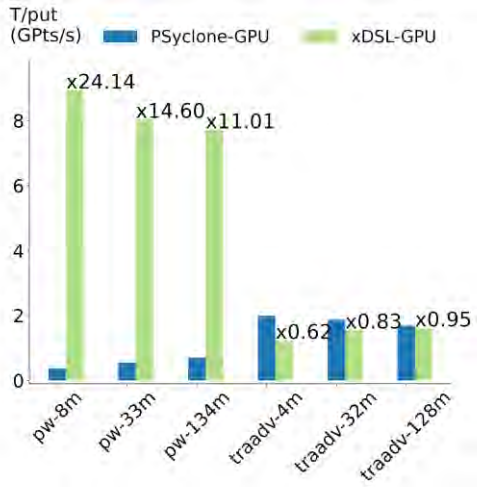
EXCALIBUR

Performance of PSyclone & Devito

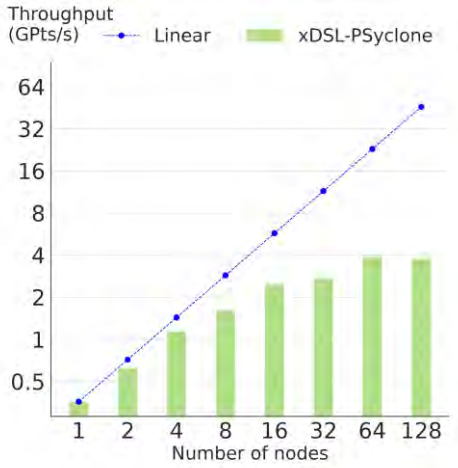
Single-node on ARCHER2

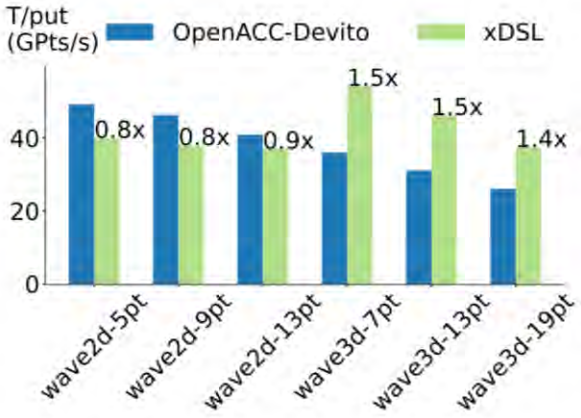


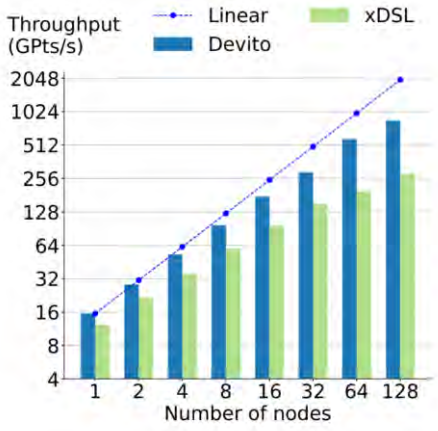
GPU on Cirrus (V100)



Strong scaling on ARCHER2



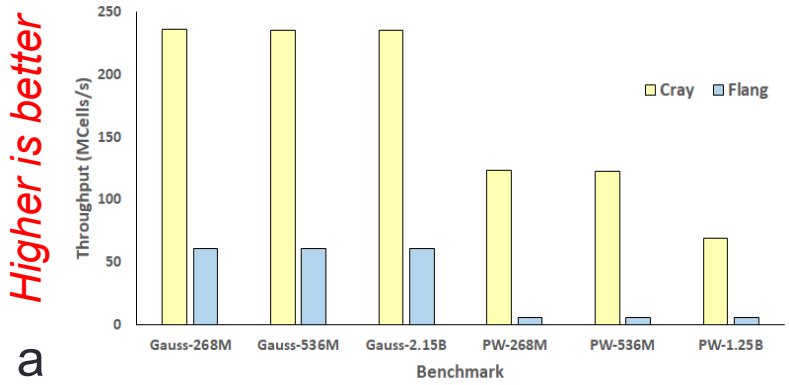


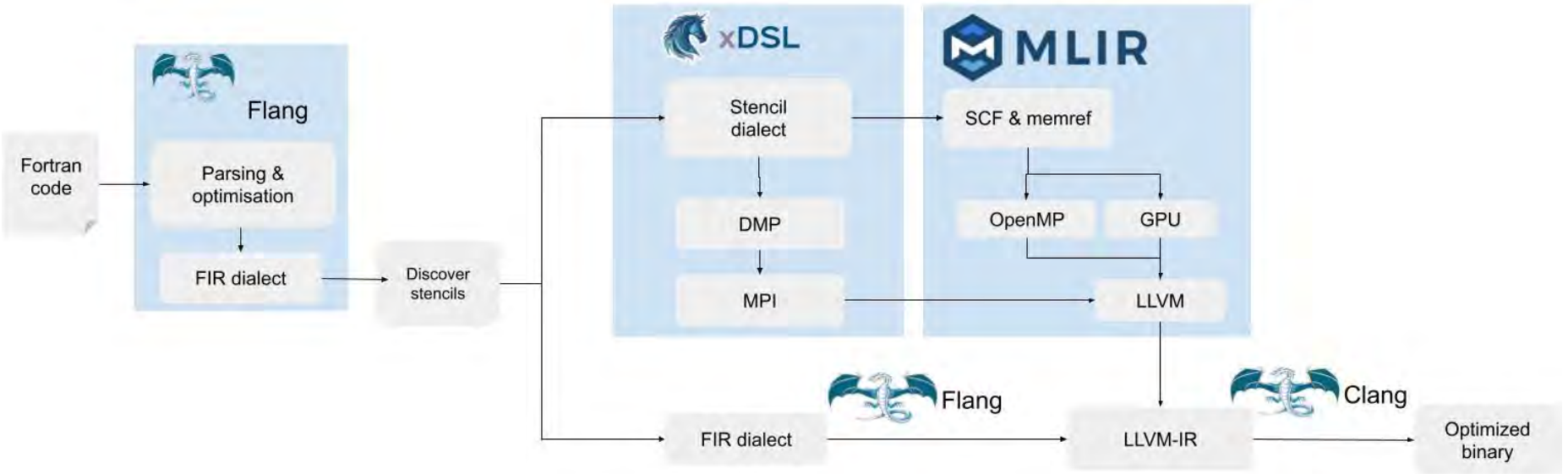


Higher is better, PSyclone top row & Devito bottom row

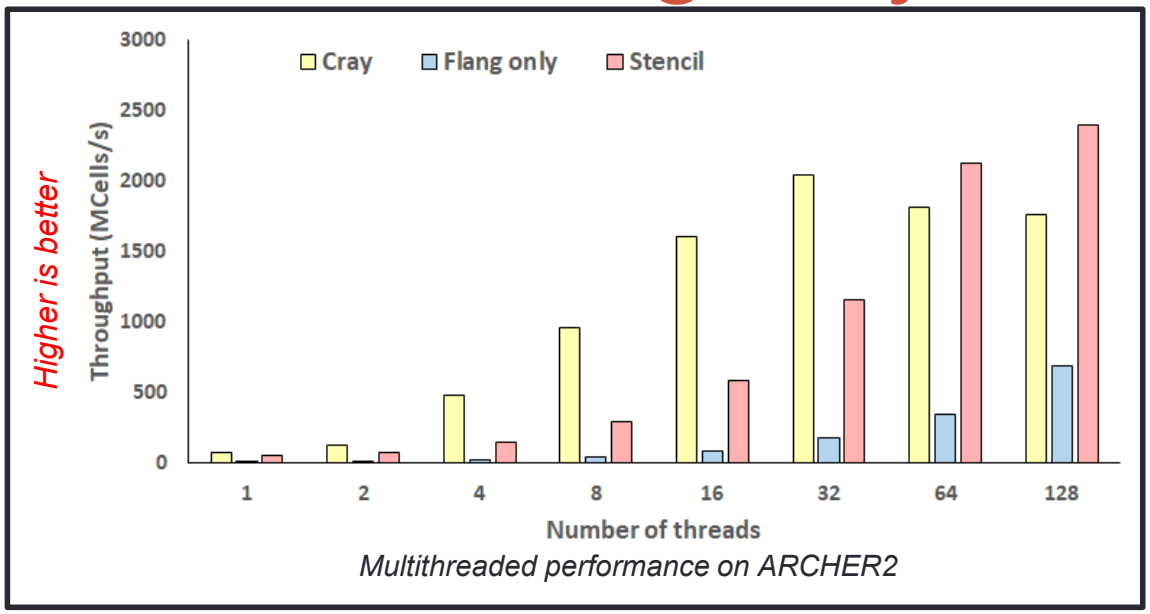
Integration with Flang: Beyond DSLs

- Performance falls short of Cray compiler for our stencil benchmarks (on a single core of ARCHER2, HPE Cray EX)
  - Our theory was that we can gain a performance improvement by combining with domain specific optimisations

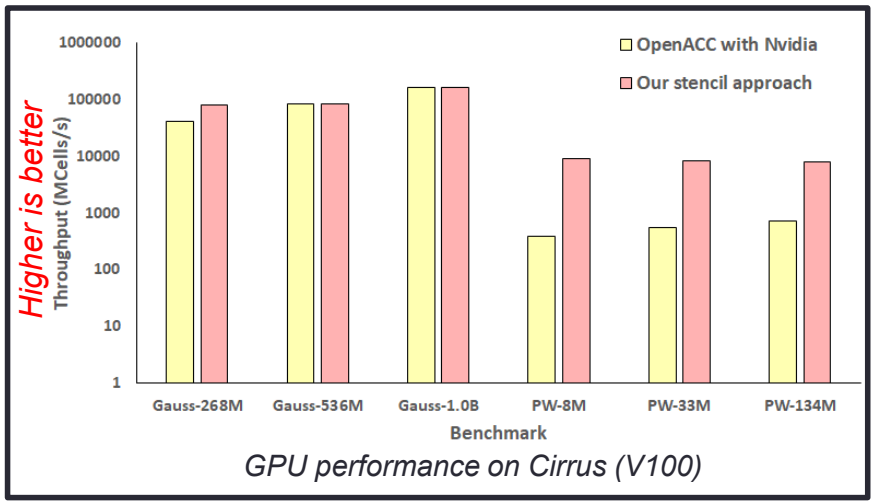




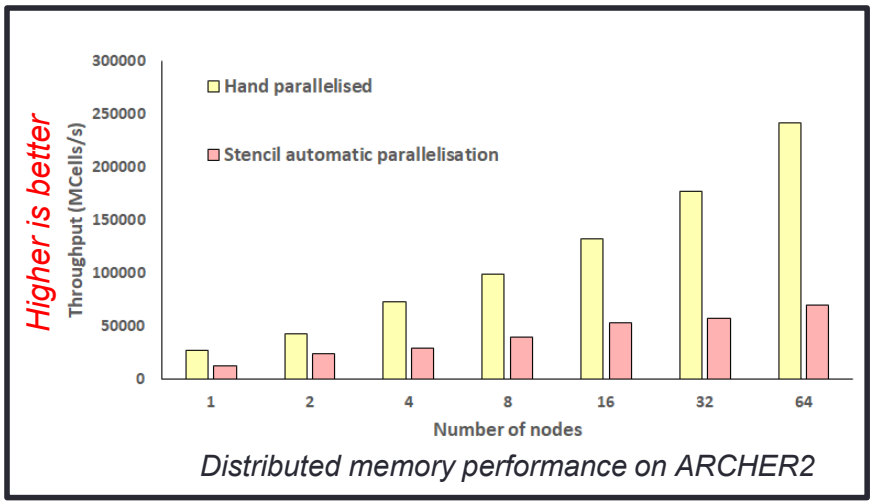
Integration with Flang: Beyond DSLs



Multithreaded performance on ARCHER2



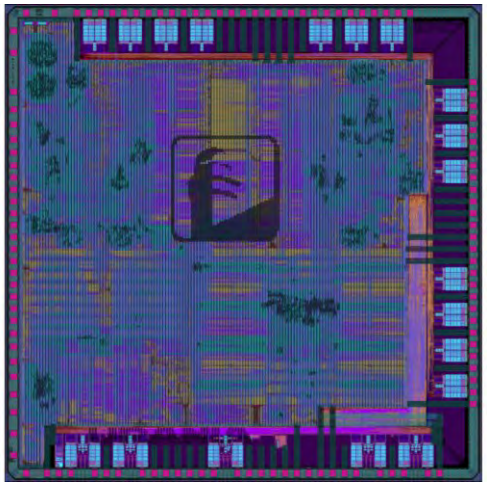
GPU performance on Cirrus (V100)



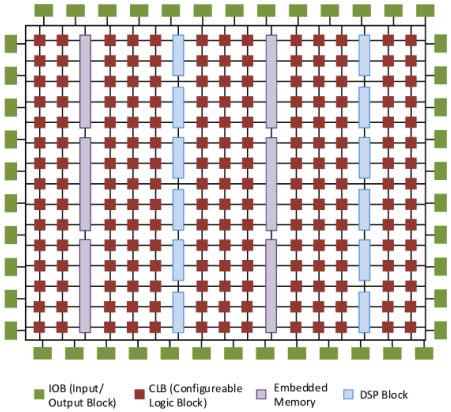
Distributed memory performance on ARCHER2

Auto-optimisation for new architectures

Field Programmable Gate Arrays (FPGAs)



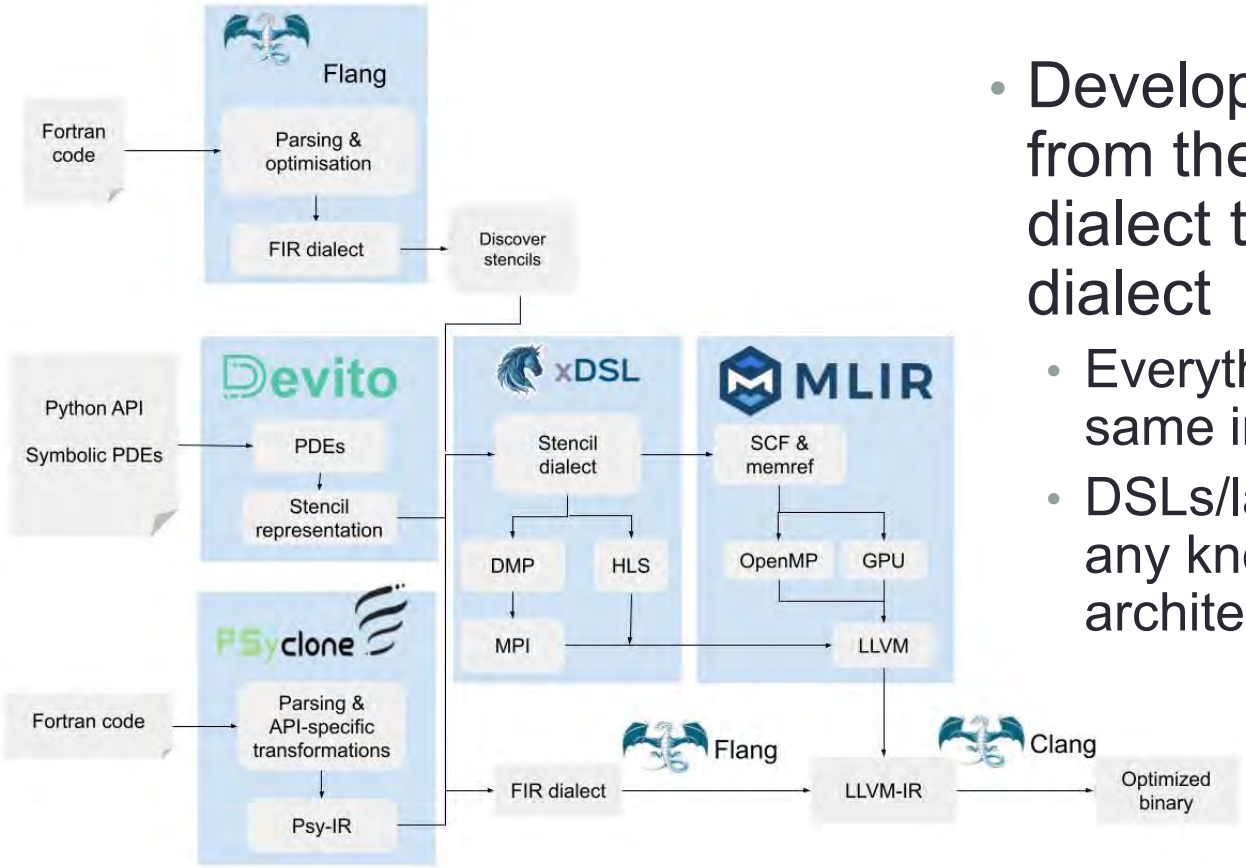
RISC-V high core-count accelerator chip



- Very different algorithm layout on FPGAs from the Von Neumann counterpart
  - Requires significant experience, expertise and time to port codes to the architecture
  - Using our existing infrastructure and domain specific abstractions, can we automatically optimise algorithms for FPGAs?
    - So there is a single, unchanged, Von Neumann version driving them?

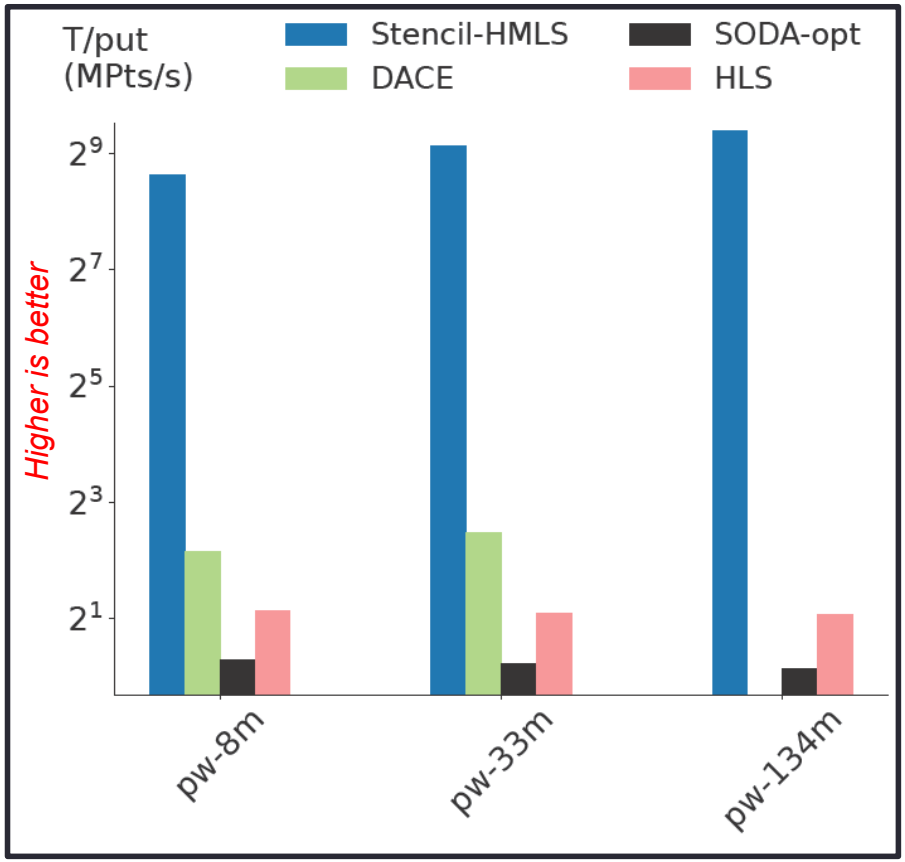
Automatic optimisation for FPGAs

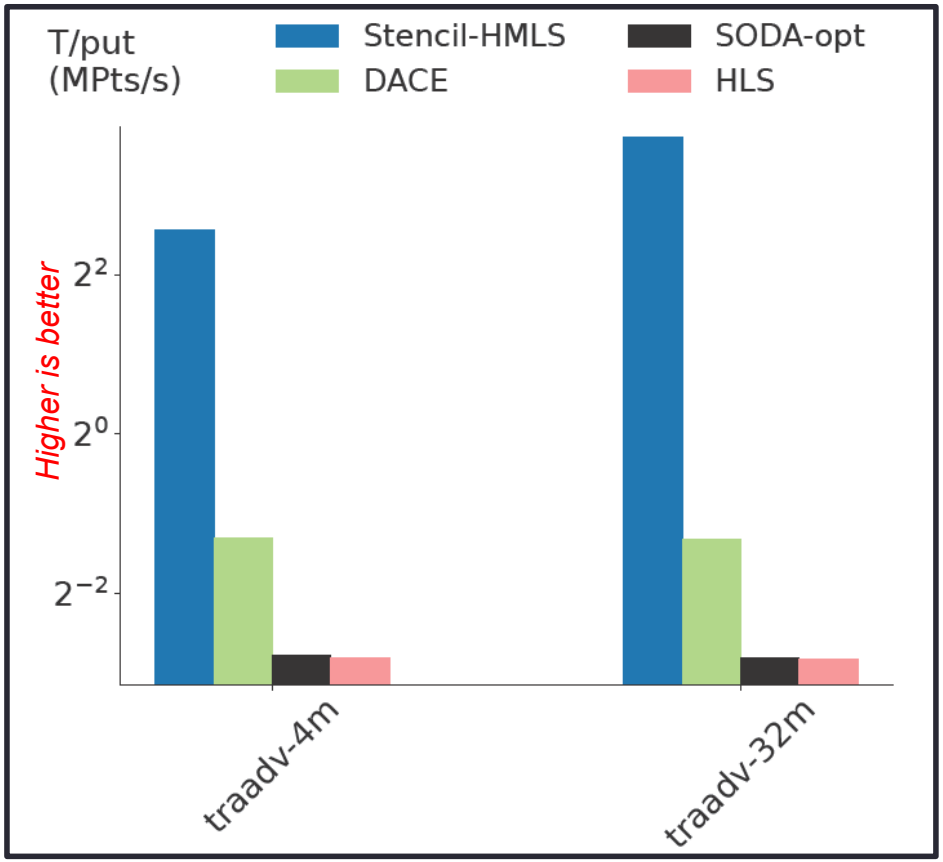
- AMD Xilinx already have an LLVM backend
  - We added a new High Level Synthesis (HLS) MLIR dialect that then lowers to IR compatible with AMD Xilinx's backend



- Developed transformations from the existing stencil dialect to this new HLS dialect
  - Everything else remains the same in the compiler pass
  - DSLs/languages don't need any knowledge of the target architecture

Automatic optimisation for FPGAs





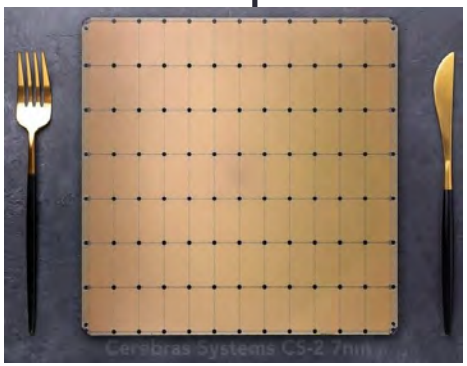
- On an AMD Xilinx U280 FPGA
- For PW advection, our approach is between 90 and 100 times faster than DaCE
- For tracer advection, our approach is between 14 and 21 times faster than DaCE

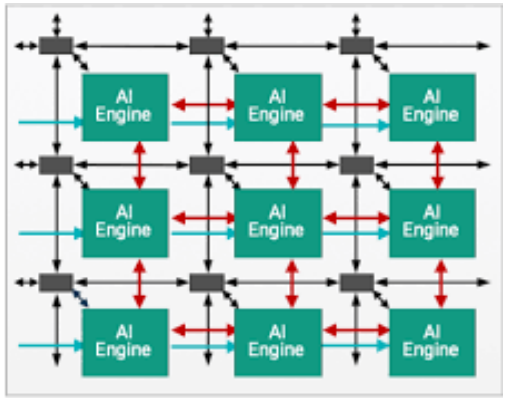
EXCALIBUR

Conclusions and next steps…

- We can't keep reinventing the wheel when it comes to compiler infrastructure for DSLs
- LLVM and MLIR are a strong alternative for sharing
  - We have developed the xDSL Python framework to lower the barrier to entry and offer key HPC components so that the ecosystems supports HPC workloads
- A lot of potential for bringing domain specific abstractions into existing languages, and we should be investing in Flang

- To date our focus has been on stencils, are now generalising this to other patterns





ExCALIBUR

Nick Brown

Emilien Bauer

Anton Lydike

- https://xdsl.dev
- https://github.com/xdslproject/xdsl
- https://xdsl.zulipchat.com/

The Jacky Pallas Memorial Presentation

Muting Hao (University of Oxford)

Advancing Aviation Efficiency and Sustainability Through CFD

Abstract: My research aims to enhance the sustainability of commercial aviation through turbomachinery Computational fluid dynamics (CFD). Commercial aviation accounts only for 2% of global emissions, but its visibility and projected growth demand action to make the sector sustainable in the future. Immediate gains in sustainability can be achieved through reductions in fuel use. My research helps reducing fuel use by improving the performance of aircraft engines. In broad terms, the performance of an aircraft jet engine can be improved in two ways: 1) increasing the fraction of the energy carried by the fuel that is transformed into the kinetic energy of the propulsive jet; 2) increasing the fraction of the kinetic energy of the propulsive jet used to propel the aircraft. The former can be attained by increasing the temperature at turbine inlet; the latter can be attained by producing most of the engine thrust using a large diameter fan. Both types of improvements pose new challenges to designers. My work on turbine cooling addresses the complexity of film cooling flows bring challenges to accurate numerical simulations and traditional Reynolds-averaged Navier–Stokes models struggle with accuracy. Large Eddy Simulation (LES) has been employed for accurate predictions of thermal performance and turbulence characteristics, enabling a parametric study, yielding insights into mixing regimes that are crucial for design considerations. Validated predictions contribute to a scaling law for the three-dimensional wall jets, guiding computational models' modifications for improved accuracy, potentially refining film cooling design methodologies. In the work on bypass design, the interactions between the fan and its exhaust duct have been studied. Proposing a flow model accommodating circumferential distortions and considering component interactions, the study emphasizes an efficient computational approach for optimal configurations, crucial for maintaining performance and integrity amid geometric variations.

Bio: Dr. Muting Hao is a Postdoctoral Researcher based in the Oxford Thermofluids Institute (OTI), at University of Oxford. Muting achieved her Dphil degree in Gas Turbine Aerodynamics from University of Oxford in 2022. Muting's Dphil was supported by Rolls-Royce funded scholarship. Her doctoral research focused on development of numerical tools and detailed turbulence statistics of film cooling, a technology of critical importance for gas turbines, under the supervision of Prof Luca di Mare. Muting then got supported by Rolls-Royce plc and by UKRI through the Horizon Guarantee Scheme for her postdoctoral research. She currently focuses on the design of outlet guide vanes for ultra-high-bypass ratio turbofan for aircraft engines.

UKRI The Jacky Pallas Memorial Award 2023

Advancing Aviation Efficiency and sustainability through CFD

Dr. Muting Hao

University of Oxford

2023.12.08

What is jet engine?

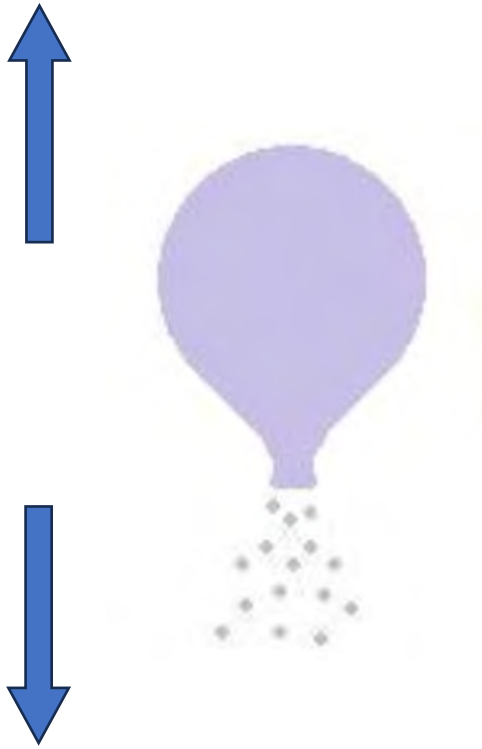
Dr. Muting Hao

Pictures © 2023 Rolls-Royce

Newton's 3rd law

Reaction:

Balloon goes up



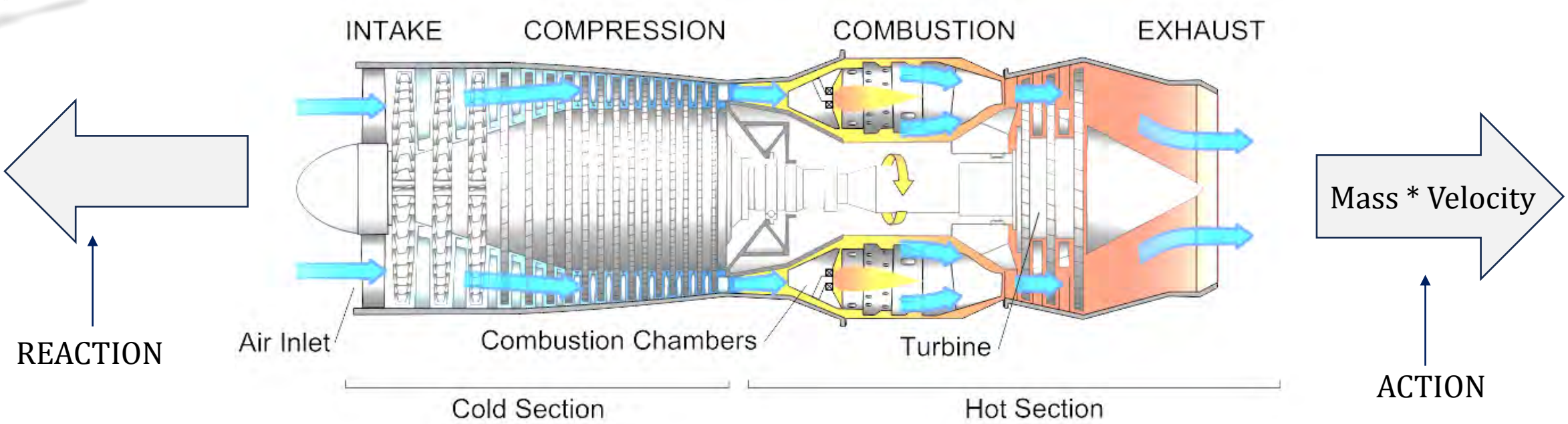
Action:

Air rushes down

Thrust= mass*($V_{aircraft} - V_{jet}$)

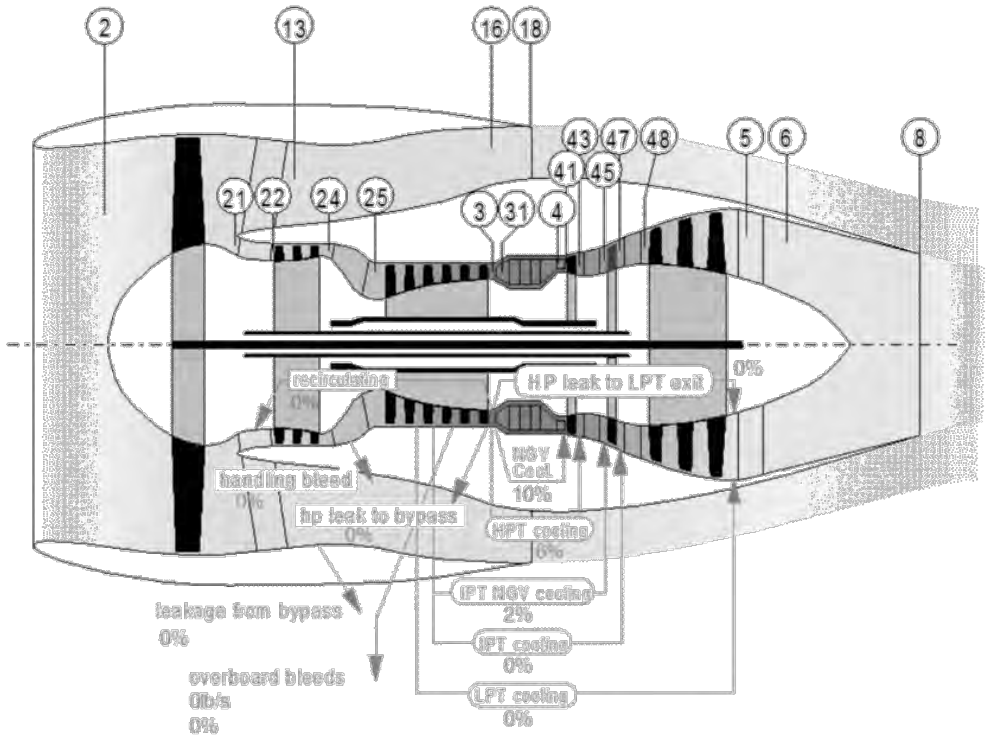
Jet – *moves small mass of gas at high velocity*

How does Jet Propulsion work?

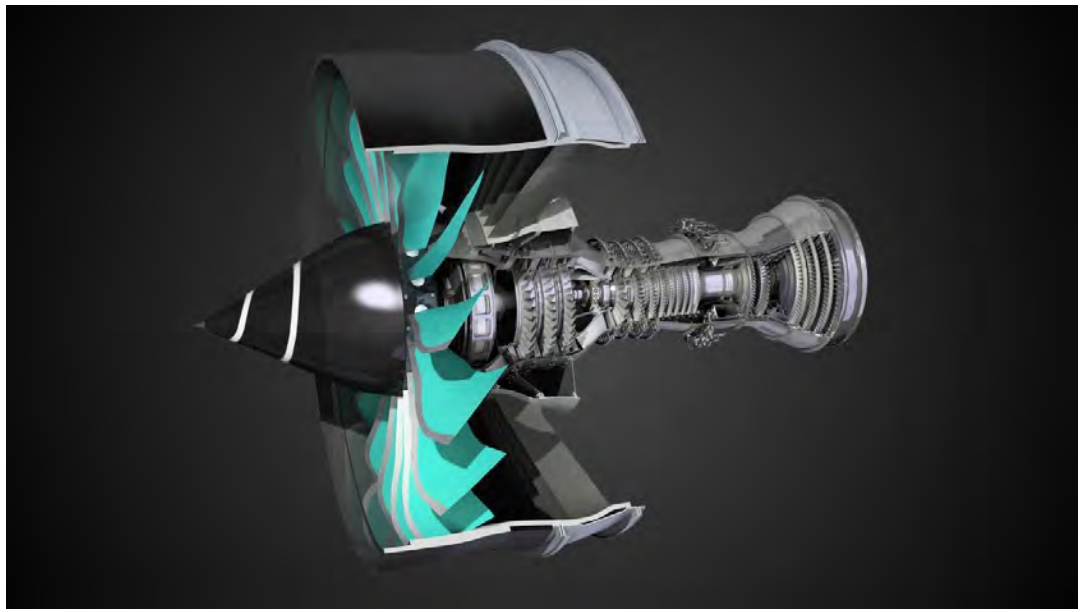


Trust= Mass*Velocity

Turbofan, Ultrafan



Trent XWB



Ultrafan

25 per cent more fuel efficient than the original Trent 700

40% less NOx

UNIVERSITY OF OXFORD

Why need the sustainability of commercial aviation?

Sole option for long-haul travel

Global emission

Economic Viability

Reductions in fuel burn per passenger per mile *are vital to decarbonise commercial flight!*

Immediate gains in sustainability - reductions in fuel use

Reductions in fuel burn per passenger per mile

The energy carried by the fuel

1) Increase

Method 1

The kinetic energy of the propulsive jet

2) Increase

Method 2

the fraction used to propel the aircraft

Increasing core temperature

Challenges

Increasing bypass ratio

Challenges

Both types of improvements pose new challenges to designers…

Improve sustainability – my research on reductions in fuel use

Method 1 **Increasing core temperature**

Solution

turbine cooling

Method 2 **Increasing bypass ratio**

Solution

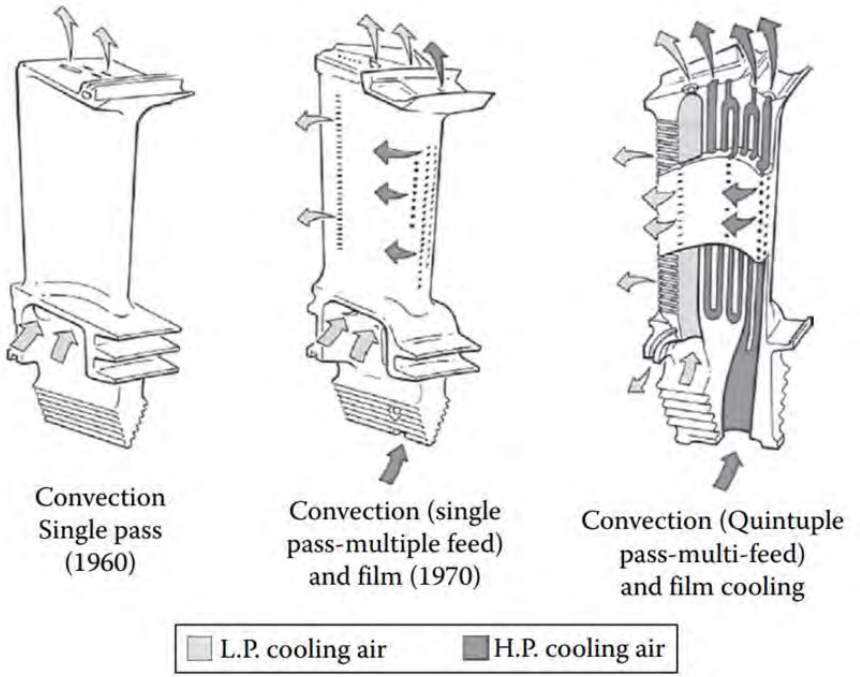
turbofan OGV design

My research addresses the most critical challenges encountered precisely in these two aspects of gas turbine performance improvement!

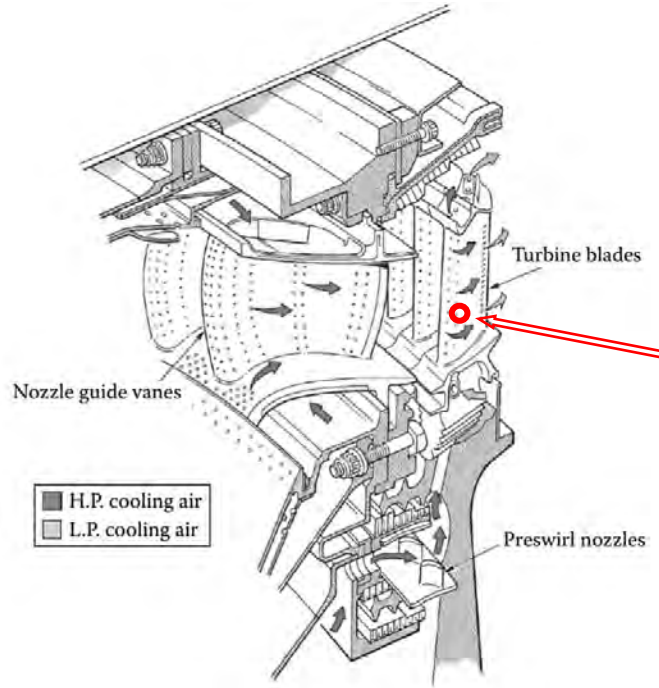
My research Part I

Turbine Film Cooling

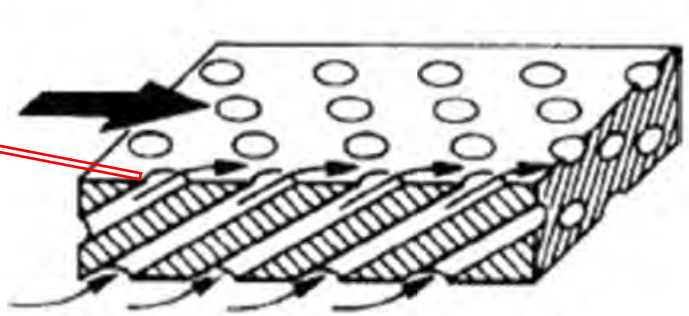
What is film cooling in gas turbine?



General development of cooling technologies, from Rolls-Royce plc, Copyright Rolls-Royce plc 2007



Cooling for a high-pressure turbine, from Rolls-Royce plc, Copyright Rolls-Royce plc 2007



Film Cooling principle

My research and code development

- ✓ Developed a CFD package for numerical modelling & computation in the LES framework
  - Resolution for small-scale features of flow

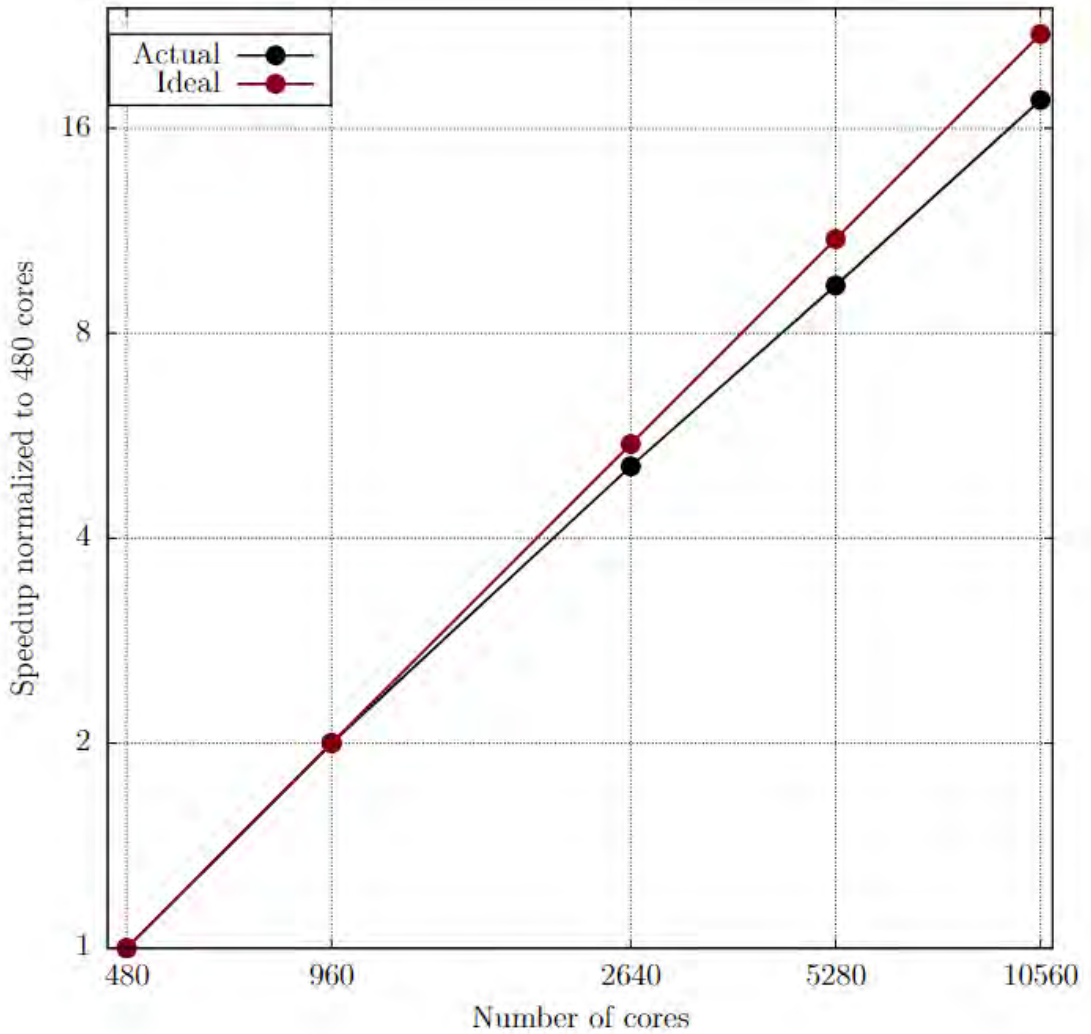
- ✓ Exploitation of all the available elements of numerical modelling and computation methods:

In-house code: HPC

- C++ ‘99 standard

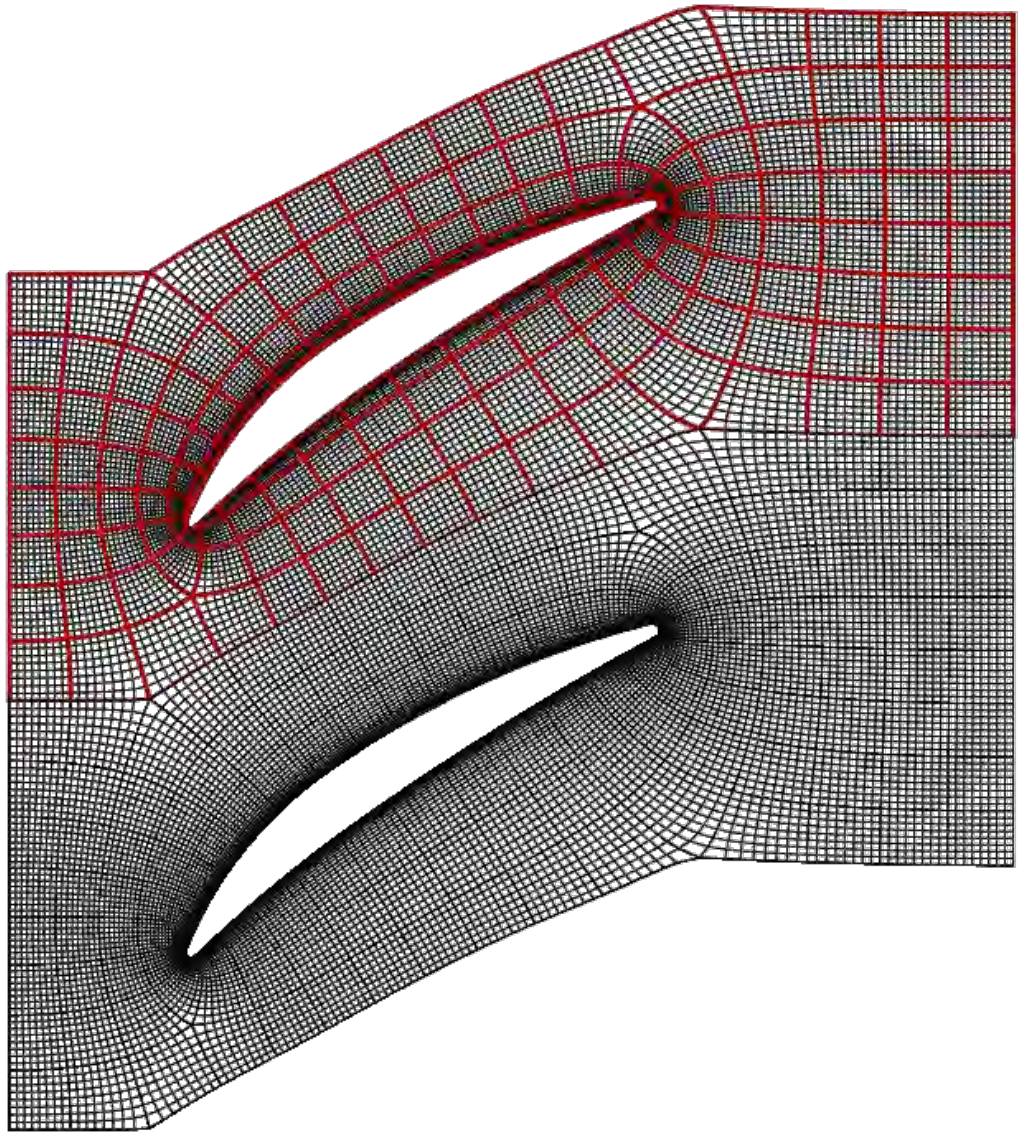
- 3 levels of parallelism
- Register level parallelism (SIMD/OpenMP #pragma SIMD)
- Thread level parallelism (OpenMP)
- Core level parallelism (MPI)

HPC implementation for the code: Speed



Mesh generator – Multiblock structured

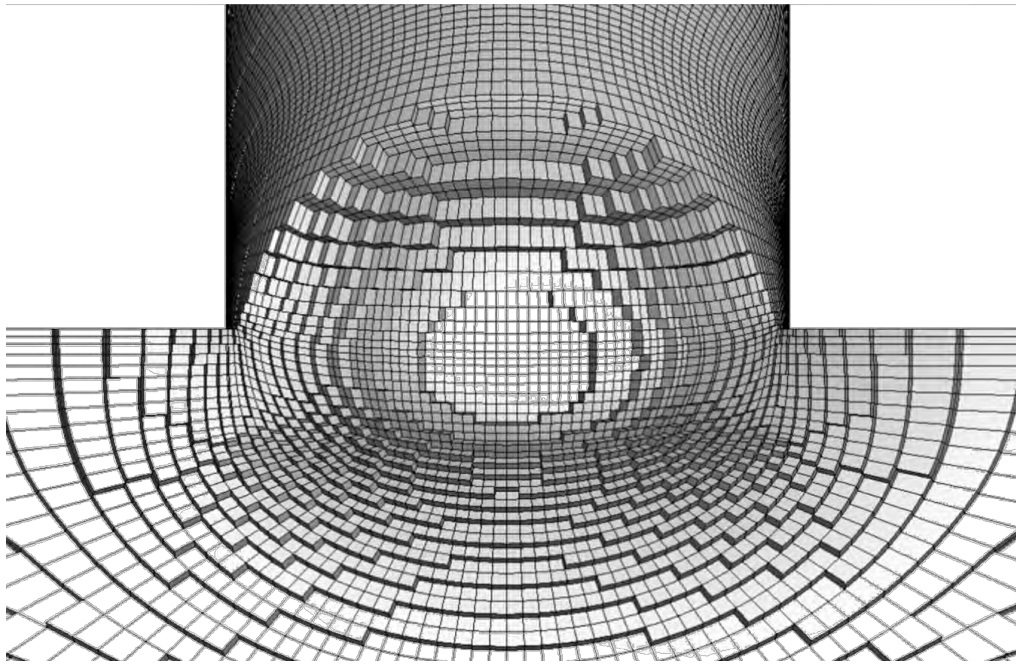
Turbine blade

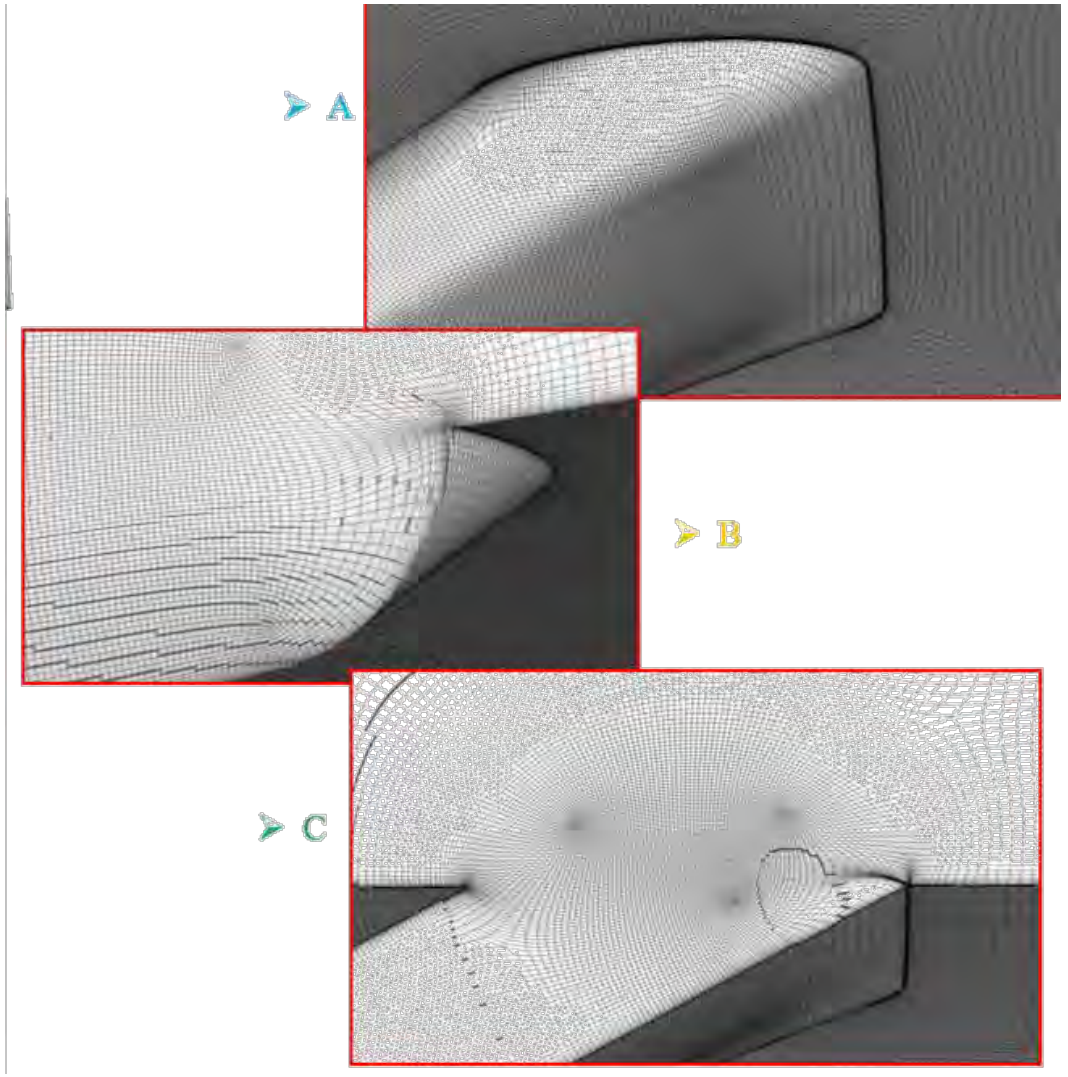


Mesh generator – Multiblock structured

Blade cooling hole

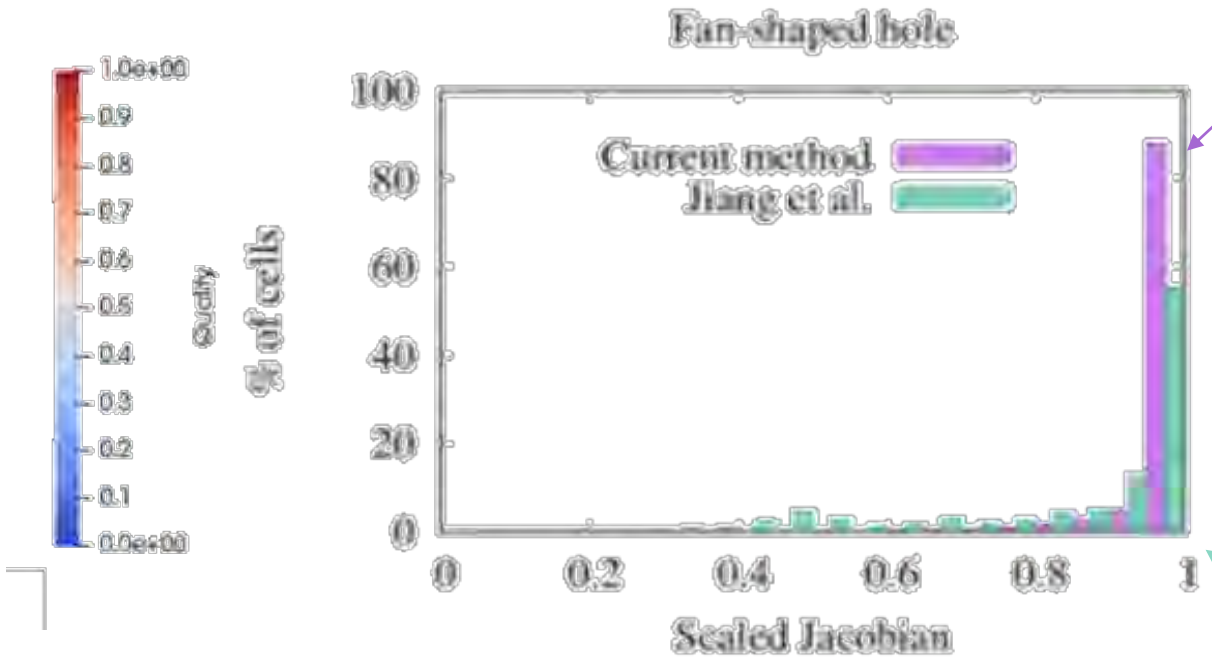
- **the block layout (unstructured)**
- **the block grids (structured)**



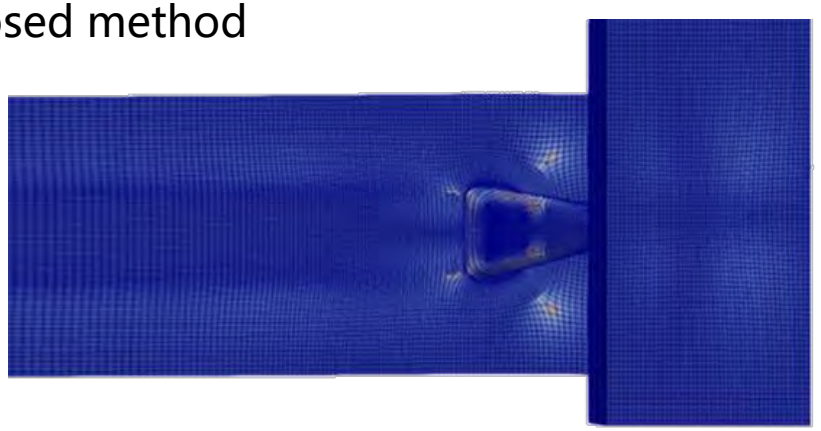


Mesh generator – Multiblock structured

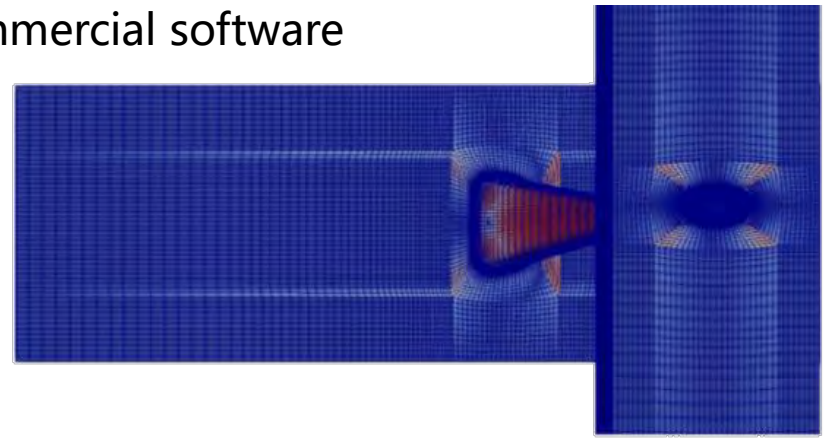
Mesh generation for hole: A truck-and branch



- Proposed method



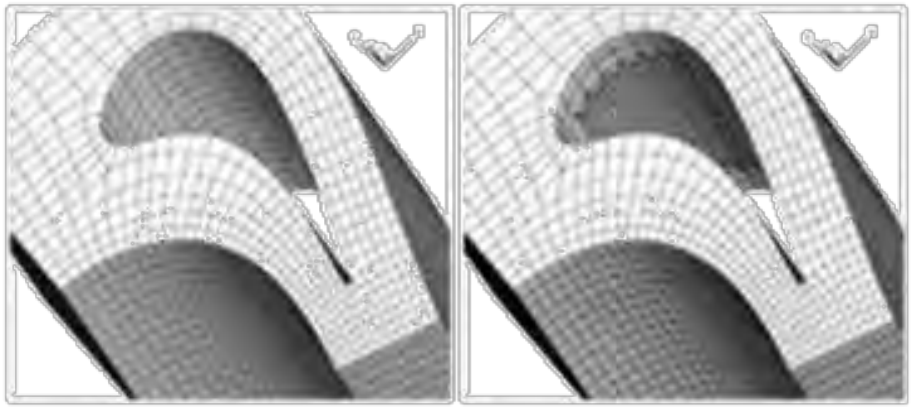
- Commercial software



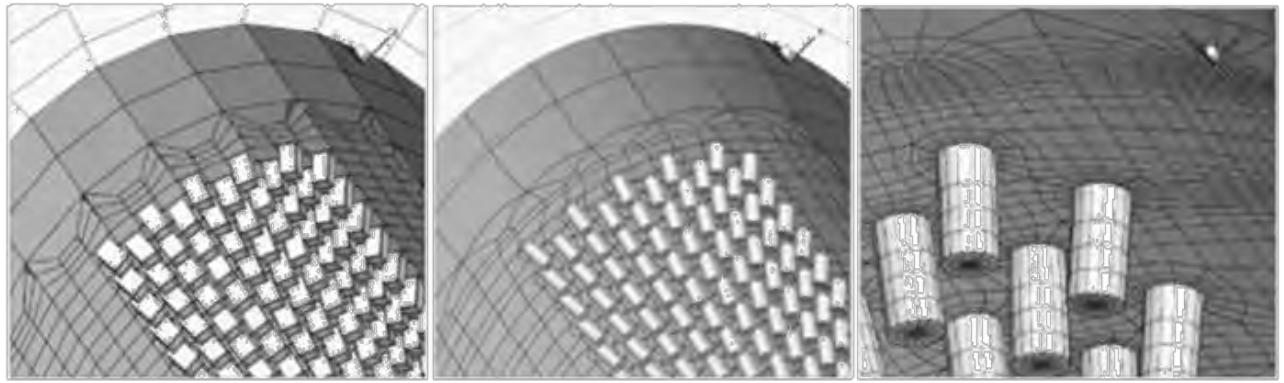
Mesh generator – Multiblock structured

Meshing procedure for turbine with holes

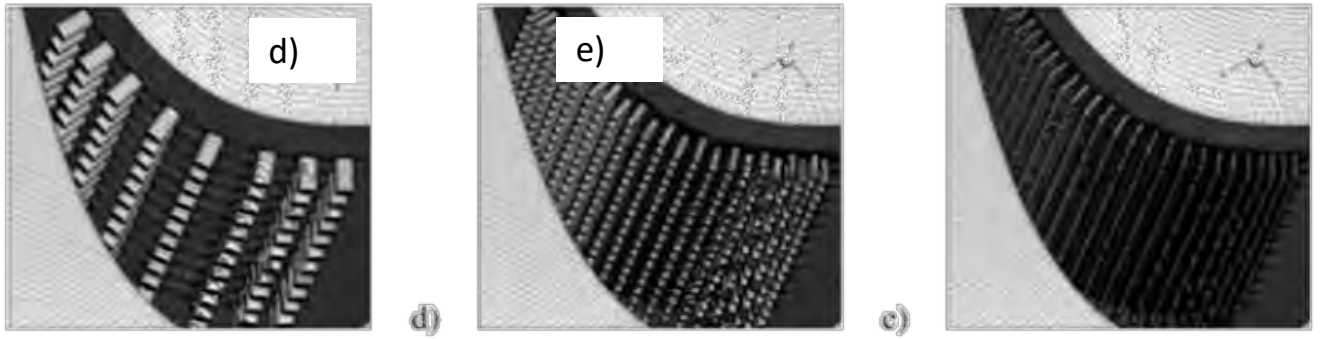
a)



b)



c)

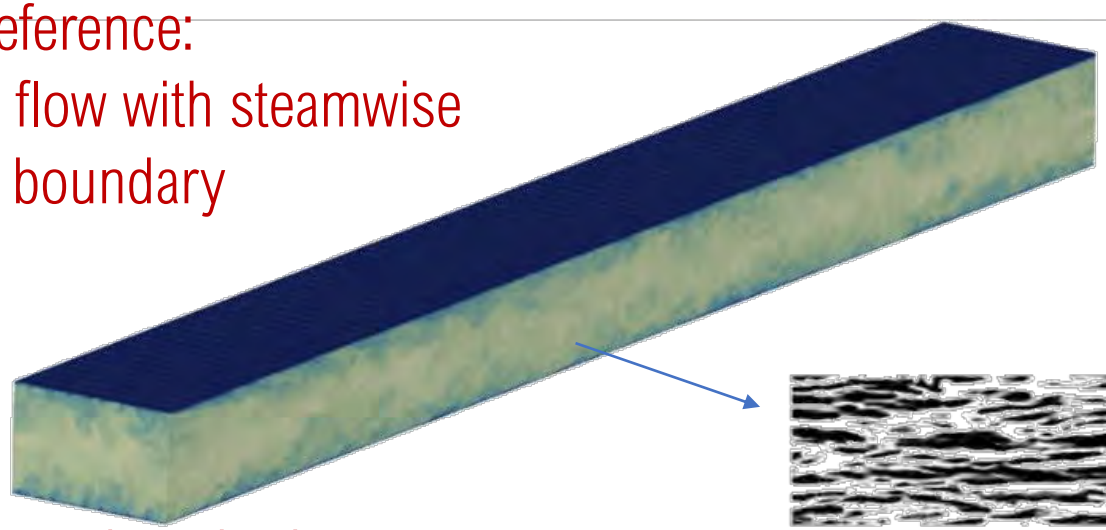


A novel synthetic turbulence Inflow generator

Hao, Muting, Joshua Hope-Collins, and Luca di Mare. "Generation of turbulent inflow data from realistic approximations of the covariance tensor." *Physics of Fluids* 34.11 (2022).

- For reference:

Channel flow with steamwise periodic boundary



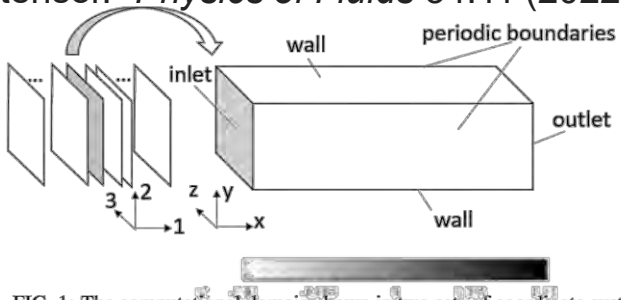
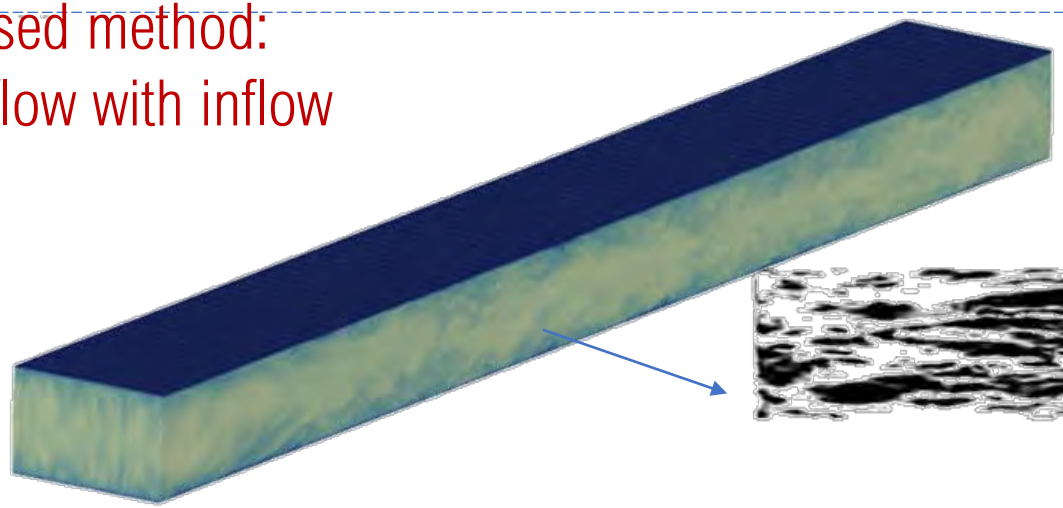


FIG. 1: The computational domain shown in two sets of coordinate systems

- Proposed method:

Channel flow with inflow generator



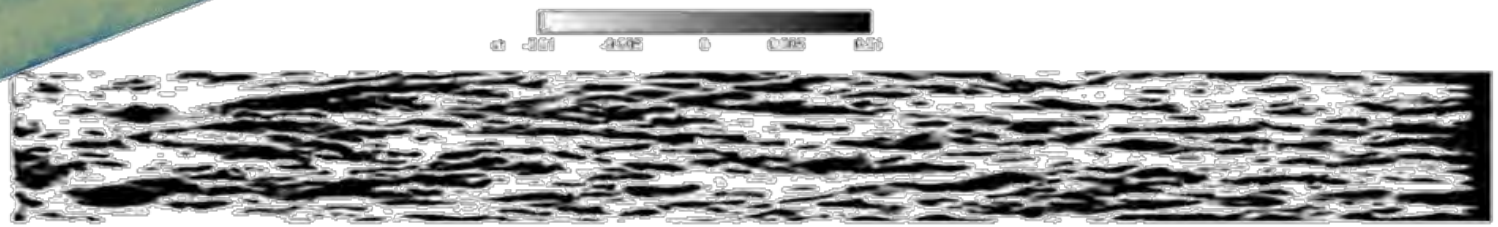


Figure 14: Downstream development of the velocity profiles downstream the inlet on the bottom half of channel

LES result of fan-shaped and cylindrical cooling films

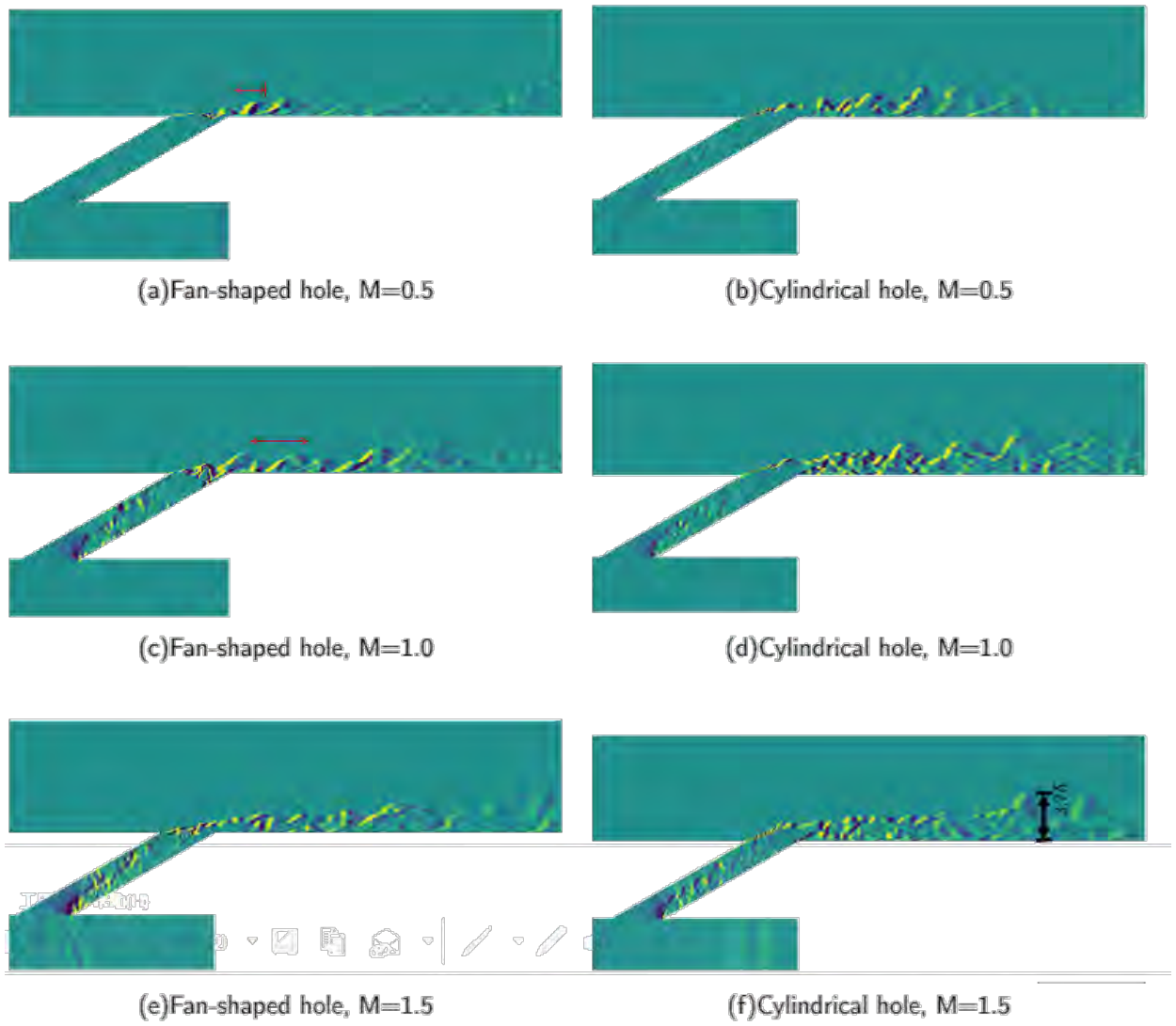


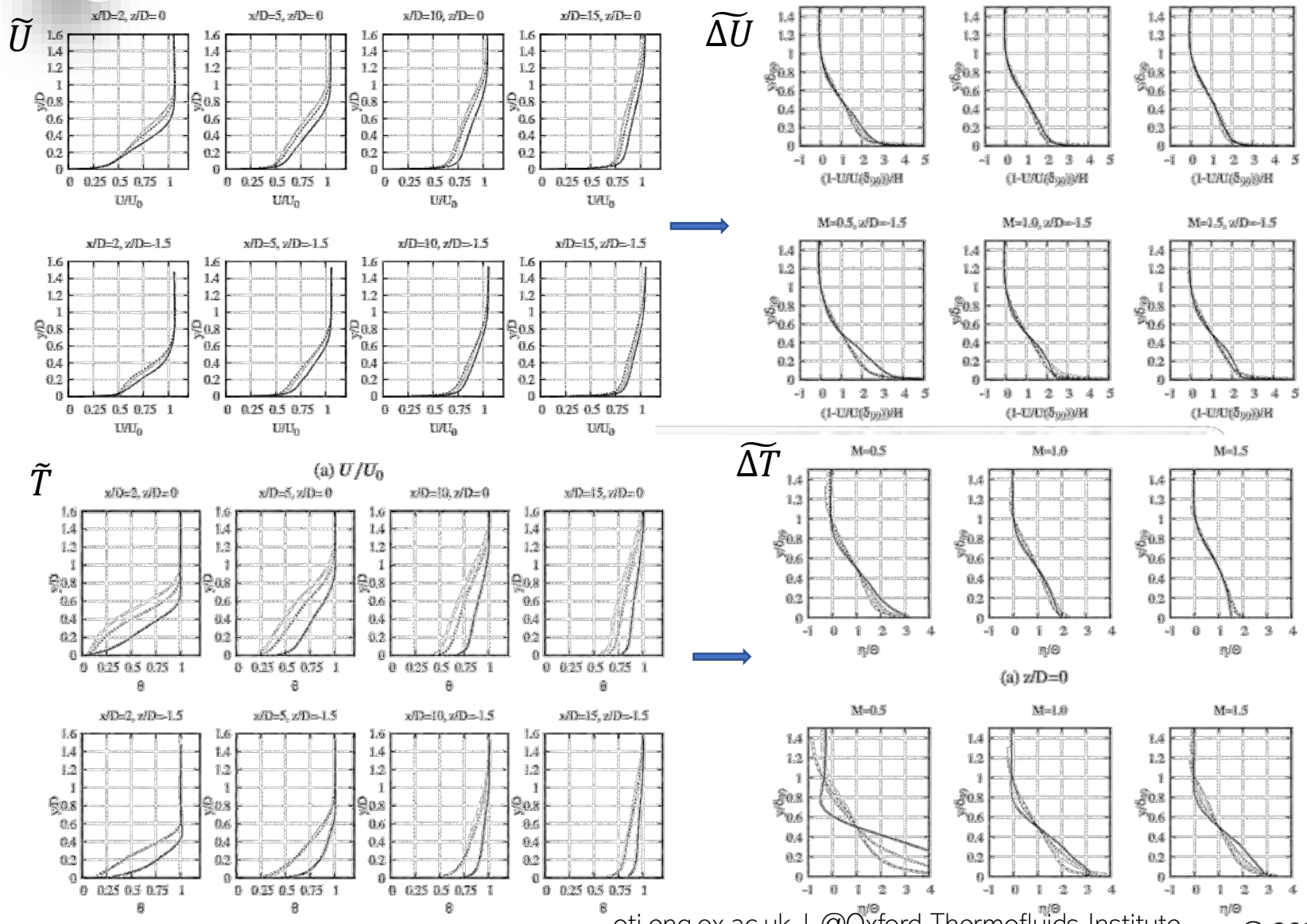
Figure 20: Instantaneous snapshot of the density gradient.

Hao, Muting, and Luca di Mare. "Reynolds stresses and turbulent heat fluxes in fan-shaped and cylindrical film cooling holes." *International Journal of Heat and Mass Transfer* 214 (2023): 124324.

LES - computations requires HPC

- All these results are based on millions of timesteps
- For 1 film cooling hole, the mesh is 10.5 million grids (large cost of computing time)
- 1 operating condition case:
  - requires 13 flow though of a cooling hole by Large Eddy Simulation
  - takes 132440 core hours
  - 448 cores are used for each run
- Cases for a wide range of operating conditions have been simulated
- The computations definitely need HPC!!!

LES results - A parametric study of film cooling hole flows

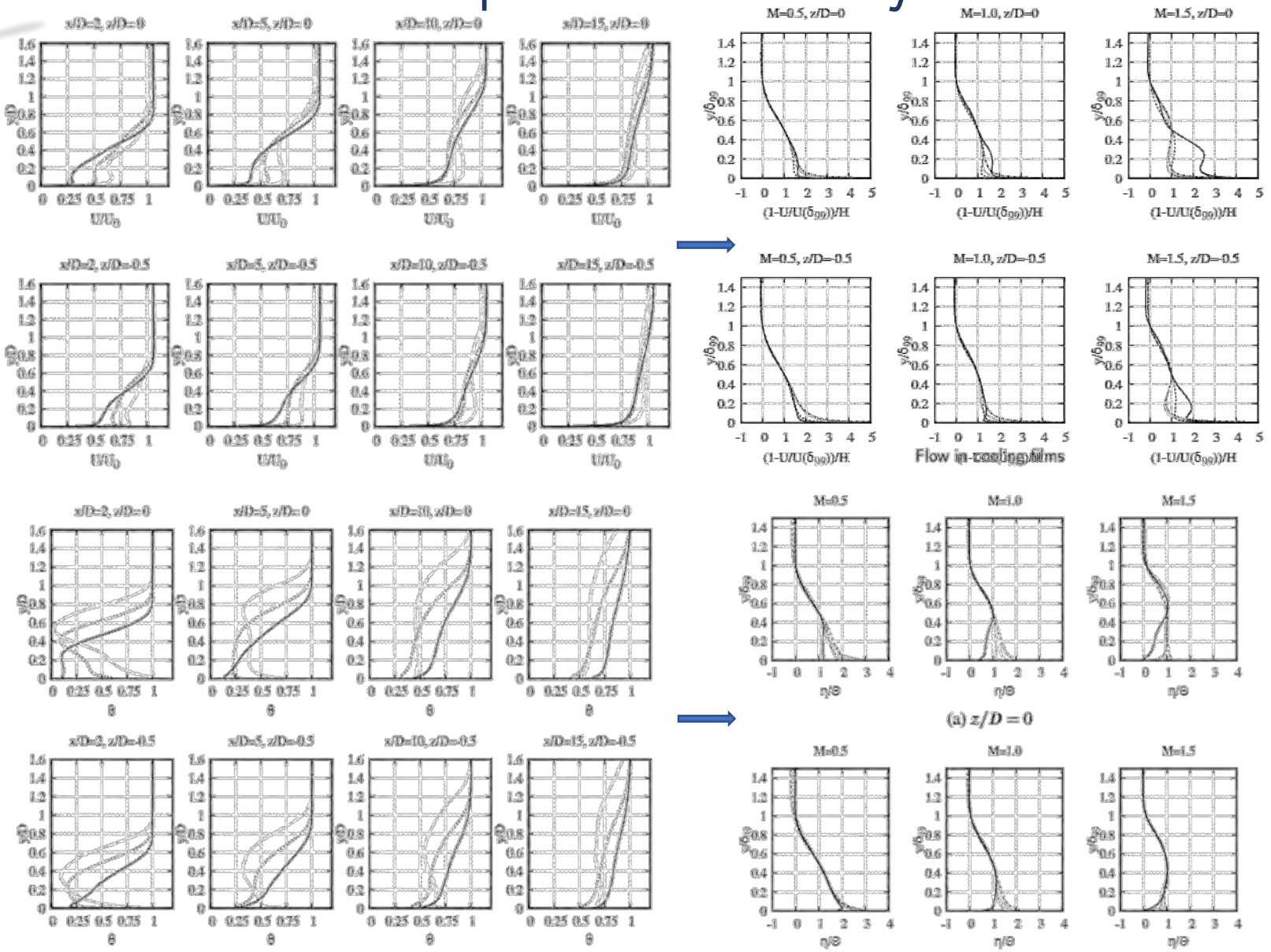


- fan-shaped film:

tend towards a self-similar flow profile

Hao, Muting, and Luca di Mare. "Reynolds stresses and turbulent heat fluxes in fan-shaped and cylindrical film cooling holes." *International Journal of Heat and Mass Transfer* 214 (2023): 124324.

LES results - A parametric study of film cooling hole flows



- cylindrical film:

only approach self-similar profiles at low blowing ratios or at high blowing ratios

Hao, Muting, and Luca di Mare. "Reynolds stresses and turbulent heat fluxes in fan-shaped and cylindrical film cooling holes." *International Journal of Heat and Mass Transfer* 214 (2023): 124324.

Theory proposed--

Scaling & similarity laws in 3D wall jets

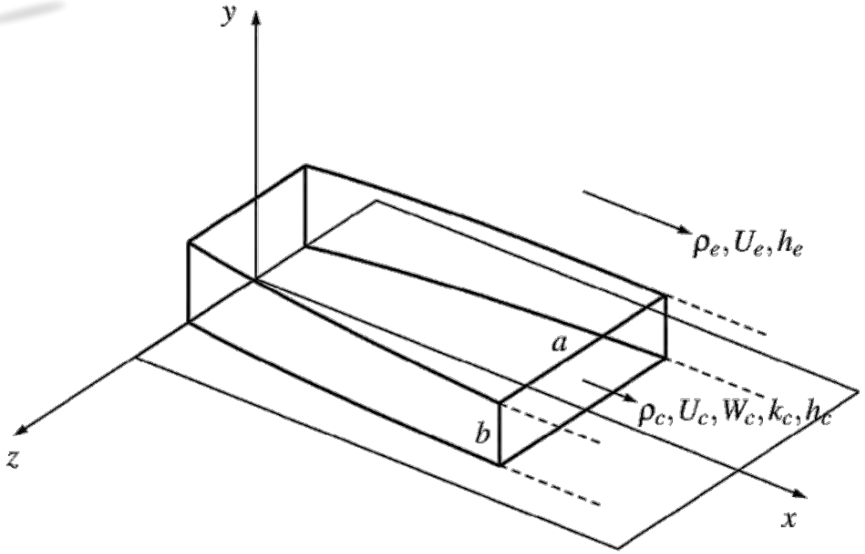


FIG. 1: Downstream development of a film cooling jet.

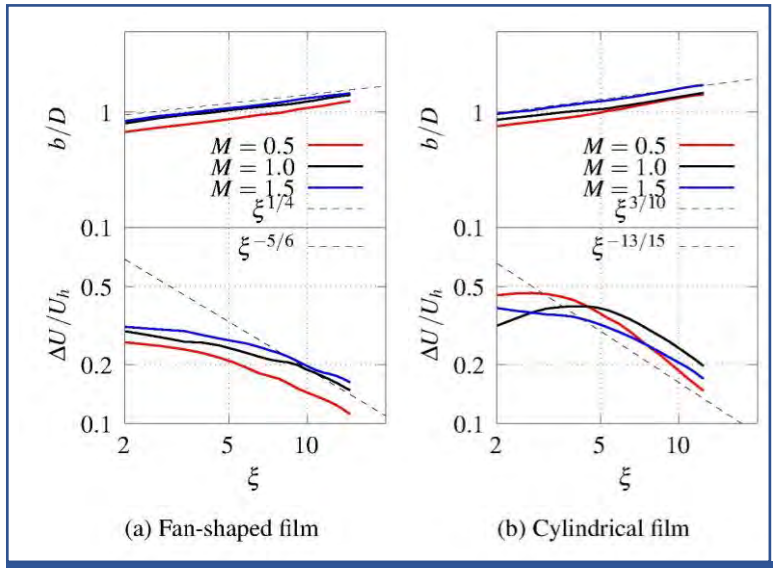
- the power laws

Hao, Muting, and Luca di Mare. "Scaling and similarity laws in three-dimensional wall jets." *Physics of Fluids* (2023).

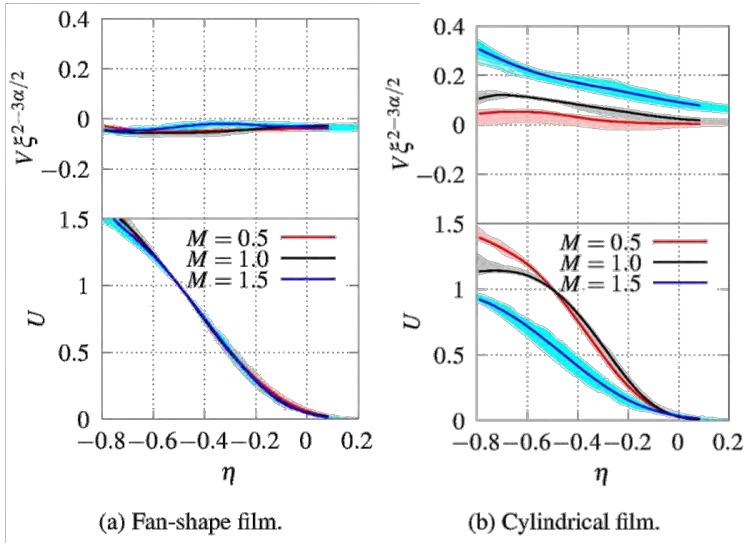
$$\rho \approx \rho_1 \xi^{-\alpha} + 1 \qquad \delta = \alpha$$
$$U \approx U_1 \xi^{-\alpha} + 1 \qquad \beta = \alpha/2$$
$$W \approx W_1 \xi^{-\beta} \qquad \gamma = \alpha$$
$$\kappa \approx \kappa_1 \xi^{-\gamma} \qquad \frac{a}{D} \approx \xi^{1-\alpha/2}$$
$$\sigma \approx \sigma_1 \xi^{\delta} \qquad \frac{b}{D} \approx \xi^{3\alpha/2-1}$$

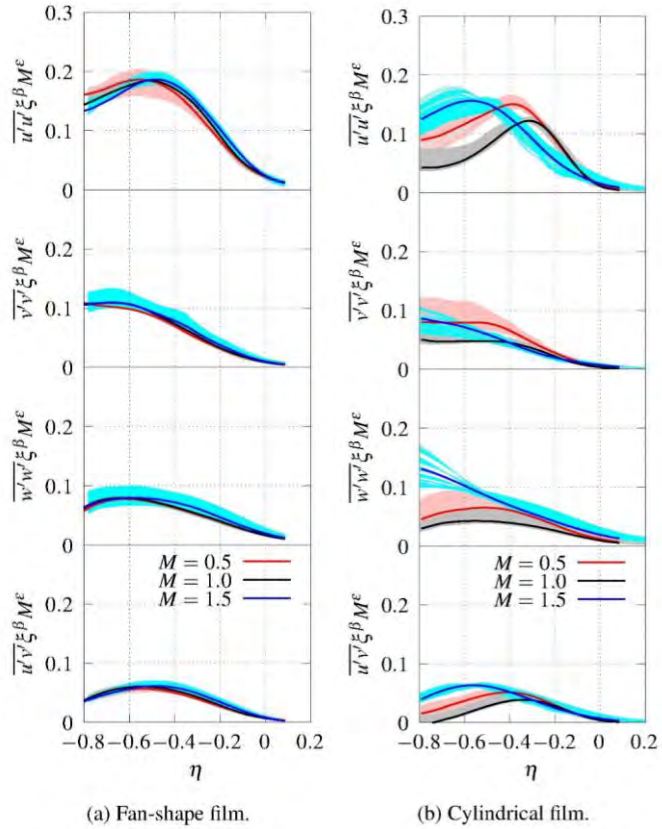
Theory proposed--

Scaling & similarity laws in 3D wall jets



Downstream evolution of velocity defect and jet layer thickness





I. Excellent agreement between the observed power law and the predicted power law]

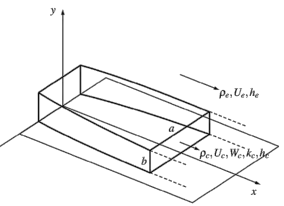


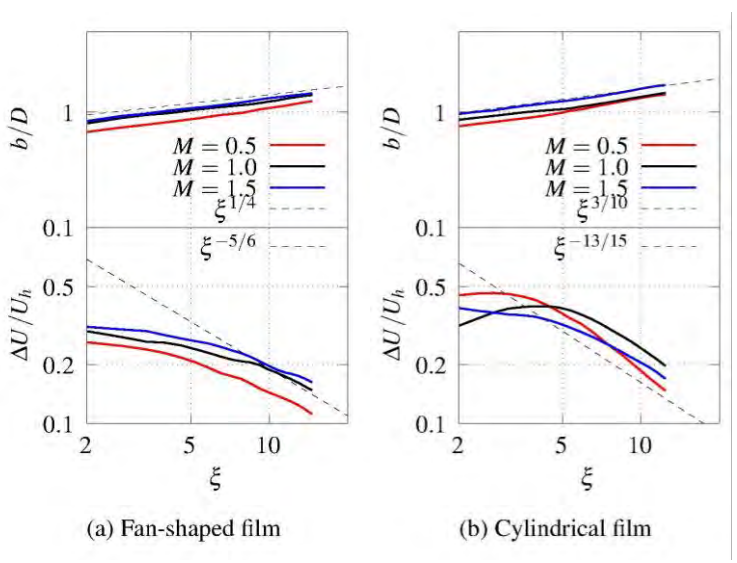
FIG. 1: Downstream development of a film cooling jet.

Hao, Muting, and Luca di Mare. "Scaling and similarity laws in three-dimensional wall jets." *Physics of Fluids* (2023).

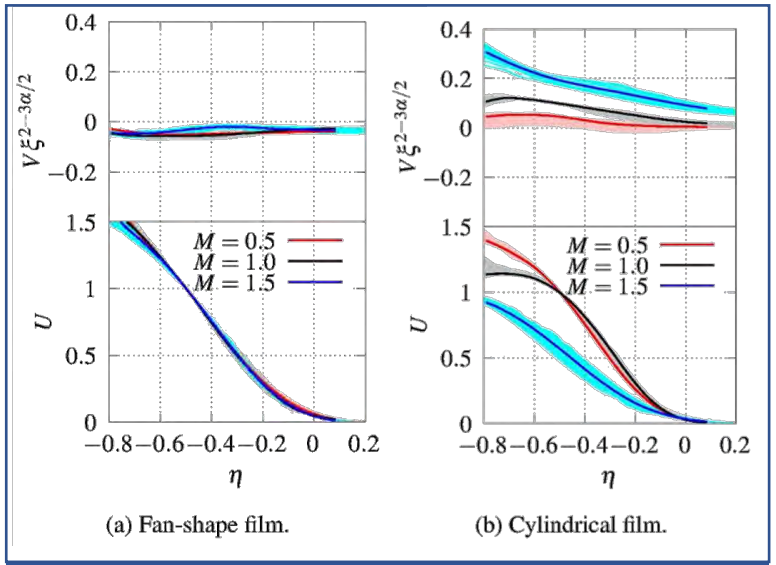
UNIVERSITY OF OXFORD

Theory proposed--

Scaling & similarity laws in 3D wall jets

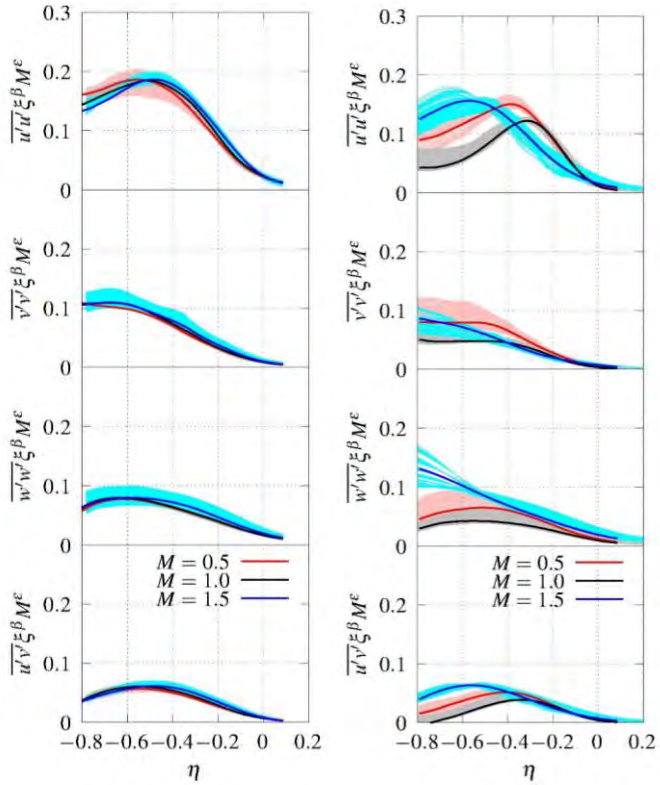


(a) Fan-shaped film (b) Cylindrical film



(a) Fan-shape film. (b) Cylindrical film.

Mean streamwise velocity & wall-normal velocity



(a) Fan-shape film. (b) Cylindrical film.

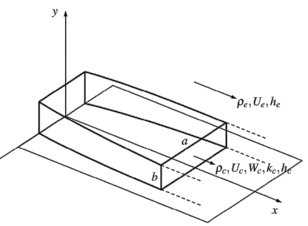


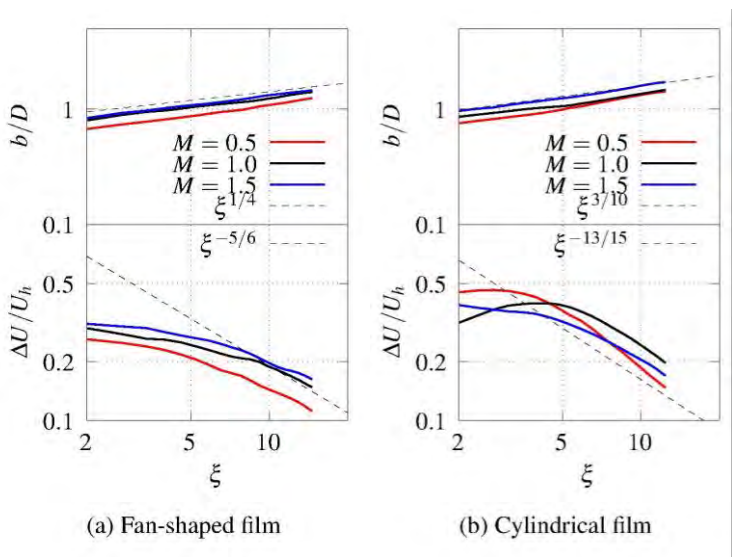
FIG. 1: Downstream development of a film cooling jet.

Hao, Muting, and Luca di Mare. "Scaling and similarity laws in three-dimensional wall jets." *Physics of Fluids* (2023).

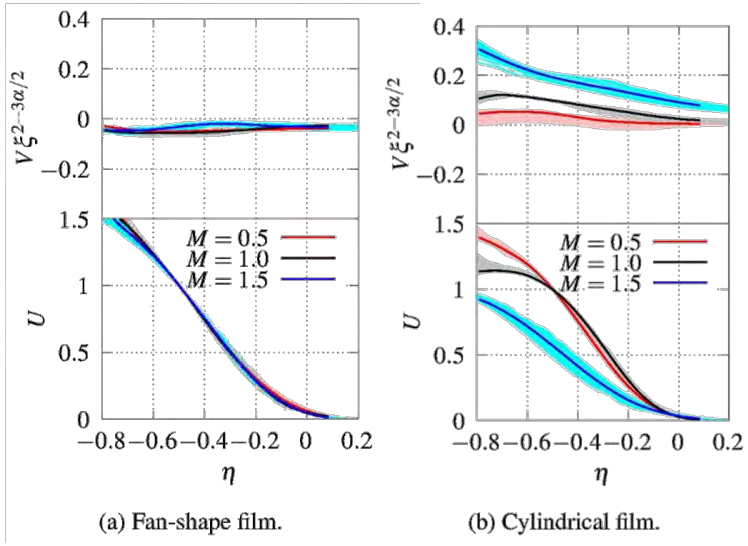
UNIVERSITY OF OXFORD

Theory proposed--

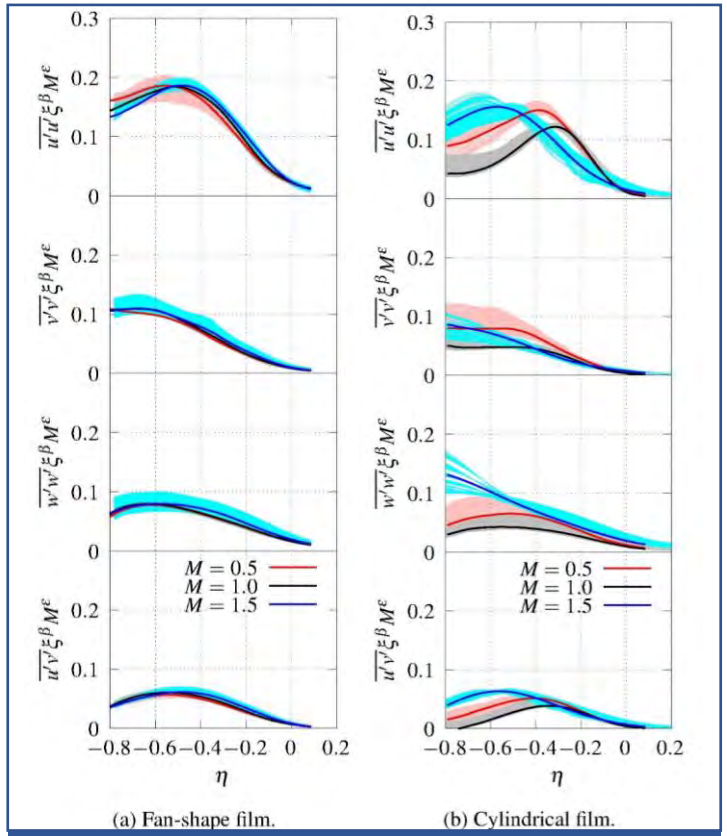
Scaling & similarity laws in 3D wall jets



(a) Fan-shaped film (b) Cylindrical film



(a) Fan-shape film. (b) Cylindrical film.



(a) Fan-shape film. (b) Cylindrical film.

Reynolds stresses

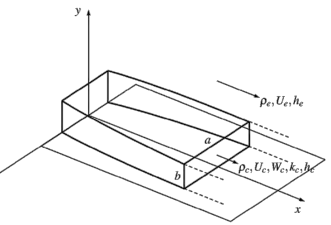


FIG. 1: Downstream development of a film cooling jet.

Hao, Muting, and Luca di Mare. "Scaling and similarity laws in three-dimensional wall jets." *Physics of Fluids* (2023).

UNIVERSITY OF OXFORD

LES statistics--

Budgets of Reynolds stresses in film cooling

Hao, Muting, and Luca di Mare. "Budgets of Reynolds stresses in film cooling with fan-shaped and cylindrical holes." *Physics of Fluids* (2023).

Fan-shaped hole

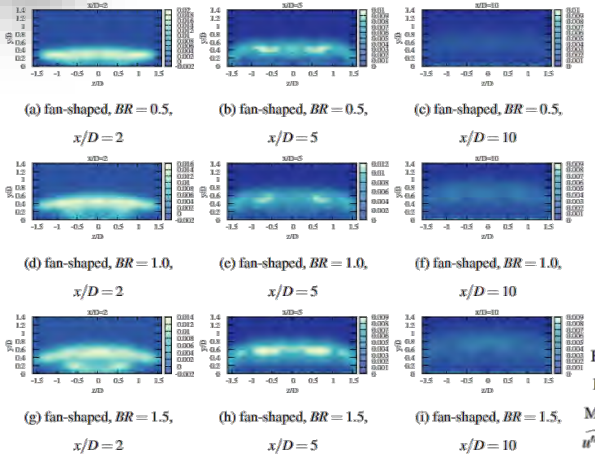


FIG. 9: Contours of Pr_{uu} for fan-shaped films at $x/D = 2, x/D = 5$ and $x/D = 10$

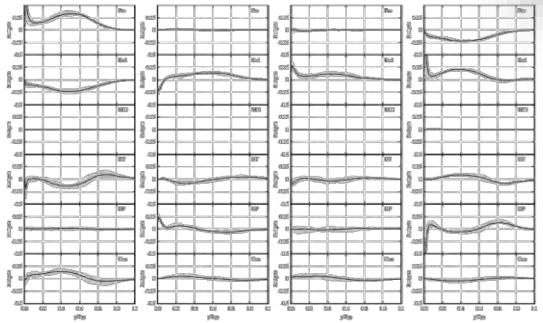


FIG. 11: Budget terms scaled by $(\bar{\rho}\Delta U^3/((x/D)^\beta \delta_{99})$ for fan-shaped film at $BR = 0.5$. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Four columns from left to right are budgets of $\widetilde{u''u''}$, $\widetilde{v''v''}$, $\widetilde{w''w''}$, $\widetilde{u''v''}$. For each column from the top to bottom are corresponding production, redistribution, mass flux term, turbulent diffusion, pressure diffusion, convection.

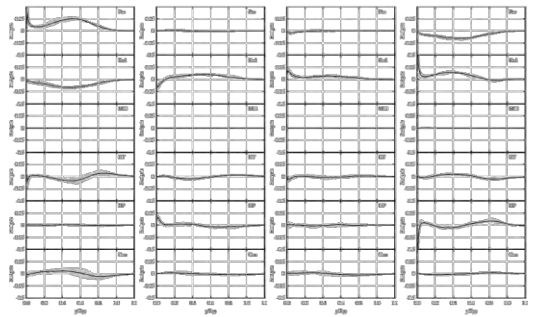


FIG. 12: Budget terms scaled by $(\bar{\rho}\Delta U^3/((x/D)^\beta \delta_{99})$ for fan-shaped film at $BR = 1.0$. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Four columns from left to right are budgets of $\widetilde{u''u''}$, $\widetilde{v''v''}$, $\widetilde{w''w''}$, $\widetilde{u''v''}$. For each column from the top to bottom are corresponding production, redistribution, mass flux term, turbulent diffusion, pressure diffusion, convection.

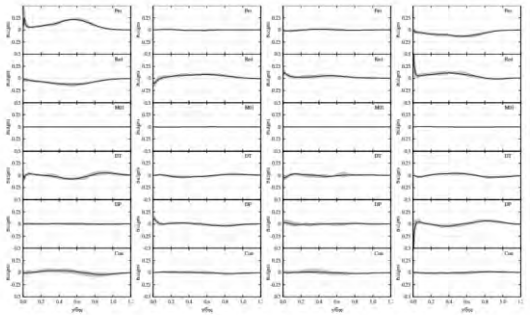


FIG. 13: Budget terms scaled by $(\bar{\rho}\Delta U^3/((x/D)^\beta \delta_{99})$ for fan-shaped film at $BR = 1.5$. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Four columns from left to right are budgets of $\widetilde{u''u''}$, $\widetilde{v''v''}$, $\widetilde{w''w''}$, $\widetilde{u''v''}$. For each column from the top to bottom are corresponding production, redistribution, mass flux term, turbulent diffusion, pressure diffusion, convection.

Cylindrical hole

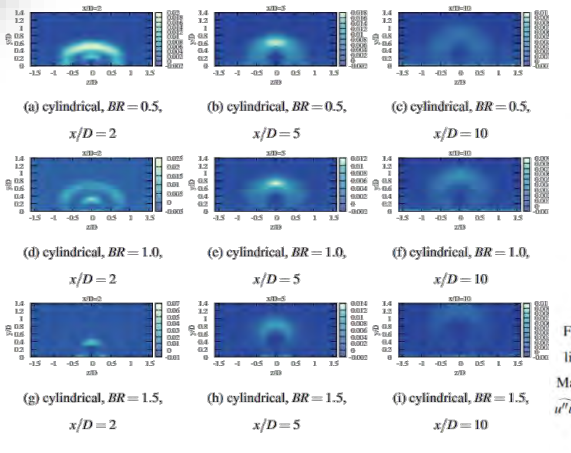


FIG. 10: Contours of Pr_{uu} for cylindrical films at $x/D = 2, x/D = 5$ and $x/D = 10$

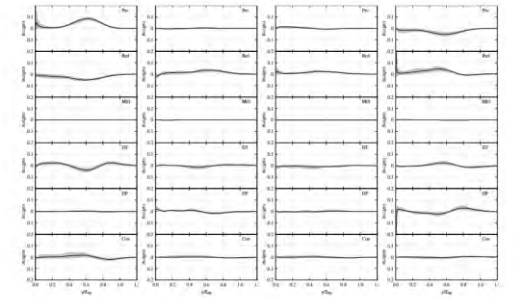


FIG. 14: Budget terms scaled by $(\bar{\rho}\Delta U^3/((x/D)^\beta \delta_{99})$ for cylindrical film at $BR = 0.5$. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Four columns from left to right are budgets of $\widetilde{u''u''}$, $\widetilde{v''v''}$, $\widetilde{w''w''}$, $\widetilde{u''v''}$. For each column from the top to bottom are corresponding production, redistribution, mass flux term, turbulent diffusion, pressure diffusion, convection.

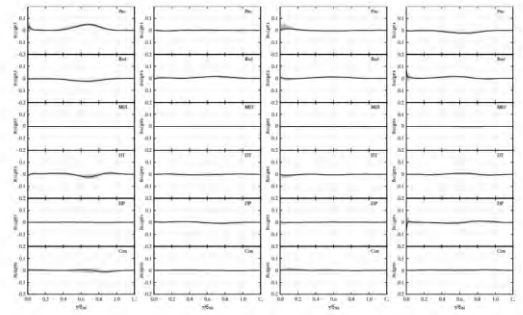


FIG. 15: Budget terms scaled by $(\bar{\rho}\Delta U^3/((x/D)^\beta \delta_{99})$ for cylindrical film at $BR = 1.0$. G
lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lin
Master curve (Average of multiple grey curves). Four columns from left to right are budge
$\widetilde{u''u''}$, $\widetilde{v''v''}$, $\widetilde{w''w''}$, $\widetilde{u''v''}$. For each column from the top to bottom are corresponding produc
redistribution, mass flux term, turbulent diffusion, pressure diffusion, convection.

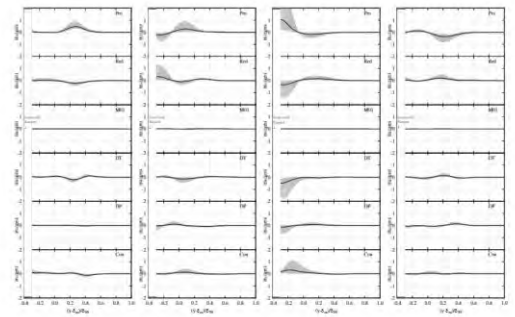


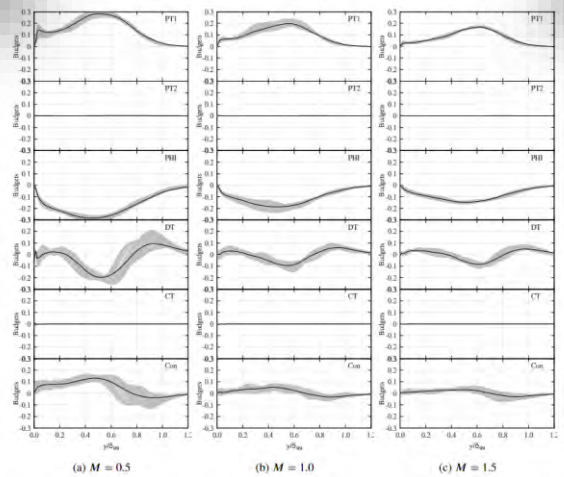
FIG. 17: Budget terms scaled by $(\bar{\rho}\Delta U_m^3/((x/D)^\beta \delta_{99})$ for cylindrical film at $BR = 1.5$. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Four columns from left to right are budgets of $\widetilde{u''u''}$, $\widetilde{v''v''}$, $\widetilde{w''w''}$, $\widetilde{u''v''}$. For each column from the top to bottom are corresponding production, redistribution, mass flux term

UNIVERSITY OF OXFORD

LES statistics--

Budgets of turbulent heat flux in film cooling

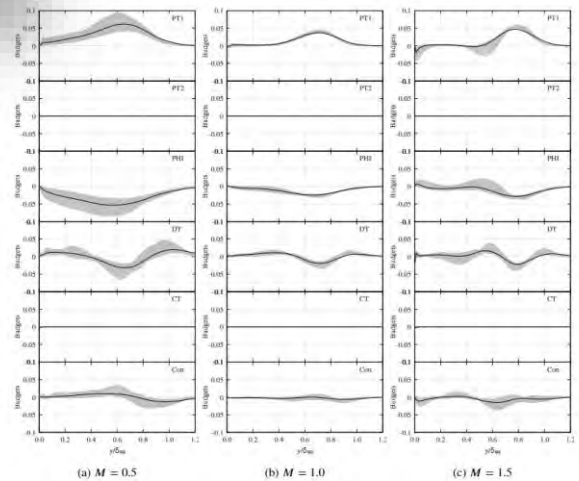
Fan-shaped hole



(a) $M = 0.5$ (b) $M = 1.0$ (c) $M = 1.5$

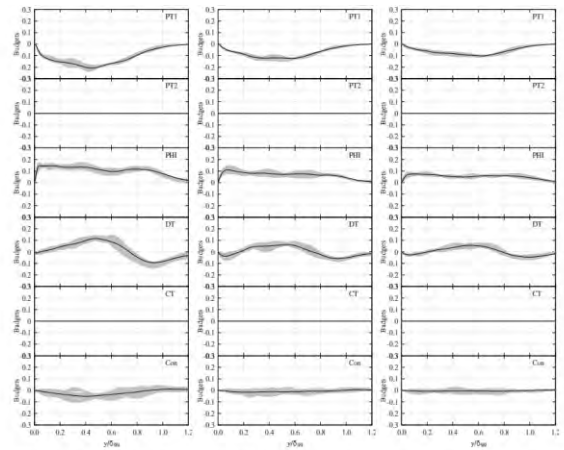
Figure 18: Budget terms of $\widetilde{u''T''}$ scaled by $(\bar{\rho}\Delta U^2\Delta T(x/D)^{-\beta}/\delta_{99}$ for fan-shaped film at three blowing ratios. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 16$, Black lines: Master curve (Average of multiple grey curves). Three columns from left to right are budgets of $M = 0.5$, $M = 1.0$, $M = 1.5$. For each column from the top to bottom are corresponding production due to the mean velocity and mean temperature $PT1$, Production due to the fluctuating strain rate $PT2$, Pressure-scrambling PHI, Turbulent viscous-thermal transport contribution DT, compressibility associated terms CT, and convection Con.

Cylindrical hole



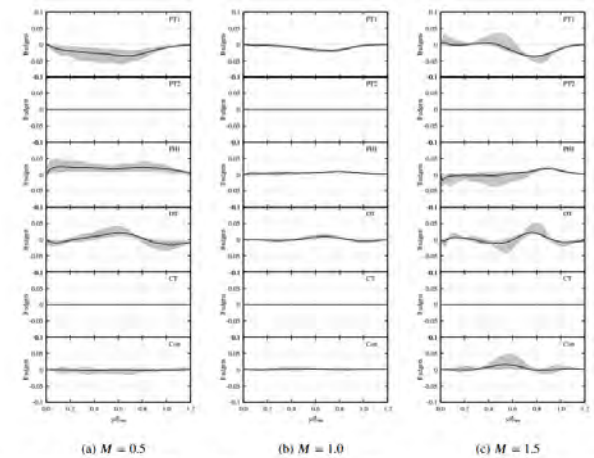
(a) $M = 0.5$ (b) $M = 1.0$ (c) $M = 1.5$

Figure 19: Budget terms of $\widetilde{u''T''}$ scaled by $(\bar{\rho}\Delta U^2\Delta T(x/D)^{-\beta}/\delta_{99}$ for cylindrical film at three blowing ratios. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Three columns from left to right are budgets of $M = 0.5$, $M = 1.0$, $M = 1.5$. For each column from the top to bottom are corresponding production due to the mean velocity and mean temperature $PT1$, Production due to the fluctuating strain rate $PT2$, Pressure-scrambling PHI, Turbulent viscous-thermal transport contribution DT, compressibility associated terms CT, and convection Con.



(a) $M = 0.5$ (b) $M = 1.0$ (c) $M = 1.5$

Figure 20: Budget terms of $\widetilde{v''T''}$ scaled by $(\bar{\rho}\Delta U^2\Delta T(x/D)^{-\beta}/\delta_{99}$ for fan-shaped film at three blowing ratios. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Three columns from left to right are budgets of $M = 0.5$, $M = 1.0$, $M = 1.5$. For each column from the top to bottom are corresponding production due to the mean velocity and mean temperature $PT1$, Production due to the fluctuating strain rate $PT2$, Pressure-scrambling PHI, Turbulent viscous-thermal transport contribution DT, compressibility associated terms CT, and convection Con.



(a) $M = 0.5$ (b) $M = 1.0$ (c) $M = 1.5$

Figure 21: Budget terms of $\widetilde{v''T''}$ scaled by $(\bar{\rho}\Delta U^2\Delta T(x/D)^{-\beta}/\delta_{99}$ for cylindrical film at three blowing ratios. Grey lines: multiple curves on streamwise positions between $x/D = 5$ and $x/D = 15$, Black lines: Master curve (Average of multiple grey curves). Three columns from left to right are budgets of $M = 0.5$, $M = 1.0$, $M = 1.5$. For each column from the top to bottom are correspondingly production due to the mean velocity and mean temperature $PT1$, Production due to the fluctuating strain rate $PT2$, Pressure-scrambling PHI, Turbulent viscous-thermal transport contribution DT, compressibility associated terms CT, and convection Con.

- First finding of budgets in film cooling!!!!!

Hao, Muting, and Luca di Mare. "Budgets of Reynolds stresses in film cooling with fan-shaped and cylindrical holes." *Physics of Fluids* (2023).

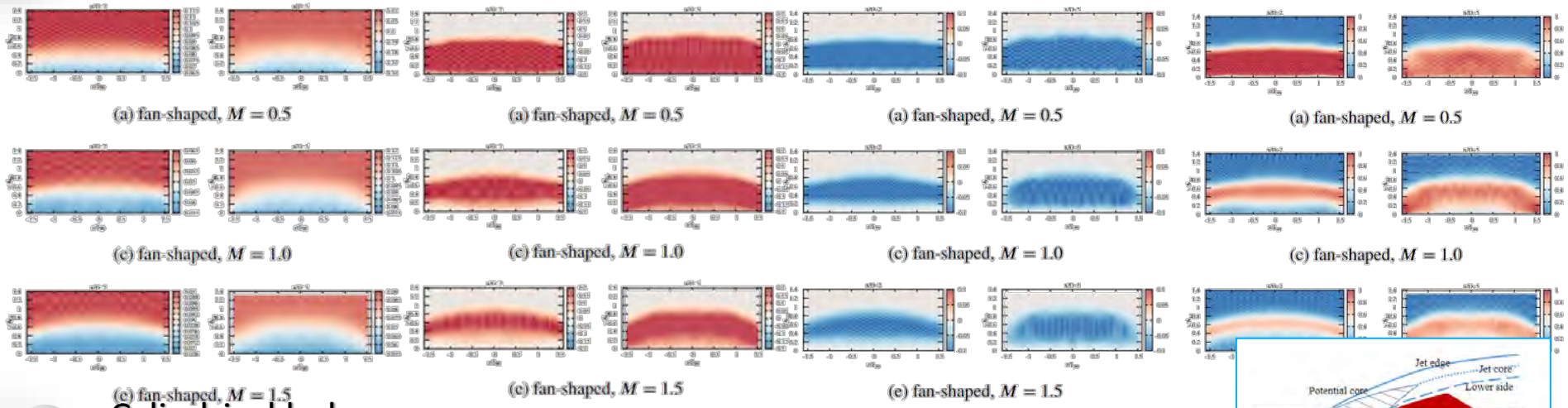
LES statistics--

Budgets of turbulent heat flux in film cooling

Fan-shaped hole

Hao, Muting, and Luca di Mare. "Budgets of Reynolds stresses in film cooling with fan-shaped and cylindrical holes." *Physics of Fluids* (2023).

$\tilde{T}$ $\widetilde{u''T''}$ $\widetilde{v''T''}$ $\widetilde{T''T''}$



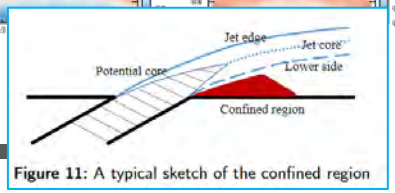
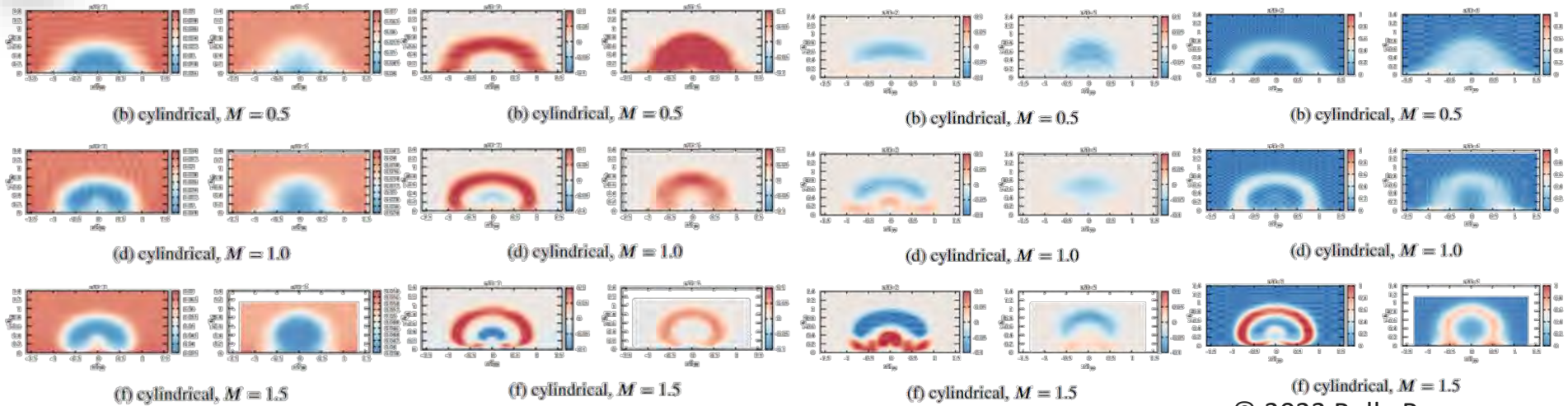


Figure 11: A typical sketch of the confined region

Cylindrical hole



UNIVERSITY OF OXFORD

The findings thanks to HPC

- Each ‘shaded & banded profiles’ include 200 positions downstream from the film cooling hole collapsing together after processing the data using a proposed power law

- The discovery replies on a great amount of statistics

- Recalling that
  - all these results are based on millions of timesteps
  - For 1 film cooling hole, the mesh is 10.5 million grids
  - 1 operating condition case: takes 132440 core hours

- The discovery of power law thanks to HPC!!!

LES results – Prospect

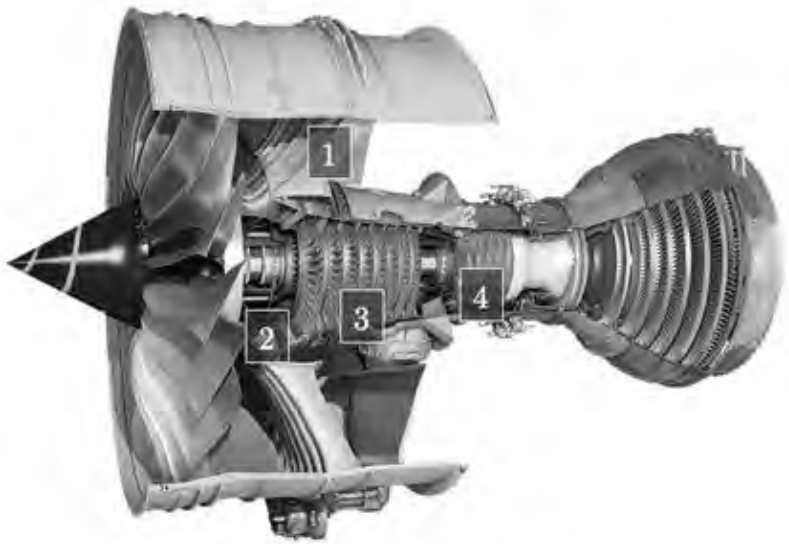
✓Implications

My research Part II

High Bypass Ratio Turbofan Design

What is high bypass?

Fan blades



1. Outlet guide vanes (OGV)
2. Engine stator section (ESS)
3. Intermediate pressure compressor (IPC)
4. High pressure compressor (HPC)

- An effective strategy to improve the performance

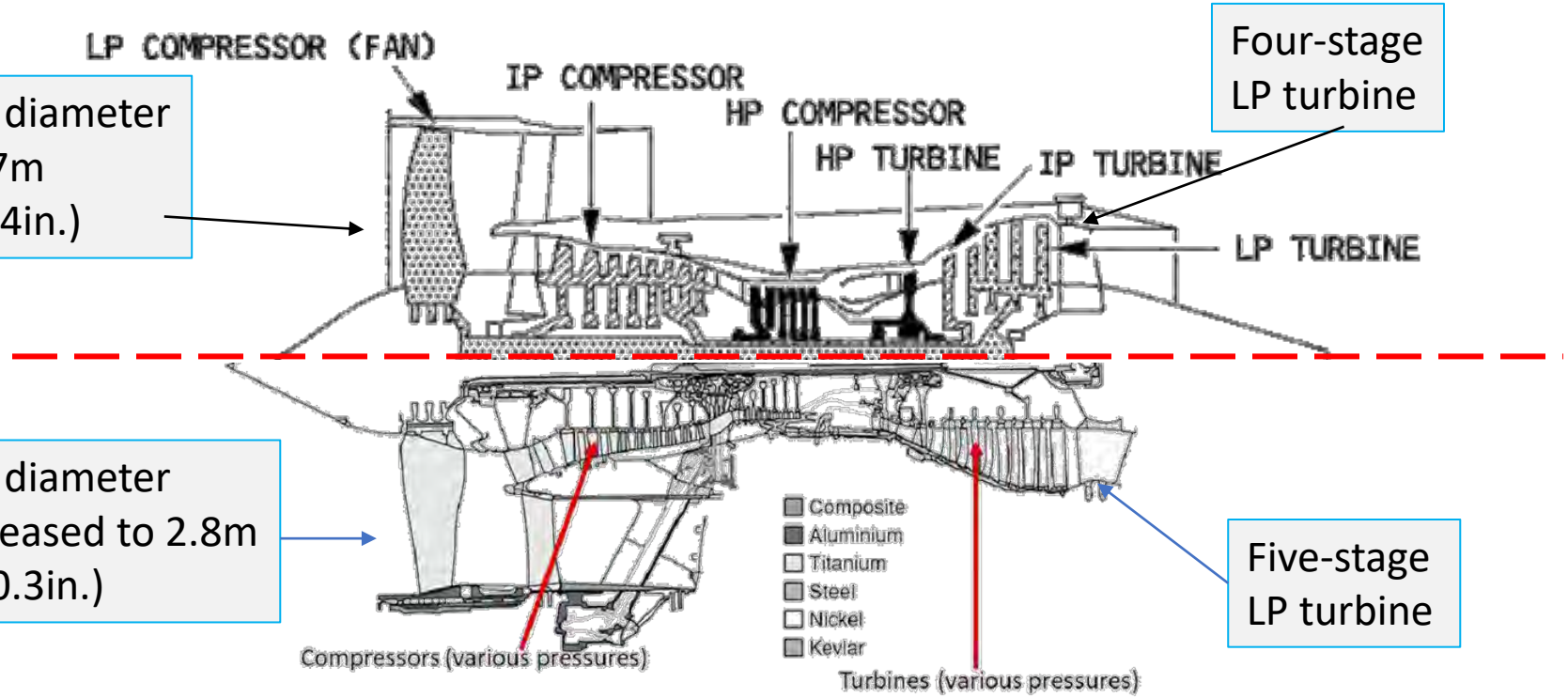
A Rolls-Royce modern, high bypass ratio, large civil aero-engine.
Printed with permission from Rolls-Royce plc. [1]

[1] Seshadri, Pranay & Parks, Geoffrey & Shahpar, Shahrokh. (2016). An Aerodynamic Analysis of a Robustly Redesigned Modern Aero-Engine Fan. arXiv.

Challenges for turbofan design due to bypass?

Trent700

Fan diameter
2.47m
(97.4in.)



Trent800

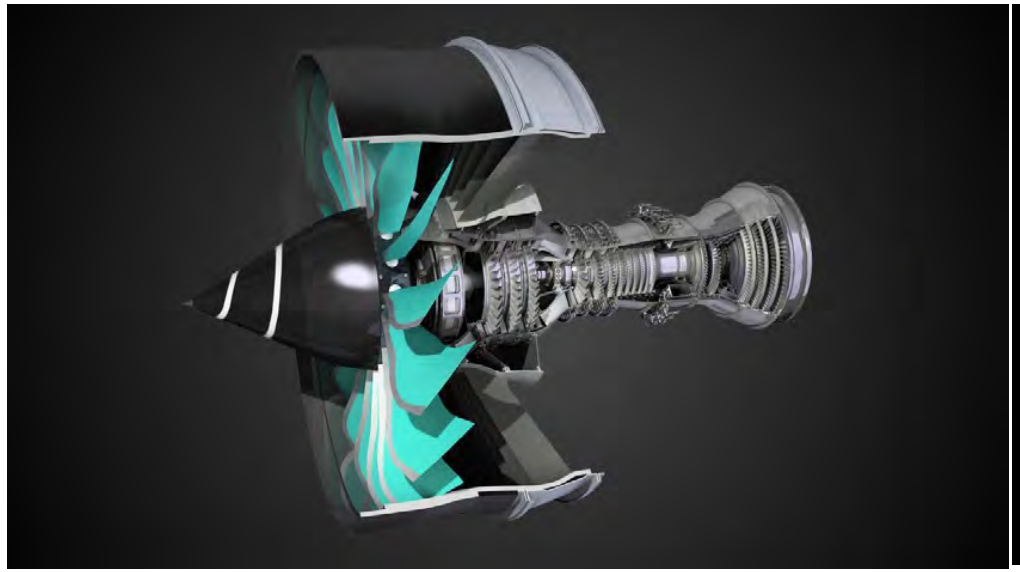
Fan diameter
Increased to 2.8m
(110.3in.)

- *stronger interactions*,

---- *performance and integrity issues!*

Challenges for future ultra turbofan design

Future Ultrafan



- *stronger interactions*

  ----to be addressed at the design stage !

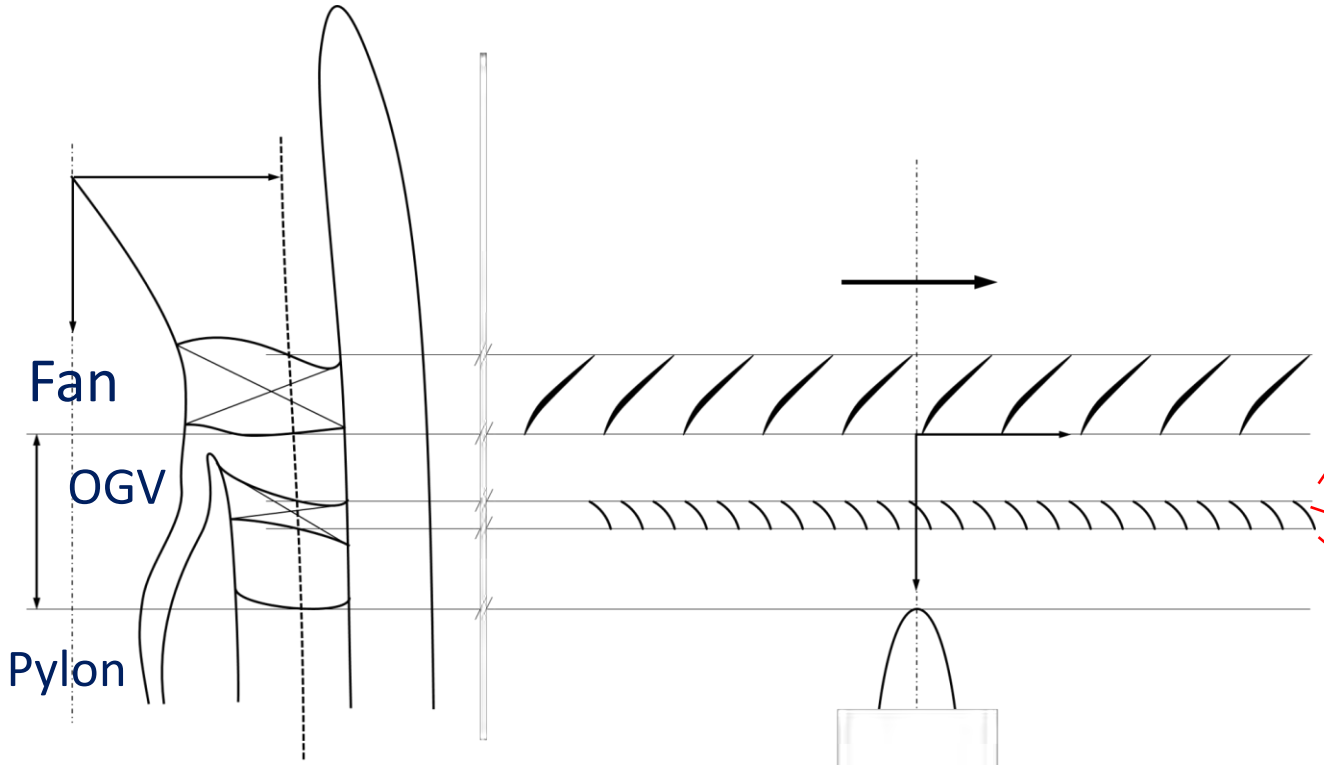
- *The direct computation of whole LPC systems through CFD is expensive*

Code configuration & efficiency

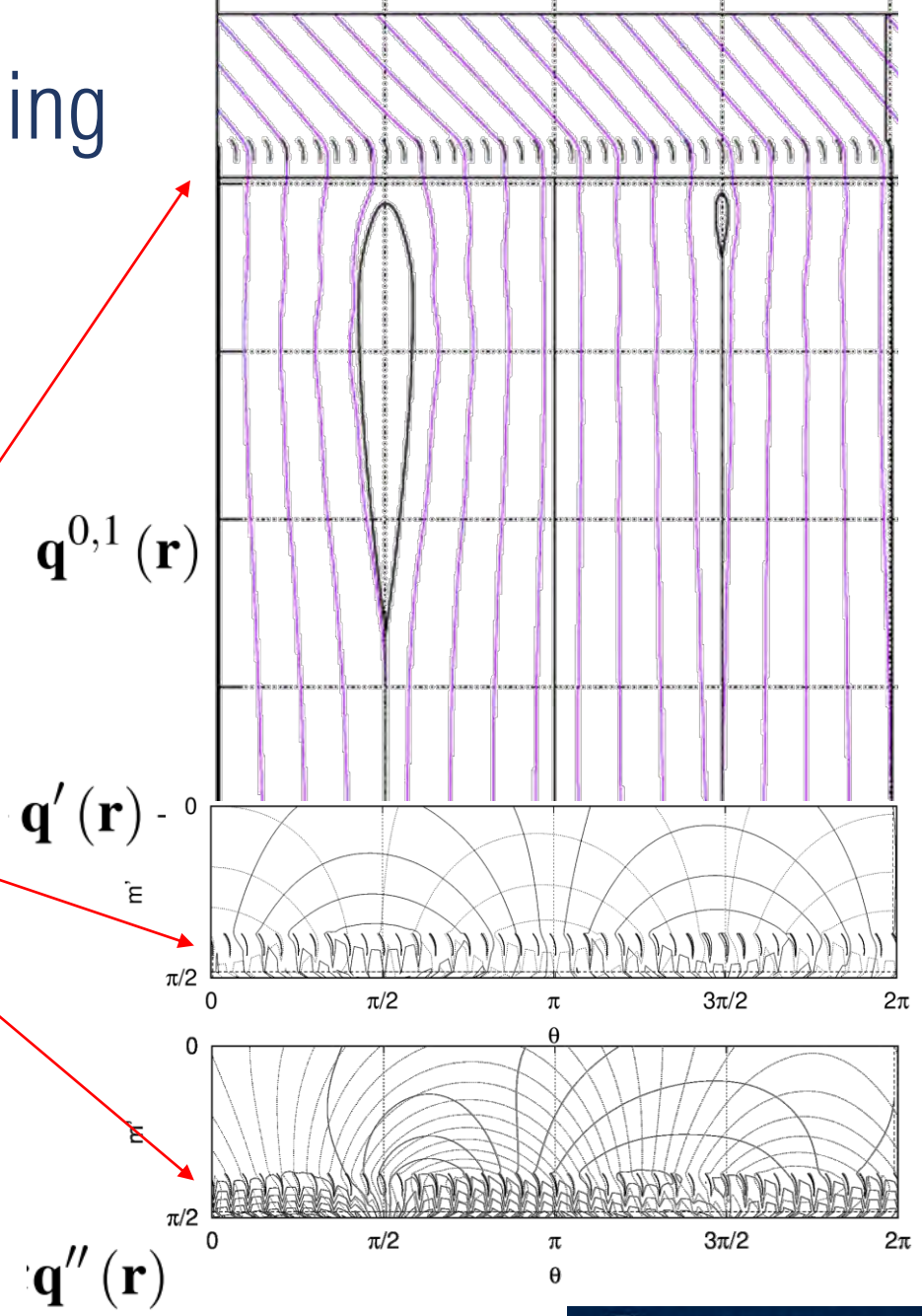
- an efficient computational approach for optimal configurations
- Flow solver fast within minutes
- *Demonstrated as a useful preliminary design tool!*

Methodology for computational modelling

- Proposed a semi-analytical model:



OGV



OGV: Outlet Guide Vane

Results and design

- *Validated against CFD results on a configuration typical of modern machines for civil aviation service!*

World's most efficient large aero-engine

Future Directions and Possibilities

1. **More capability of code**

   Faster

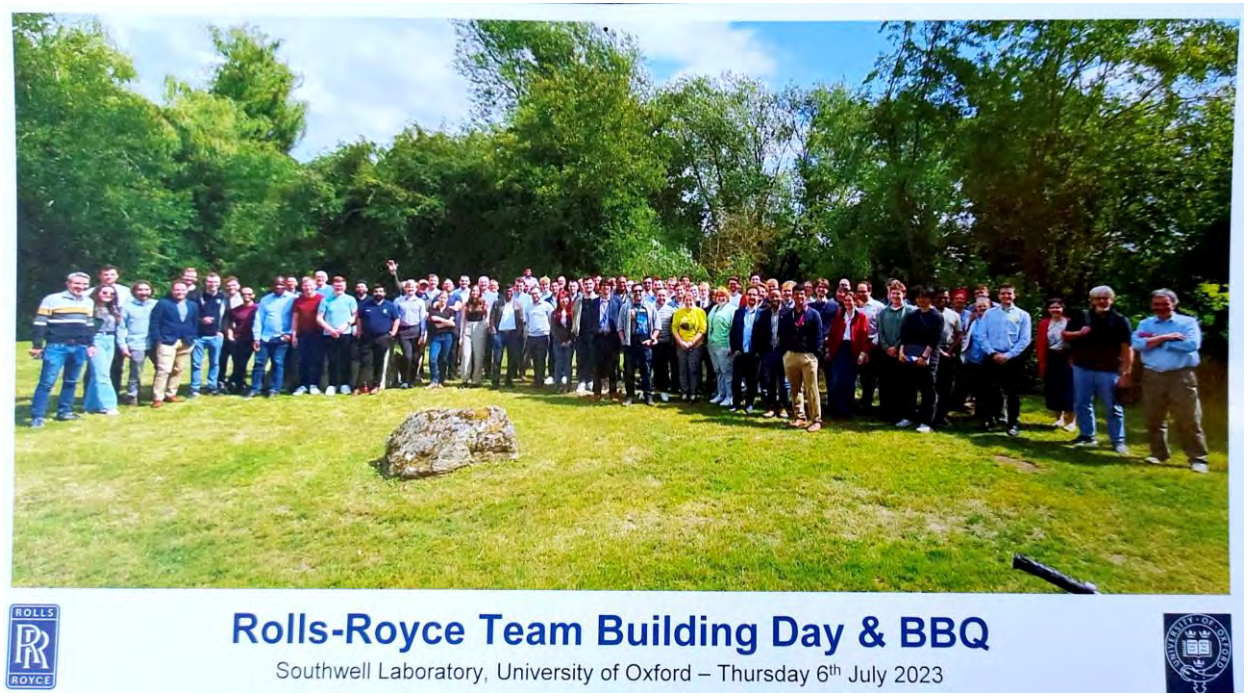
2. **Collaboration and Partnerships**

3. **Continued**

   HPC of the code

Acknowledgement

University of Oxford
Oxford Thermalfluid Institute
Rolls-Royce plc.
UKRI
Specialists in Rolls Royce:
Dr. Haidong Li
Dr. Adami Paolo
Dr. Davendu Y. Kulkarni
Dr. Bharat Lad
Dr. Goenaga Frederic
Dr. Romero Eduardo



Thank you!

Special thanks to
Prof. Luca di Mare
University of Oxford

and
each wonderful person
in our Numerical Analysis Group!

muting.hao@eng.ox.ac.uk
Muting Hao

UNIVERSITY OF OXFORD

UNIVERSITY OF OXFORD

Thank you!

muting.hao@eng.ox.ac.uk

CIUK 2023 Presentations

Mark Parson (EPCC, University of Edinburgh)

The UK's Exascale Supercomputer

Bio: Prof. Mark Parsons joined EPCC, the supercomputing centre at The University of Edinburgh, in 1994 as a software developer working on several industrial contracts following a PhD in Particle Physics undertaken on the LEP accelerator at CERN in Geneva. In 1997 he became the Centre's Commercial Manager and subsequently its Commercial Director. Today he is EPCC's Director and also the Dean of Research Computing at the University of Edinburgh.

He has many interests in distributed computing ranging from its industrial use to the provision of pan-European HPC services through the PRACE Research Infrastructure. His research interests include highly distributed data intensive computing and novel hardware design.

UK Research and Innovation

The UK Exascale Project CIUK Conference

8th December 2023

Professor Mark Parsons
EPSRC Director of Research Computing
EPCC Director

Future of Compute review

- Review report published on 6th March “Tech Moment”
- Published alongside Science & Technology Framework
- Report makes 10 recommendations split into 3 themes
  - Theme 1: Unlock the world-leading, high-growth potential of UK compute
  - Theme 2: Build world-class, sustainable compute capabilities
  - Theme 3: Empower the compute community
- Only two recommendations focus on immediate investment – Exascale and AI Research resources

Future of Compute review

"Have a long term rolling 10-year strategy"

"Invest in infrastructure for compute and AI"

- Review repo[illegible]th Marc[illegible]
- Published al[illegible]gside Science &[illegible]echnology Framework
- Report makes 10 recom[illegible]endations split into 3 themes
  - Theme 1: Unlock the world-leading, high-growth potential of UK compute
  - Theme 2: Build world-class, sustainable compute capabilities
  - Theme 3: Empower the compute community
- Only two recommendations focus on immediate investment – Exascale and AI Research resources

"Invest in software and skills"


Department for Science, Innovation & Technology

Independent report

Independent Review of The Future of Compute: Final report and recommendations

Updated 6 March 2023

Contents

Introduction from the expert panel
Report outline
List of recommendations
Proposed timeframe
Glossary of terms
1. The significance of compute for the UK
2. The international landscape of compute
3. The demand for compute in the UK
4. Meeting the UK's compute needs
5. Creating a vibrant compute ecosystem
6. Recommendations
Acknowledgements

Future of Compute Review – phased approach

- Phase 1: Immediately deliver hardware that supports a wide range of demands from research and business communities. This should provide at least 250 petaflops with enough performance and capacity to support current and future user requirements.
- Phase 2: Deliver hardware that has at least one exaflop of processing power … This should be delivered no later than 2026, and within 2 years of phase 1 to maximise investment.

- Review proposed Phase 1 by 2024 and Phase 2 by 2026

Recent progress

- Project bodies fully established
  - Exascale Project Board
  - Scientific and Industrial Advisory Board
  - Technical Procurement Group (TPG)
  - EPCC at University of Edinburgh announced as hosting site
- Development of Outline Business Case to Government almost complete
  - Approval expected in late March / early April
- Procurement process for 250PF system started
  - Procurement Information Notice (PIN) published in October
  - Detailed responses received from multiple vendors

ACF Computer Room 4 – space for 270 standard racks

TPG at Supercomputing 2023

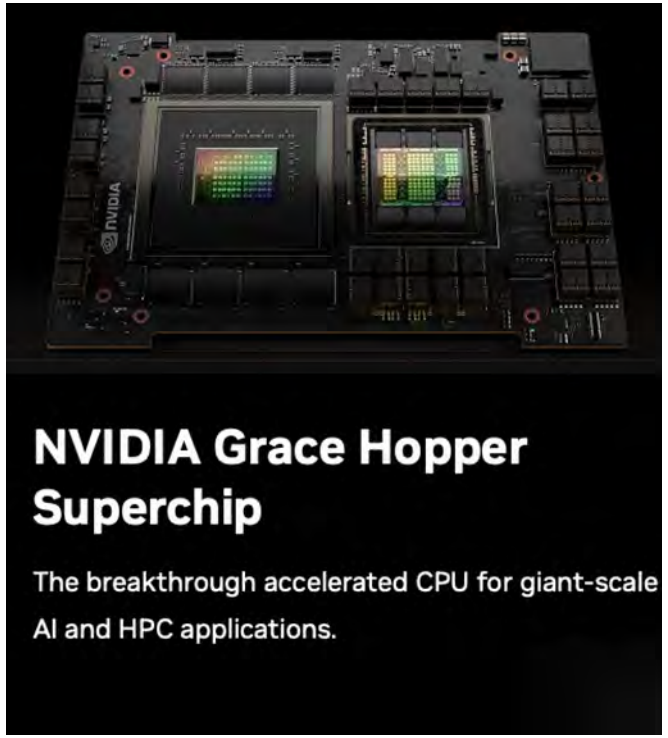
- TPG went to Denver for Supercomputing 2023 in November
- Formal pre-market engagement
- Over 14,000 attendees
- 438 Exhibitors
- Formal meetings with 11 vendors
  - 1 – 2 hours
- Lots of information also gathered from Exhibition
- Very successful trip

Likely technologies

- Only a small number of vendors can supply dense, fully liquid cooled systems which are required for Exascale
- There are two main silicon options in the timescale of Phase 1:

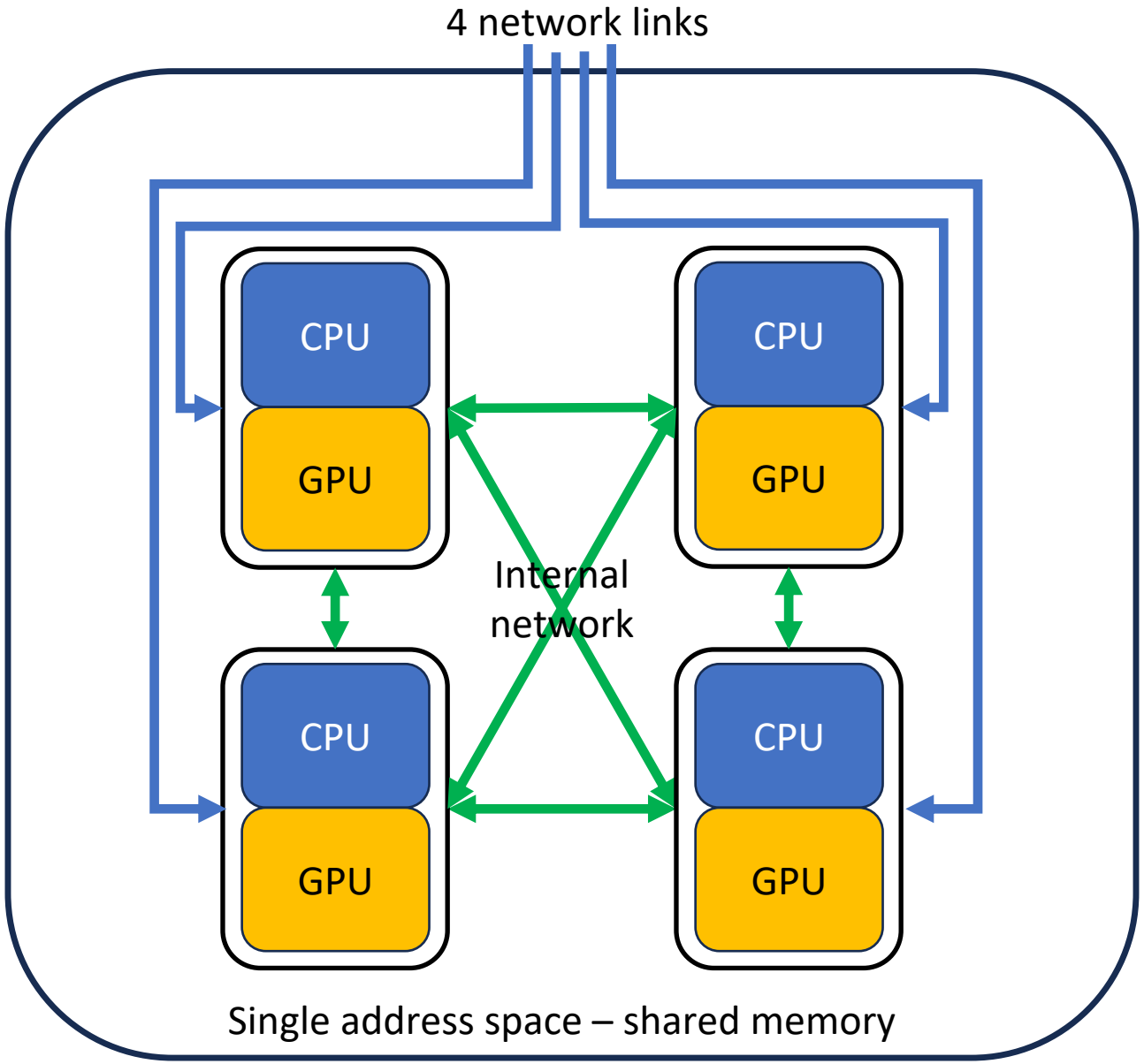
AMD

NVIDIA



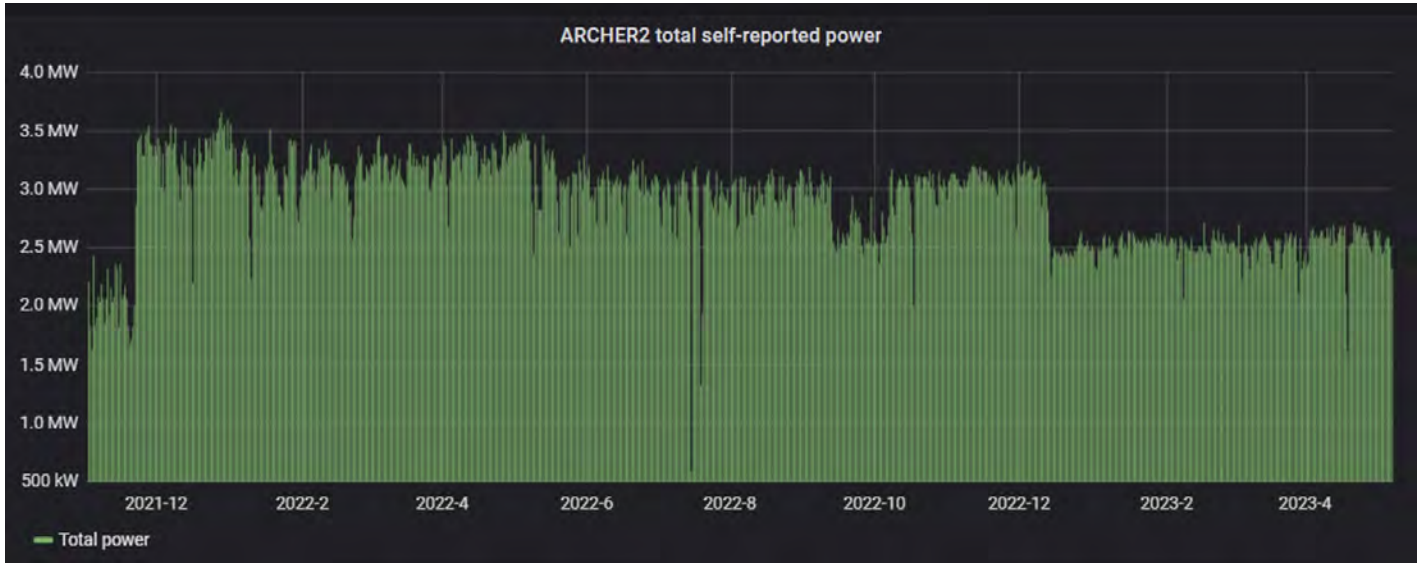
A "Typical" Exascale node

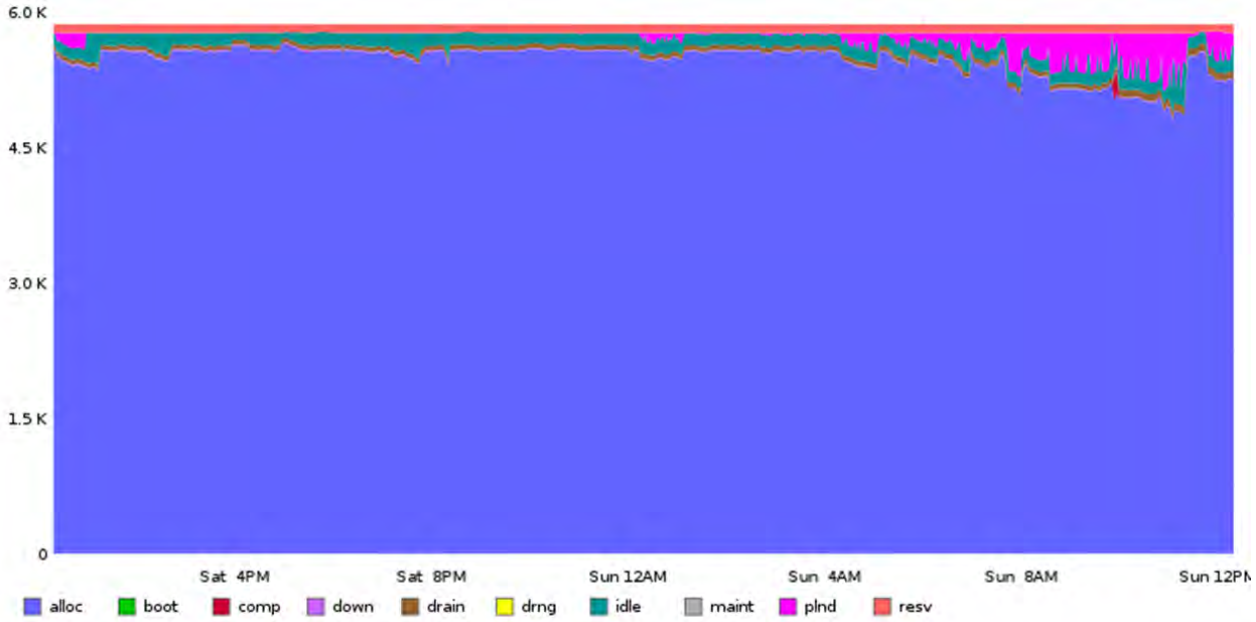
- 4 CPU/GPU modules
- Connected by very fast internal network
- 4 network links for interconnect topology
- Shared memory
- Key innovations
  - Very high-performance GPUs
  - Power performance ratio 10X better than ARCHER2
  - Single address space
  - Coherent shared memory
- Key challenges
  - NUMA effects
  - Programming models

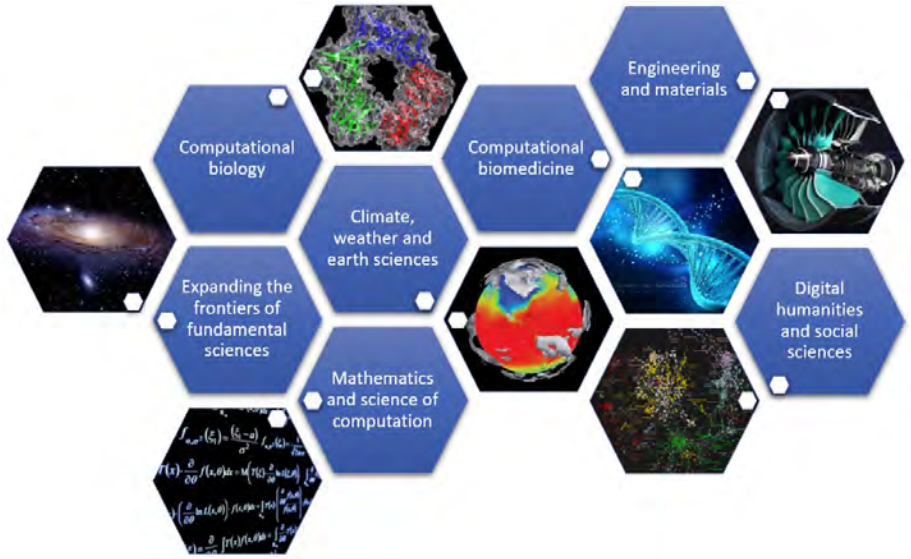


ARCHER2 is full of science …

This is the most consistently busy system EPSRC has ever brought into operation

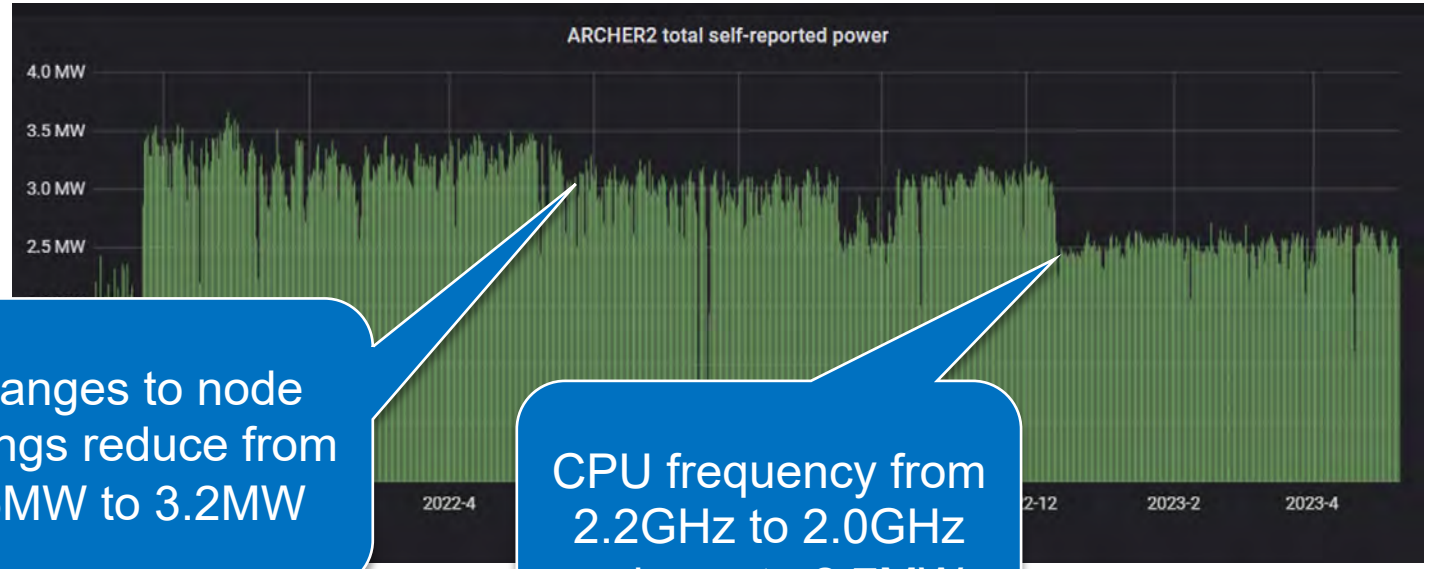






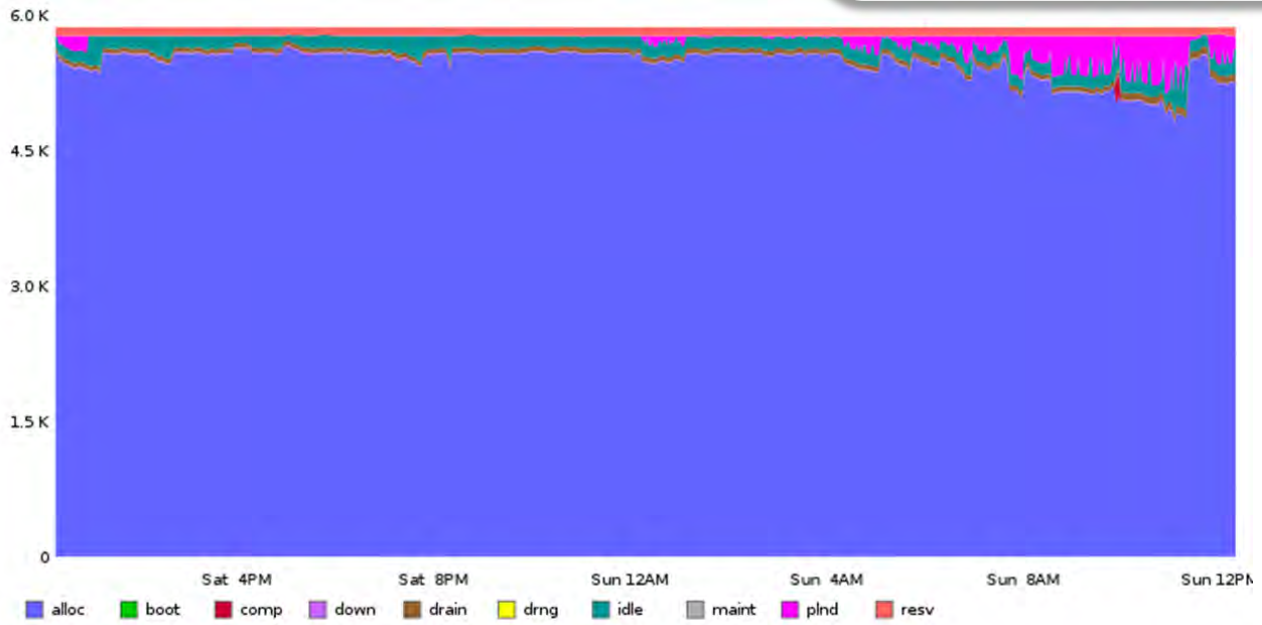
ARCHER2 is full of science …

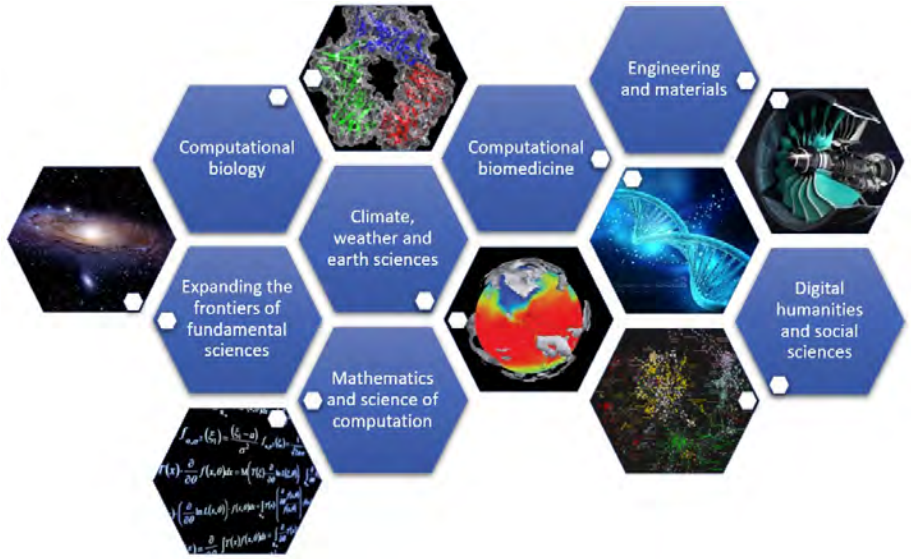
This is the most consistently busy system EPSRC has ever brought into operation



Changes to node settings reduce from 3.5MW to 3.2MW

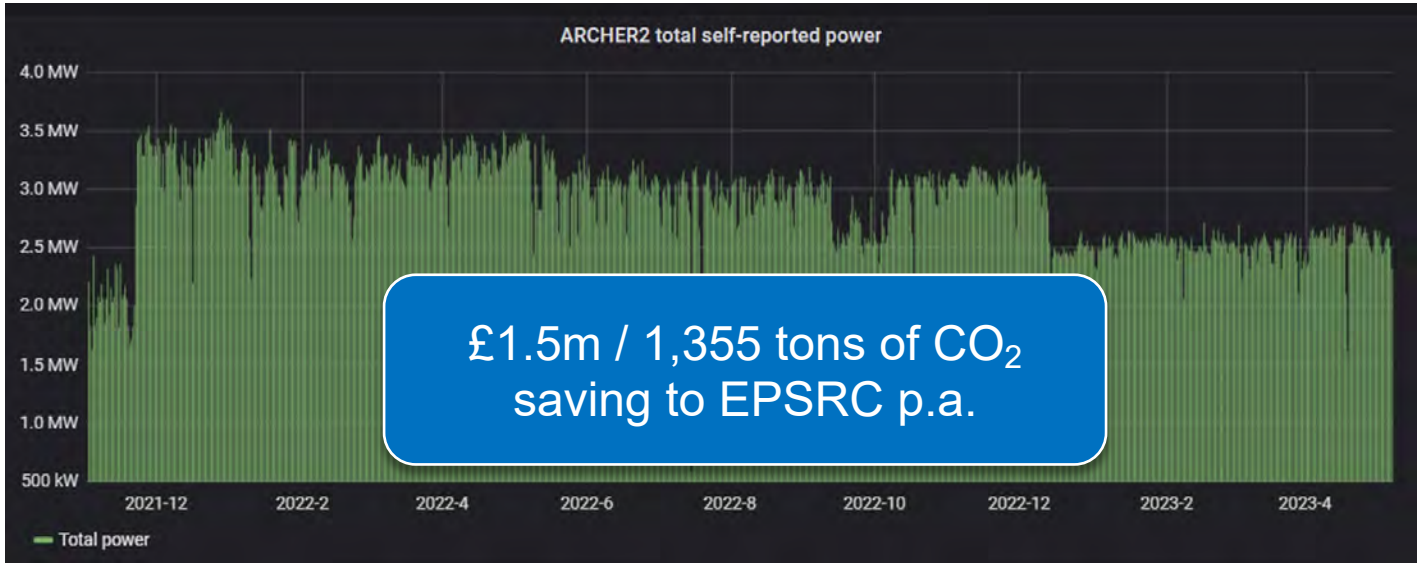
CPU frequency from 2.2GHz to 2.0GHz reduces to 2.7MW



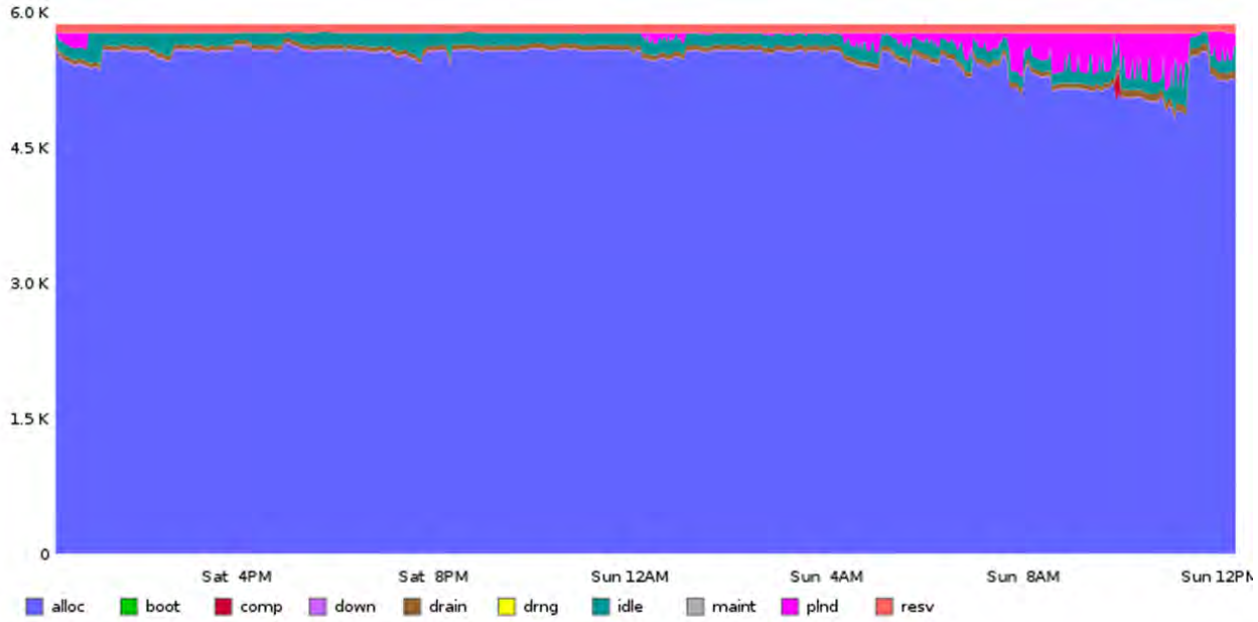


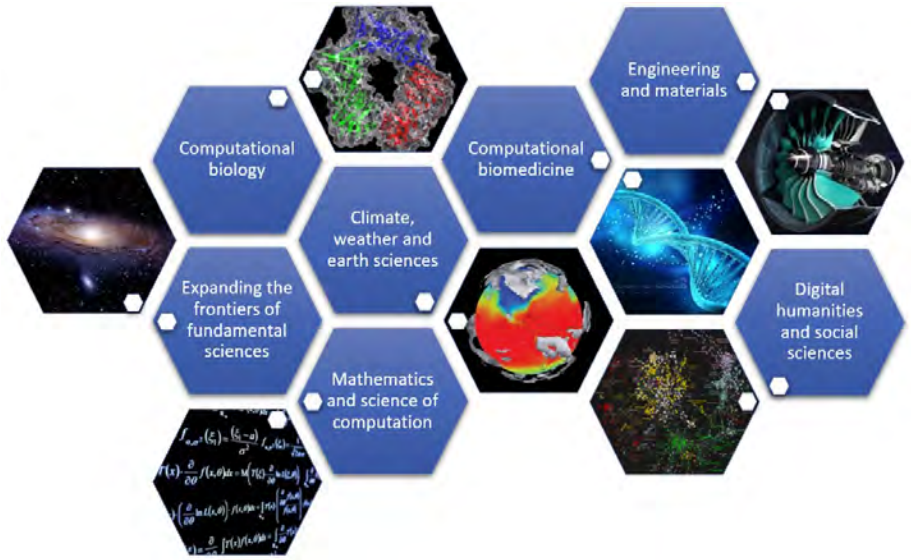
ARCHER2 is full of science …

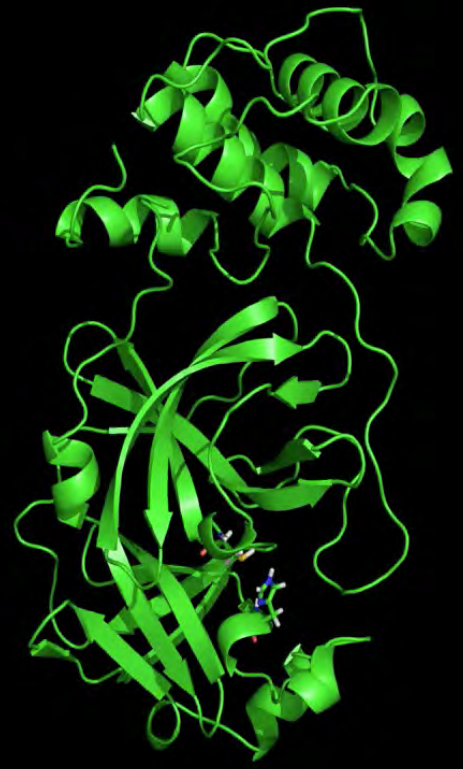
This is the most consistently busy system EPSRC has ever brought into operation

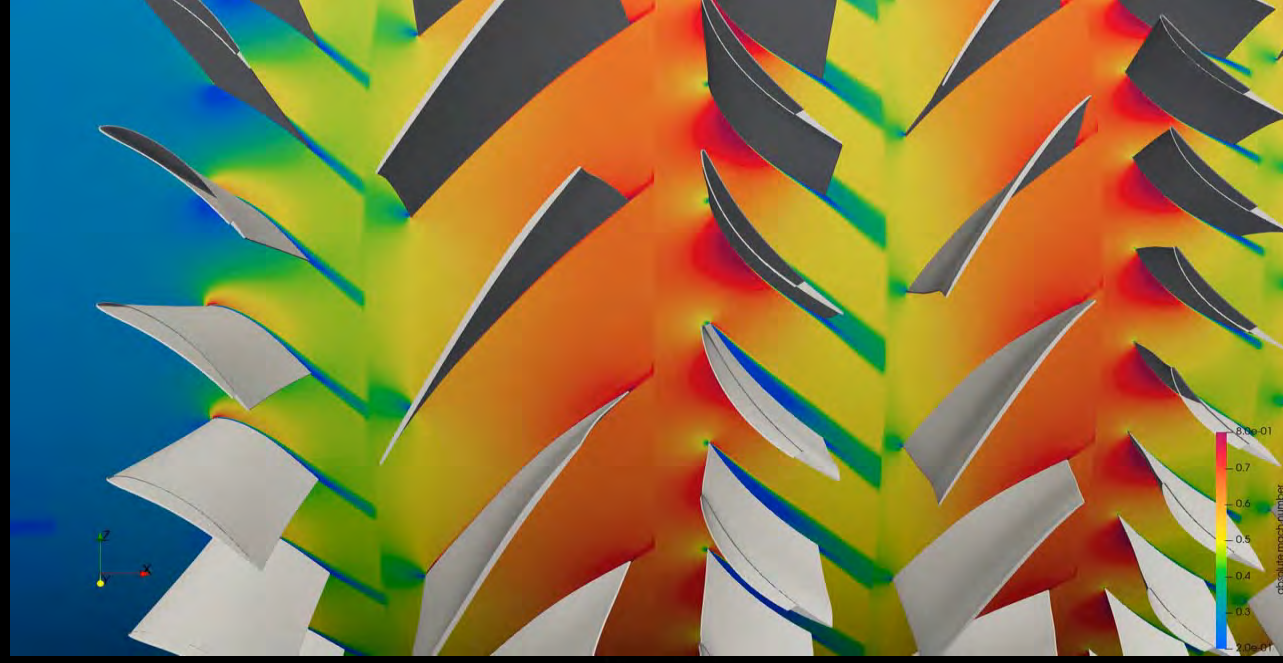


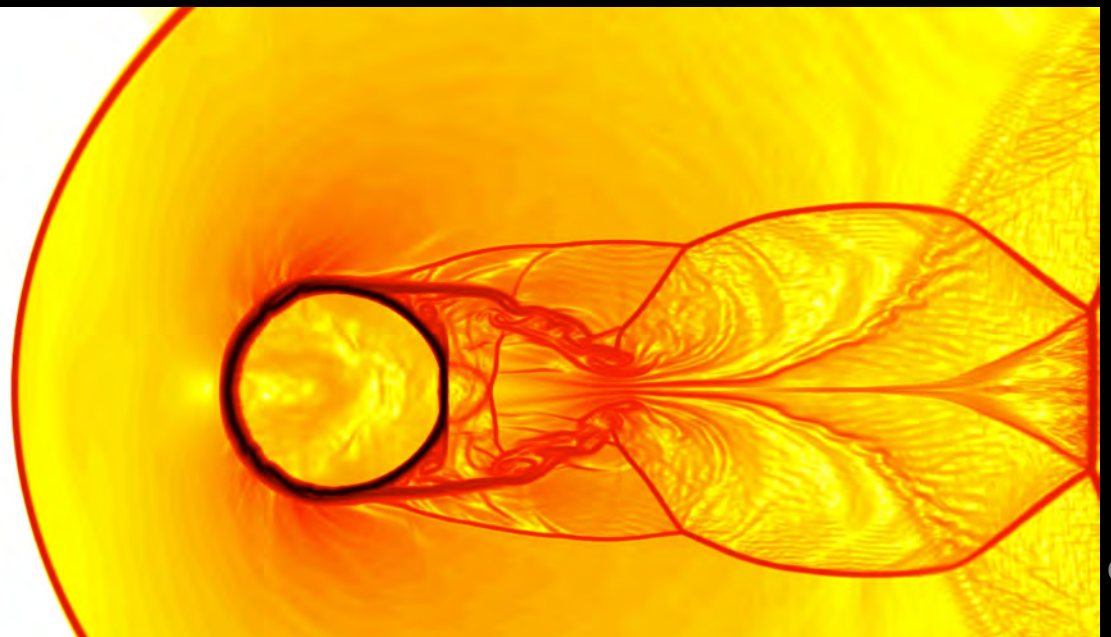
£1.5m / 1,355 tons of CO_2 saving to EPSRC p.a.

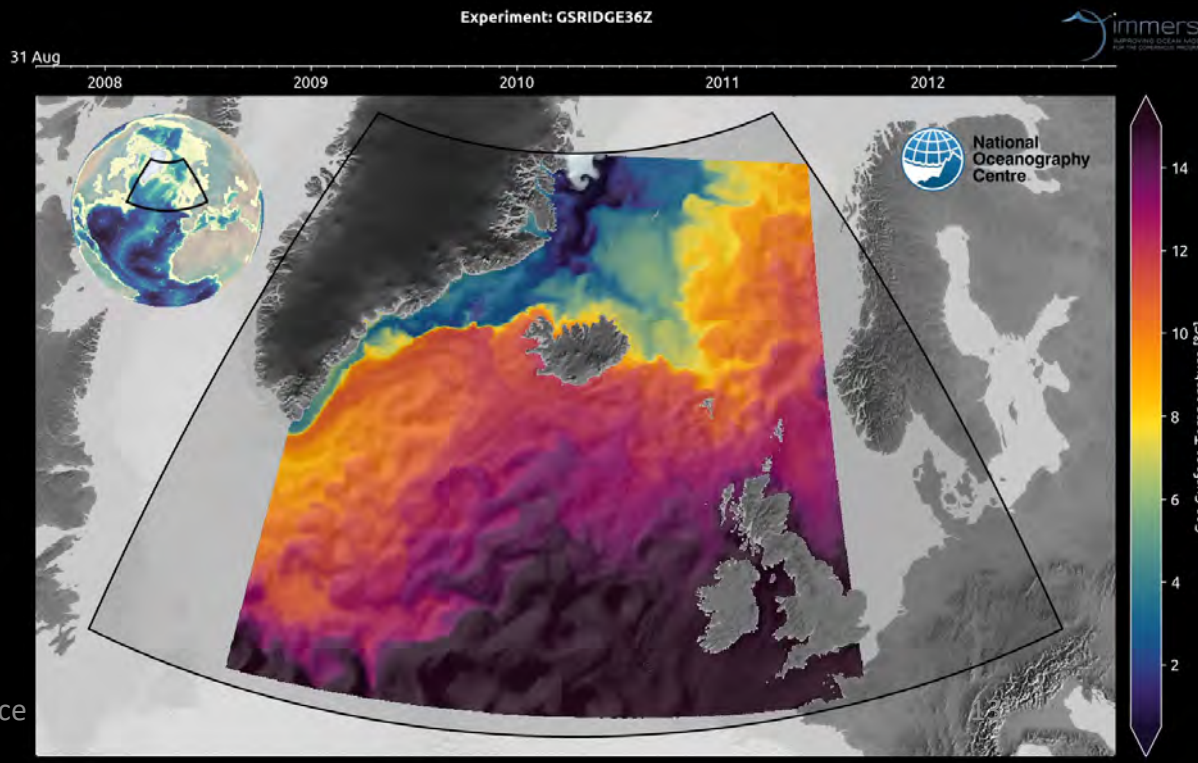




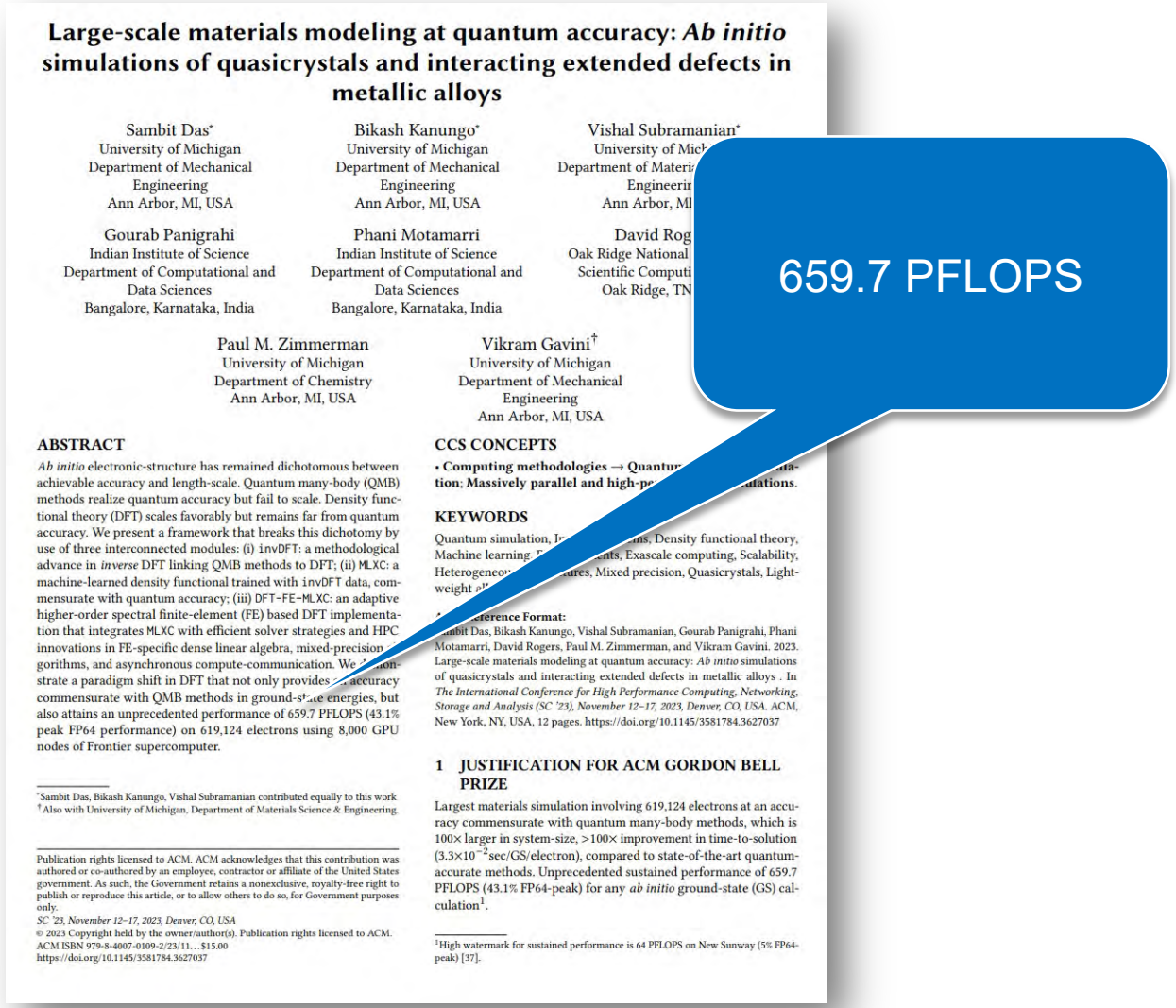







Experiment: GSRIDGE36Z
31 Aug
2008
2009
2010
2011
2012
National Oceanography Centre
14
12
10
8
6
4
2

Scientific impact worldwide



Large-scale materials modeling at quantum accuracy: *Ab initio* simulations of quasicrystals and interacting extended defects in metallic alloys

Sambit Das*
University of Michigan
Department of Mechanical Engineering
Ann Arbor, MI, USA

Bikash Kanungo*
University of Michigan
Department of Mechanical Engineering
Ann Arbor, MI, USA

Vishal Subramanian*
University of Mich
Department of Materi
Engineerir
Ann Arbor, MI

Gourab Panigrahi
Indian Institute of Science
Department of Computational and Data Sciences
Bangalore, Karnataka, India

Phani Motamarri
Indian Institute of Science
Department of Computational and Data Sciences
Bangalore, Karnataka, India

David Rog
Oak Ridge National
Scientific Computi
Oak Ridge, TN

Paul M. Zimmerman
University of Michigan
Department of Chemistry
Ann Arbor, MI, USA

Vikram Gavini†
University of Michigan
Department of Mechanical Engineering
Ann Arbor, MI, USA

ABSTRACT

Ab initio electronic-structure has remained dichotomous between achievable accuracy and length-scale. Quantum many-body (QMB) methods realize quantum accuracy but fail to scale. Density functional theory (DFT) scales favorably but remains far from quantum accuracy. We present a framework that breaks this dichotomy by use of three interconnected modules: (i) `invDFT`: a methodological advance in *inverse* DFT linking QMB methods to DFT; (ii) `MLXC`: a machine-learned density functional trained with `invDFT` data, commensurate with quantum accuracy; (iii) `DFT-FE-MLXC`: an adaptive higher-order spectral finite-element (FE) based DFT implementation that integrates `MLXC` with efficient solver strategies and HPC innovations in FE-specific dense linear algebra, mixed-precision ... gorithms, and asynchronous compute-communication. We ... strate a paradigm shift in DFT that not only provides ... accuracy commensurate with QMB methods in ground-st... energies, but also attains an unprecedented performance of 659.7 PFLOPS (43.1% peak FP64 performance) on 619,124 electrons using 8,000 GPU nodes of Frontier supercomputer.

CCS CONCEPTS

• **Computing methodologies** → **Quantu... tion**; **Massively parallel and high-p... ulations**.

KEYWORDS

Quantum simulation, I... ms, Density functional theory, Machine learning, ... nts, Exascale computing, Scalability, Heterogeneou... ures, Mixed precision, Quasicrystals, Lightweight ...

...ference Format:

...mbit Das, Bikash Kanungo, Vishal Subramanian, Gourab Panigrahi, Phani Motamarri, David Rogers, Paul M. Zimmerman, and Vikram Gavini. 2023. Large-scale materials modeling at quantum accuracy: *Ab initio* simulations of quasicrystals and interacting extended defects in metallic alloys . In *The International Conference for High Performance Computing, Networking, Storage and Analysis (SC '23), November 12–17, 2023, Denver, CO, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3581784.3627037

1 JUSTIFICATION FOR ACM GORDON BELL PRIZE

Largest materials simulation involving 619,124 electrons at an accuracy commensurate with quantum many-body methods, which is 100× larger in system-size, >100× improvement in time-to-solution (3.3×10^{-2} sec/GS/electron), compared to state-of-the-art quantum-accurate methods. Unprecedented sustained performance of 659.7 PFLOPS (43.1% FP64-peak) for any *ab initio* ground-state (GS) calculation[1].

*Sambit Das, Bikash Kanungo, Vishal Subramanian contributed equally to this work
†Also with University of Michigan, Department of Materials Science & Engineering.

[1]High watermark for sustained performance is 64 PFLOPS on New Sunway (5% FP64-peak) [37].

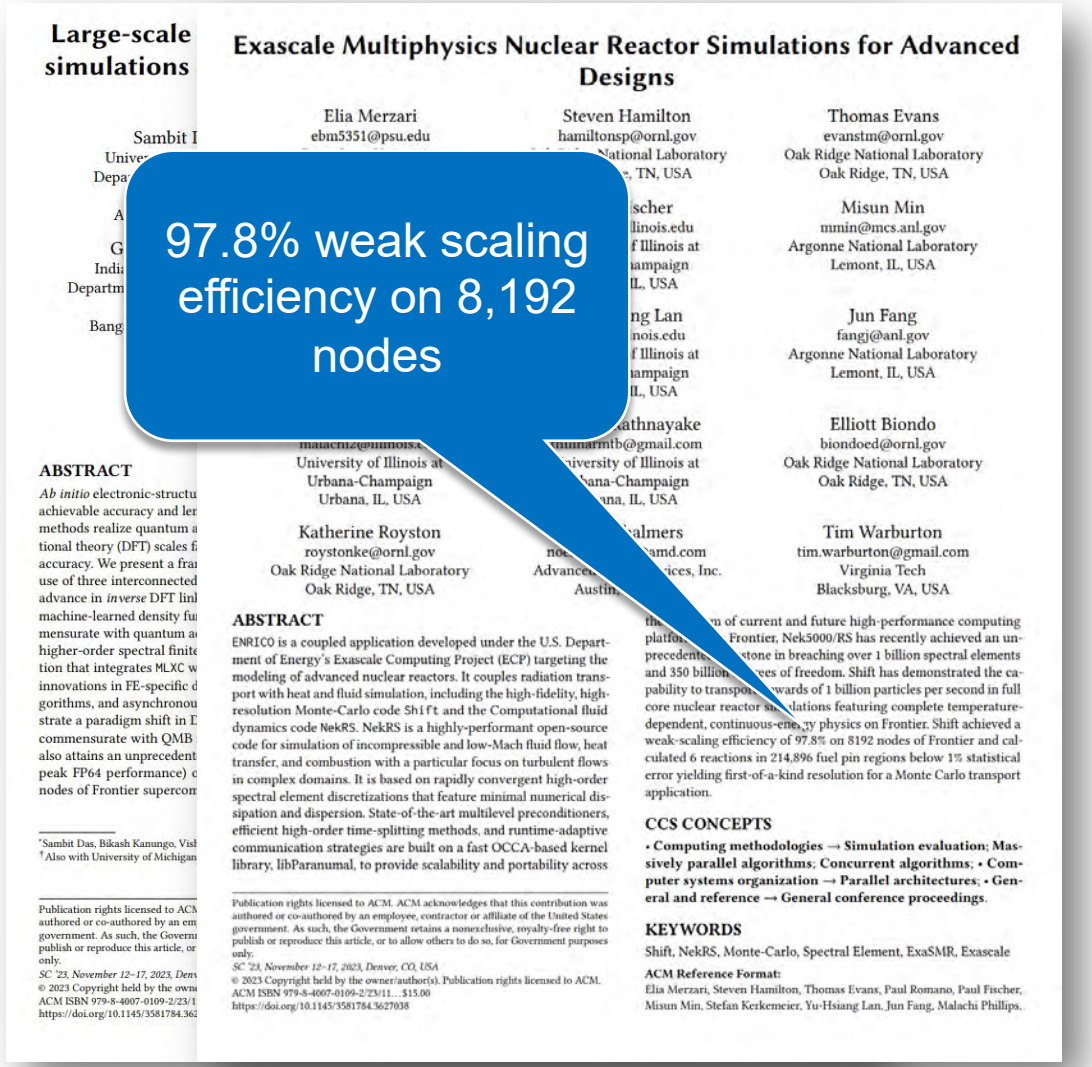
SC '23, November 12–17, 2023, Denver, CO, USA

ACM ISBN 979-8-4007-0109-2/23/11…$15.00
https://doi.org/10.1145/3581784.3627037

659.7 PFLOPS

Provide the capability and scientists will use it

Scientific impact worldwide



Large-scale simulations

Exascale Multiphysics Nuclear Reactor Simulations for Advanced Designs

97.8% weak scaling efficiency on 8,192 nodes

Elia Merzari
ebm5351@psu.edu

Steven Hamilton
hamiltonsp@ornl.gov
National Laboratory
TN, USA

Thomas Evans
evanstm@ornl.gov
Oak Ridge National Laboratory
Oak Ridge, TN, USA

Misun Min
mmin@mcs.anl.gov
Argonne National Laboratory
Lemont, IL, USA

Jun Fang
fangj@anl.gov
Argonne National Laboratory
Lemont, IL, USA

Elliott Biondo
biondoed@ornl.gov
Oak Ridge National Laboratory
Oak Ridge, TN, USA

Tim Warburton
tim.warburton@gmail.com
Virginia Tech
Blacksburg, VA, USA

Katherine Royston
roystonke@ornl.gov
Oak Ridge National Laboratory
Oak Ridge, TN, USA

University of Illinois at Urbana-Champaign
Urbana, IL, USA

ABSTRACT

ENRICO is a coupled application developed under the U.S. Department of Energy's Exascale Computing Project (ECP) targeting the modeling of advanced nuclear reactors. It couples radiation transport with heat and fluid simulation, including the high-fidelity, high-resolution Monte-Carlo code Shift and the Computational fluid dynamics code NekRS. NekRS is a highly-performant open-source code for simulation of incompressible and low-Mach fluid flow, heat transfer, and combustion with a particular focus on turbulent flows in complex domains. It is based on rapidly convergent high-order spectral element discretizations that feature minimal numerical dissipation and dispersion. State-of-the-art multilevel preconditioners, efficient high-order time-splitting methods, and runtime-adaptive communication strategies are built on a fast OCCA-based kernel library, libParanumal, to provide scalability and portability across

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of the United States government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.
SC '23, November 12–17, 2023, Denver, CO, USA
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0109-2/23/11…$15.00
https://doi.org/10.1145/3581784.3627038

CCS CONCEPTS

• Computing methodologies → Simulation evaluation; Massively parallel algorithms; Concurrent algorithms; • Computer systems organization → Parallel architectures; • General and reference → General conference proceedings.

KEYWORDS

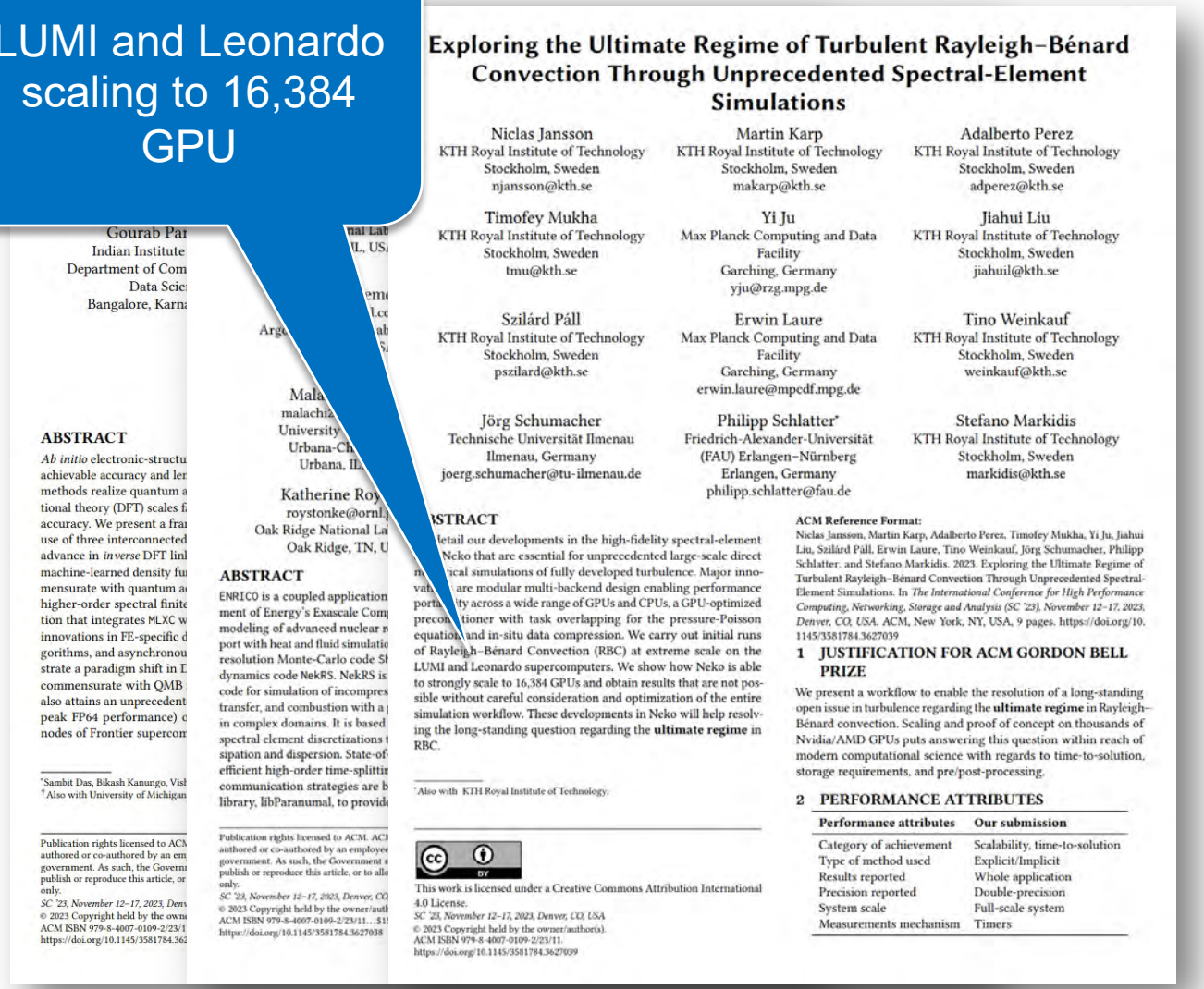
Shift, NekRS, Monte-Carlo, Spectral Element, ExaSMR, Exascale

ACM Reference Format:
Elia Merzari, Steven Hamilton, Thomas Evans, Paul Romano, Paul Fischer, Misun Min, Stefan Kerkemeier, Yu-Hsiang Lan, Jun Fang, Malachi Phillips.

Provide the capability and scientists will use it

Scientific impact worldwide

LUMI and Leonardo scaling to 16,384 GPU



Exploring the Ultimate Regime of Turbulent Rayleigh–Bénard Convection Through Unprecedented Spectral-Element Simulations

Niclas Jansson
KTH Royal Institute of Technology
Stockholm, Sweden
njansson@kth.se

Martin Karp
KTH Royal Institute of Technology
Stockholm, Sweden
makarp@kth.se

Adalberto Perez
KTH Royal Institute of Technology
Stockholm, Sweden
adperez@kth.se

Timofey Mukha
KTH Royal Institute of Technology
Stockholm, Sweden
tmu@kth.se

Yi Ju
Max Planck Computing and Data Facility
Garching, Germany
yju@rzg.mpg.de

Jiahui Liu
KTH Royal Institute of Technology
Stockholm, Sweden
jiahuil@kth.se

Szilárd Páll
KTH Royal Institute of Technology
Stockholm, Sweden
pszilard@kth.se

Erwin Laure
Max Planck Computing and Data Facility
Garching, Germany
erwin.laure@mpcdf.mpg.de

Tino Weinkauf
KTH Royal Institute of Technology
Stockholm, Sweden
weinkauf@kth.se

Jörg Schumacher
Technische Universität Ilmenau
Ilmenau, Germany
joerg.schumacher@tu-ilmenau.de

Philipp Schlatter*
Friedrich-Alexander-Universität (FAU) Erlangen–Nürnberg
Erlangen, Germany
philipp.schlatter@fau.de

Stefano Markidis
KTH Royal Institute of Technology
Stockholm, Sweden
markidis@kth.se

ABSTRACT

detail our developments in the high-fidelity spectral-element Neko that are essential for unprecedented large-scale direct ical simulations of fully developed turbulence. Major innova- are modular multi-backend design enabling performance porta ity across a wide range of GPUs and CPUs, a GPU-optimized precon tioner with task overlapping for the pressure-Poisson equation and in-situ data compression. We carry out initial runs of Rayleigh–Bénard Convection (RBC) at extreme scale on the LUMI and Leonardo supercomputers. We show how Neko is able to strongly scale to 16,384 GPUs and obtain results that are not possible without careful consideration and optimization of the entire simulation workflow. These developments in Neko will help resolving the long-standing question regarding the **ultimate regime** in RBC.

*Also with KTH Royal Institute of Technology.

SC '23, November 12–17, 2023, Denver, CO, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0109-2/23/11.
https://doi.org/10.1145/3581784.3627039

ACM Reference Format:
Niclas Jansson, Martin Karp, Adalberto Perez, Timofey Mukha, Yi Ju, Jiahui Liu, Szilárd Páll, Erwin Laure, Tino Weinkauf, Jörg Schumacher, Philipp Schlatter, and Stefano Markidis. 2023. Exploring the Ultimate Regime of Turbulent Rayleigh–Bénard Convection Through Unprecedented Spectral-Element Simulations. In *The International Conference for High Performance Computing, Networking, Storage and Analysis (SC '23), November 12–17, 2023, Denver, CO, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3581784.3627039

1 JUSTIFICATION FOR ACM GORDON BELL PRIZE

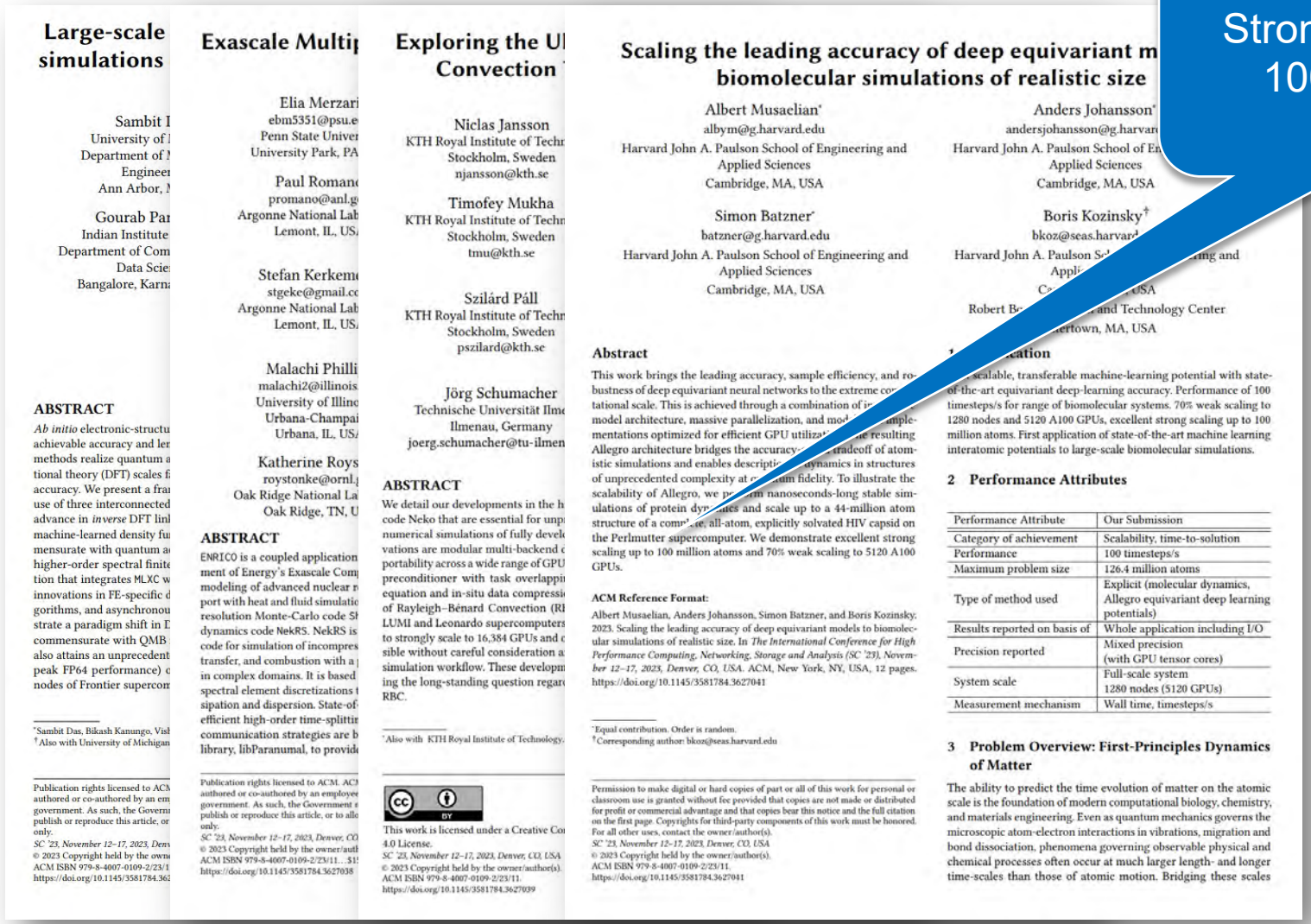
We present a workflow to enable the resolution of a long-standing open issue in turbulence regarding the **ultimate regime** in Rayleigh–Bénard convection. Scaling and proof of concept on thousands of Nvidia/AMD GPUs puts answering this question within reach of modern computational science with regards to time-to-solution, storage requirements, and pre/post-processing.

2 PERFORMANCE ATTRIBUTES

Performance attributes	Our submission
Category of achievement	Scalability, time-to-solution
Type of method used	Explicit/Implicit
Results reported	Whole application
Precision reported	Double-precision
System scale	Full-scale system
Measurements mechanism	Timers

Provide the capability and scientists will use it

Scientific impact worldwi



Large-scale simulations

Sambit I
University of
Department of
Engineer
Ann Arbor,

Gourab Par
Indian Institute
Department of Com
Data Scie
Bangalore, Karna

ABSTRACT

Ab initio electronic-structu
achievable accuracy and ler
methods realize quantum a
tional theory (DFT) scales f
accuracy. We present a fra
use of three interconnected
advance in *inverse* DFT lin
machine-learned density fu
mensurate with quantum a
higher-order spectral finite
tion that integrates MLXC w
innovations in FE-specific d
gorithms, and asynchronou
strate a paradigm shift in D
commensurate with QMB
also attains an unprecedent
peak FP64 performance) o
nodes of Frontier supercom

*Sambit Das, Bikash Kanungo, Vish
†Also with University of Michigan

Publication rights licensed to ACM
authored or co-authored by an em
government. As such, the Governm
publish or reproduce this article, or
only.
SC '23, November 12–17, 2023, Den
© 2023 Copyright held by the own
ACM ISBN 979-8-4007-0109-2/23/1
https://doi.org/10.1145/3581784.362

Exascale Multi

Elia Merzari
ebm5351@psu.e
Penn State Univer
University Park, PA

Paul Romano
promano@anl.g
Argonne National Lab
Lemont, IL, US

Stefan Kerkeme
stgeke@gmail.co
Argonne National Lab
Lemont, IL, US

Malachi Philli
malachi2@illinois
University of Illino
Urbana-Champai
Urbana, IL, USA

Katherine Roys
roystonke@ornl.
Oak Ridge National La
Oak Ridge, TN, U

ABSTRACT

ENRICO is a coupled application
ment of Energy's Exascale Comp
modeling of advanced nuclear r
port with heat and fluid simulatio
resolution Monte-Carlo code Sh
dynamics code NekRS. NekRS is
code for simulation of incompres
transfer, and combustion with a
in complex domains. It is based
spectral element discretizations t
sipation and dispersion. State-of
efficient high-order time-splittin
communication strategies are b
library, libParanumal, to provid

Publication rights licensed to ACM. AC
authored or co-authored by an employee
government. As such, the Government
publish or reproduce this article, or to all
only.
SC '23, November 12–17, 2023, Denver, CO
© 2023 Copyright held by the owner/auth
ACM ISBN 979-8-4007-0109-2/23/11…$15
https://doi.org/10.1145/3581784.3627038

Exploring the U
Convection

Niclas Jansson
KTH Royal Institute of Techn
Stockholm, Sweden
njansson@kth.se

Timofey Mukha
KTH Royal Institute of Techn
Stockholm, Sweden
tmu@kth.se

Szilárd Páll
KTH Royal Institute of Techn
Stockholm, Sweden
pszilard@kth.se

Jörg Schumacher
Technische Universität Ilme
Ilmenau, Germany
joerg.schumacher@tu-ilmen

ABSTRACT

We detail our developments in the h
code Neko that are essential for unp
numerical simulations of fully develo
vations are modular multi-backend
portability across a wide range of GPU
preconditioner with task overlappi
equation and in-situ data compressi
of Rayleigh–Bénard Convection (R
LUMI and Leonardo supercomputers
to strongly scale to 16,384 GPUs and
sible without careful consideration a
simulation workflow. These developm
ing the long-standing question regar
RBC.

*Also with KTH Royal Institute of Technology.

This work is licensed under a Creative Co
4.0 License.
SC '23, November 12–17, 2023, Denver, CO, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0109-2/23/11.
https://doi.org/10.1145/3581784.3627039

Scaling the leading accuracy of deep equivariant m
biomolecular simulations of realistic size

Albert Musaelian*
albym@g.harvard.edu
Harvard John A. Paulson School of Engineering and Applied Sciences
Cambridge, MA, USA

Anders Johansson*
andersjohansson@g.harvar
Harvard John A. Paulson School of En
Applied Sciences
Cambridge, MA, USA

Simon Batzner*
batzner@g.harvard.edu
Harvard John A. Paulson School of Engineering and Applied Sciences
Cambridge, MA, USA

Boris Kozinsky†
bkoz@seas.harvard
Harvard John A. Paulson Sc ng and
Appli
C USA
Robert Bo and Technology Center
ertown, MA, USA

Abstract

This work brings the leading accuracy, sample efficiency, and ro-
bustness of deep equivariant neural networks to the extreme co
tational scale. This is achieved through a combination of i
model architecture, massive parallelization, and mod mple-
mentations optimized for efficient GPU utiliza e resulting
Allegro architecture bridges the accuracy tradeoff of atom-
istic simulations and enables descripti dynamics in structures
of unprecedented complexity at um fidelity. To illustrate the
scalability of Allegro, we p m nanoseconds-long stable sim-
ulations of protein dy ics and scale up to a 44-million atom
structure of a comp e, all-atom, explicitly solvated HIV capsid on
the Perlmutter supercomputer. We demonstrate excellent strong
scaling up to 100 million atoms and 70% weak scaling to 5120 A100
GPUs.

ACM Reference Format:

Albert Musaelian, Anders Johansson, Simon Batzner, and Boris Kozinsky. 2023. Scaling the leading accuracy of deep equivariant models to biomolecular simulations of realistic size. In *The International Conference for High Performance Computing, Networking, Storage and Analysis (SC '23), November 12–17, 2023, Denver, CO, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3581784.3627041

*Equal contribution. Order is random.
†Corresponding author: bkoz@seas.harvard.edu

SC '23, November 12–17, 2023, Denver, CO, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0109-2/23/11.
https://doi.org/10.1145/3581784.3627041

2 Performance Attributes

Performance Attribute	Our Submission
Category of achievement	Scalability, time-to-solution
Performance	100 timesteps/s
Maximum problem size	126.4 million atoms
Type of method used	Explicit (molecular dynamics, Allegro equivariant deep learning potentials)
Results reported on basis of	Whole application including I/O
Precision reported	Mixed precision (with GPU tensor cores)
System scale	Full-scale system 1280 nodes (5120 GPUs)
Measurement mechanism	Wall time, timesteps/s

3 Problem Overview: First-Principles Dynamics of Matter

The ability to predict the time evolution of matter on the atomic scale is the foundation of modern computational biology, chemistry, and materials engineering. Even as quantum mechanics governs the microscopic atom-electron interactions in vibrations, migration and bond dissociation, phenomena governing observable physical and chemical processes often occur at much larger length- and longer time-scales than those of atomic motion. Bridging these scales

Strong scaling to 100m atoms

Provide the capability and scientists will use it

Scientific impact worldwide

Large-scale simulations

Exascale Multi

Exploring the U Convection

Scaling the lea biom

The Simple Cloud-Resolving E3SM Atmosphere Model Running on the Frontier Exascale System

Mark A. Taylor, Peter M. Caldwell, Luca Ber, Conrad Cl, Aaron S. Donahue, James G. Foucar, Oksana Guba, Benjamin R. Hillman, Noel K, Jayesh Krishna, Matthew R. Norman, Sarat Sreepathi, Christopher R. Terai, James B. White III, Danqing Wu, Andrew G. Sa, Renata B. McCoy, L. Ruby Leung, David ader

ABSTRACT

SCREAM is one of only a few global atmosphere models to be ported to GPUs. As far as we know, SCREAM is the first such model to run on both AMD GPUs and NVIDIA GPUs, as well as the first to run on nearly an entire Exascale system (Frontier). On Frontier, we obtained a record setting performance of 1.26 simulated years per day for a realistic cloud resolving simulation.

CCS CONCEPTS

• Applied computing → Earth and atmospheric sciences.

KEYWORDS

atmospheric modeling, global cloud resolving, high performance computing, exascale, GPU

First global atmosphere model to run on full Exascale system

Provide the capability and scientists will use it

Scientific impact worldwide

Large-scale simulations

Sambit I
University of
Department of
Engineer
Ann Arbor,

Gourab Pa
Indian Institute
Department of Com
Data Scie
Bangalore, Karna

ABSTRACT

Ab initio electronic-structu
achievable accuracy and le
methods realize quantum a
tional theory (DFT) scales f
accuracy. We present a fra
use of three interconnected
advance in *inverse* DFT lin
machine-learned density fu
mensurate with quantum a
higher-order spectral finite
tion that integrates MLXC w
innovations in FE-specific d
gorithms, and asynchronou
strate a paradigm shift in D
commensurate with QMB
also attains an unprecedent
peak FP64 performance) o
nodes of Frontier supercom

*Sambit Das, Bikash Kanungo, Vish
†Also with University of Michigan

Publication rights licensed to ACM. AC
authored or co-authored by an em
government. As such, the Governm
publish or reproduce this article, or
only.
SC '23, November 12–17, 2023, Den
© 2023 Copyright held by the owne
ACM ISBN 979-8-4007-0109-2/23/1
https://doi.org/10.1145/3581784.362

Exascale Multi

Elia Merzari
ebm5351@psu.e
Penn State Univer
University Park, PA

Paul Romano
promano@anl.g
Argonne National Lab
Lemont, IL, US

Stefan Kerkem
stgeke@gmail.co
Argonne National Lab
Lemont, IL, US

Malachi Philli
malachi2@illinois.
University of Illino
Urbana-Champai
Urbana, IL, USA

Katherine Roys
roystonke@ornl.
Oak Ridge National La
Oak Ridge, TN, U

ABSTRACT

ENRICO is a coupled application
ment of Energy's Exascale Comp
modeling of advanced nuclear r
port with heat and fluid simulatio
resolution Monte-Carlo code Sh
dynamics code NekRS. NekRS is
code for simulation of incompres
transfer, and combustion with a
in complex domains. It is based
spectral element discretizations t
sipation and dispersion. State-of
efficient high-order time-splittin
communication strategies are b
library, libParanumal, to provid

Publication rights licensed to ACM. AC
authored or co-authored by an employee
government. As such, the Government
publish or reproduce this article, or to all
only.
SC '23, November 12–17, 2023, Denver, CO
© 2023 Copyright held by the owner/auth
ACM ISBN 979-8-4007-0109-2/23/11…$15
https://doi.org/10.1145/3581784.3627038

Exploring the U Convection

Niclas Jansson
KTH Royal Institute of Techn
Stockholm, Sweden
njansson@kth.se

Timofey Mukha
KTH Royal Institute of Techn
Stockholm, Sweden
tmu@kth.se

Szilárd Páll
KTH Royal Institute of Techn
Stockholm, Sweden
pszilard@kth.se

Jörg Schumacher
Technische Universität Ilme
Ilmenau, Germany
joerg.schumacher@tu-ilmen

ABSTRACT

We detail our developments in the h
code Neko that are essential for unpr
numerical simulations of fully develo
vations are modular multi-backend c
portability across a wide range of GPU
preconditioner with task overlappi
equation and in-situ data compressi
of Rayleigh–Bénard Convection (RB
LUMI and Leonardo supercomputers
to strongly scale to 16,384 GPUs and c
sible without careful consideration a
simulation workflow. These developm
ing the long-standing question regar
RBC.

*Also with KTH Royal Institute of Technology.

SC '23, November 12–17, 2023, Denver, CO, USA

ACM ISBN 979-8-4007-0109-2/23/11.
https://doi.org/10.1145/3581784.3627039

Scaling the lea biom

Albert Mu
albym@g.ha
Harvard John A. Paulson Sc
Applied S
Cambridge,

Simon Ba
batzner@g.ha
Harvard John A. Paulson Sc
Applied S
Cambridge,

Abstract

This work brings the leading accura
bustness of deep equivariant neural ne
tational scale. This is achieved through
model architecture, massive parallel
mentations optimized for efficient G
Allegro architecture bridges the accu
istic simulations and enables descript
of unprecedented complexity at quan
scalability of Allegro, we perform n
ulations of protein dynamics and s
structure of a complete, all-atom, exp
the Perlmutter supercomputer. We d
scaling up to 100 million atoms and 7
GPUs.

ACM Reference Format:
Albert Musaelian, Anders Johansson, Sim
2023. Scaling the leading accuracy of deep
ular simulations of realistic size. In *The I
Performance Computing, Networking, Stor
ber 12–17, 2023, Denver, CO, USA.* ACM,
https://doi.org/10.1145/3581784.3627041

*Equal contribution. Order is random.
†Corresponding author: bkoz@seas.harvard.edu

SC '23, November 12–17, 2023, Denver, CO, USA

ACM ISBN 979-8-4007-0109-2/23/11.
https://doi.org/10.1145/3581784.3627041

The Simple Clou

Mark A. Taylor
mataylo@sandia.gov
Sandia National Laborat
Albuquerque, NM, US

Aaron S. Donahue
donahue5@llnl.gov
Lawrence Livermore Nati
Laboratory
Livermore, CA, USA

Jayesh Krishna
jayesh@anl.gov
Argonne National Labora
Lemont, IL, USA

James B. White III
trey.white@hpe.com
Hewlett Packard Enterp
San Diego, CA, USA

Renata B. McCoy
mccoy20@llnl.gov
Lawrence Livermore Nati
Laboratory
Livermore, CA, USA

ABSTRACT

We present an efficient and perform
of the Simple Cloud Resolving E3SM
SCREAM is a full featured atmosph
with a nonhydrostatic dynamical co
eterizations for microphysics, mois
has been written from scratch in C++
to abstract the on-node execution m

SC '23, November 12–17, 2023, Denver, CO, USA

ACM ISBN 979-8-4007-0109-2/23/11.
https://doi.org/10.1145/3581784.3627044

Digital transformation of droplet/aerosol infection risk assessment realized on "Fugaku" for the fight against COVID-19

Supercomputing '21, November 14–19, 2021
XX(X):1–14
©The Author(s) 2021
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Kazuto Ando*[1], Rahul Bale*[1], ChungGang Li*[12], Satoshi Matsuoka*[13], Keiji Onishi*[1], and Makoto Tsubokura*[12]

Abstract

The fastest supercomputer in 2020, Fugaku, has not only achieved digital transformation of epidemiology in allowing end-to-end, detailed quantitative modeling of COVID-19 transmissions for the first time, but also transformed the behavior of the entire Japanese public through its detailed analysis of transmission risks in multitudes of societal situations entailing heavy risks. A novel aerosol simulation methodology was synthesized out of a combination of a new CFD methods meeting industrial demands, CUBE[1], which not only allowed the simulations to scale massively with high resolution required for micrometer virus-containing aerosol particles, but also extremely rapid time-to-solution due to its ability to generate the digital twins representing multitudes of societal situations in minutes not week, attaining true overall application high performance; such simulations have been running for the past 1.5 years on Fugaku, cumulatively consuming top supercomputer-class resources and the result communicated by the media as well as becoming official public policies.

Keywords

COVID-19, Computational fluid dynamics, Building cube method, Immersed boundary method, Dirty CAD, Droplet/Aerosol transmission, societal behavioral change

1 Justification for ACM Gordon Bell Special Prize for HPC-Based COVID-19 Research

Our work best fulfills the criteria of the ACM Gordon Bell Prize on three pillars of awarding metrics: *performance*, in that whole machine scalability and efficiency as well as orders of magnitude speedup in digital twin mesh generation was achieved, *technology*, in that a these being the result of the new CFD method combined with Fugaku/A64FX's HPC application-centric design, and *science and societal impact* in achieving transformational epidemiological simulation of droplets/aerosols and impacted the entire Japanese society's behavior to curtail the COVID-19 pandemic.

2 Performance Attributes

Performance attribute	Our submission
Category of achievement	Time-to-solution, Scalability
Type of method used	Both explicit and implicit
Results reported on basis of	Whole application except I/O
Precision reported	Double precision
System scale	Measured on full system
Measurements	Hardware performance counter

3 Overview of the Problem

3.1 COVID-19 Droplet/Aerosol Infection

COVID-19, initially discovered at the end of 2019, quickly spread globally and changed our lives—the main question of this unknown virus was its main mode of transmission. In particular, in the early stages of the pandemic, various theories existed, some extrapolated from traditional epidemiological observations, but turning out to be somewhat incorrect and/or on shaky scientific grounds, and such information coming from seemingly authoritative sources significantly disrupted the socio-economic activities due to lockdown etc. One might still recall that, in the early stages of the pandemic, even institutions such as the WHO as well as the US CDC gave somewhat skeptical views regarding the effectiveness of commercial surgical masks as preventive measures, which might have misdirected the behavior of individuals causing pandemic to worsen.

A month after WHO declared the COVID-19 pandemic in March 2020, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and RIKEN Center for Computational Science (R-CCS), jointly announced a program to rapidly exploit the computational capability of the new Fugaku supercomputer, which was still in the early days of its installation, to combat COVID-19. As Fugaku would get deployed, the COVID-19 applications would receive priority status with both resources and support from R-CCS and Fujitsu, the design and manufacturing partner; in the end, *each* of the project will be granted computing resources equalling the entire dominance of a top-tier class supercomputer, i.e., equivalent to tens of millions of node

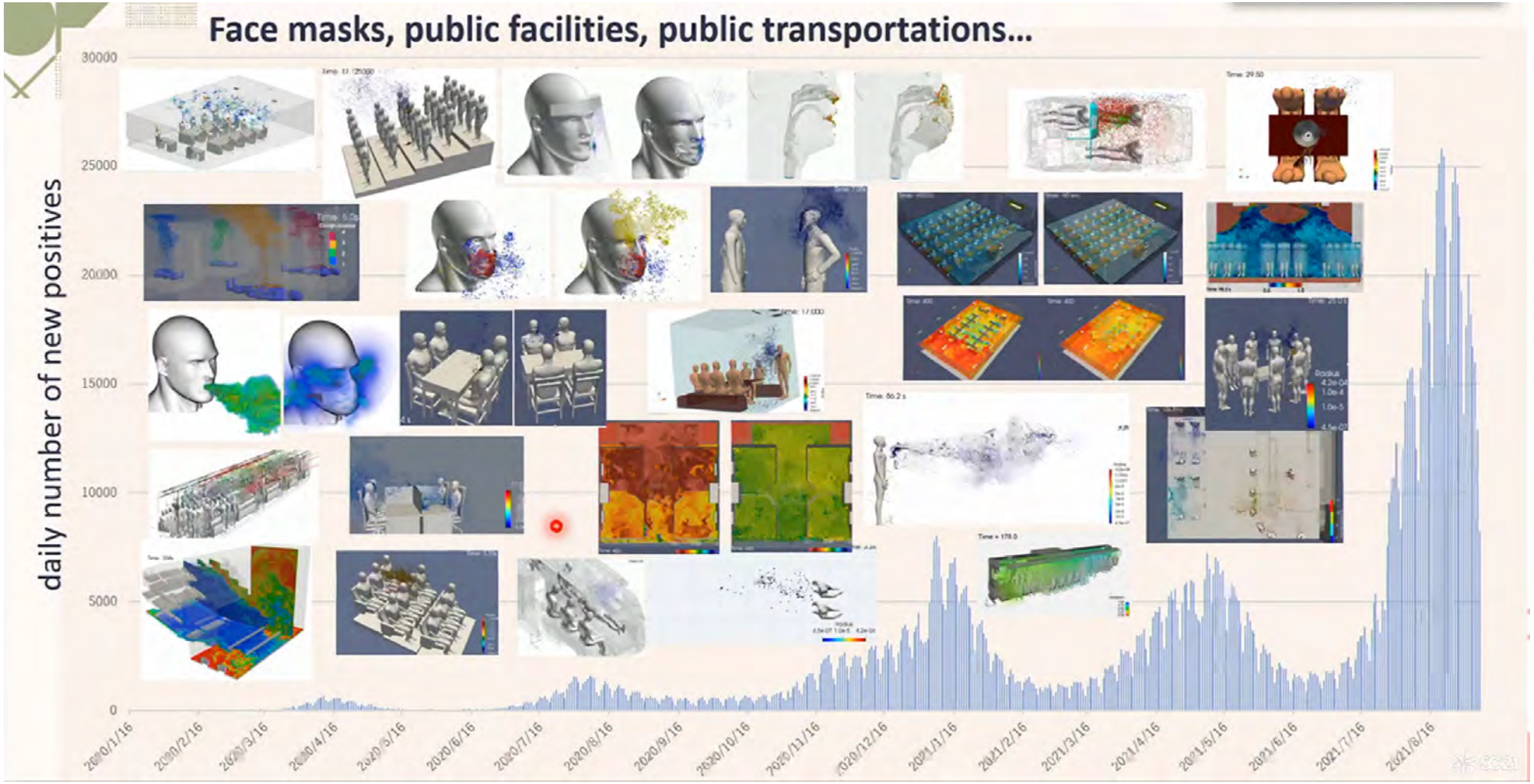
[1]RIKEN Center for Computational Science (R-CCS), Japan
[2]Kobe University, Japan
[3]Tokyo Institute of Technology, Japan
*All authors are listed in alphabetical order by surnames.

Corresponding author:
Makoto Tsubokura, Complex Phenomena Unified Simulation Research Team, RIKEN Center for Computational Science (R-CCS), Kobe, Japan.
Email: mtsubo@riken.jp

Prepared using sagej.cls *[Version: 2017/01/17 v1.20]*

Provide the capability and scientists will use it

Example: Digital Transformation of Droplet/Aerosol Infection Risk Assessment Realised on Fugaku for the Fight against COVID-19



From SC21 Gordon Bell Prize Presentation

UK Research and Innovation

Planned application software programme

- UKRI delivered software programme for Exascale
  - Will benefit many parts of computational science ecosystem
- Outline plan - mixture of
  - Expanded eCSE programme – eCSE++
  - Longer software development projects for specific applications development – new applications / rewritten applications
  - Challenge driven software development – to put UK computational science at forefront internationally
- Building on ExCALIBUR activities

Summary

- The UK Exascale Project is moving forward
- The project is not about buying a computer … it's about the **science and innovation benefits** that computer will bring
- A phased approach is being taken with full Exascale service planned by 2026/27
- **Phase 1** expected to be **installed in late 2025** – it will have around 1,500 nodes and 6,000 CPU/GPUs
- ACF is ready to host Phase 1 and preparing for Phase 2

CIUK 2023 Presentations

Sadaf Alam (University of Bristol)

Isambard AI - a National AI Research Infrastructure

Abstract: The talk overviews Isambard AI digital research infrastructure (DRI), a national AI research resource, including unique capabilities of its hardware and software stacks, a highly sustainable and energy efficient modular data centre solution, and cloud-native user interfaces to support contemporary AI frameworks and to improve productivity of AI, ML and HPC workflows.

Bio: Dr Sadaf R. Alam is the University of Bristol's Director of Advanced Computing Strategy. Sadaf joined Bristol University in 2022 from the Swiss National Supercomputing Centre (CSCS) where she was the Chief Technology Officer (CTO). Dr. Alam studied computer science at the University of Edinburgh, UK, where she received her Ph.D. Until March 2009, she was a computer scientist at the Oak Ridge National Laboratory, USA.

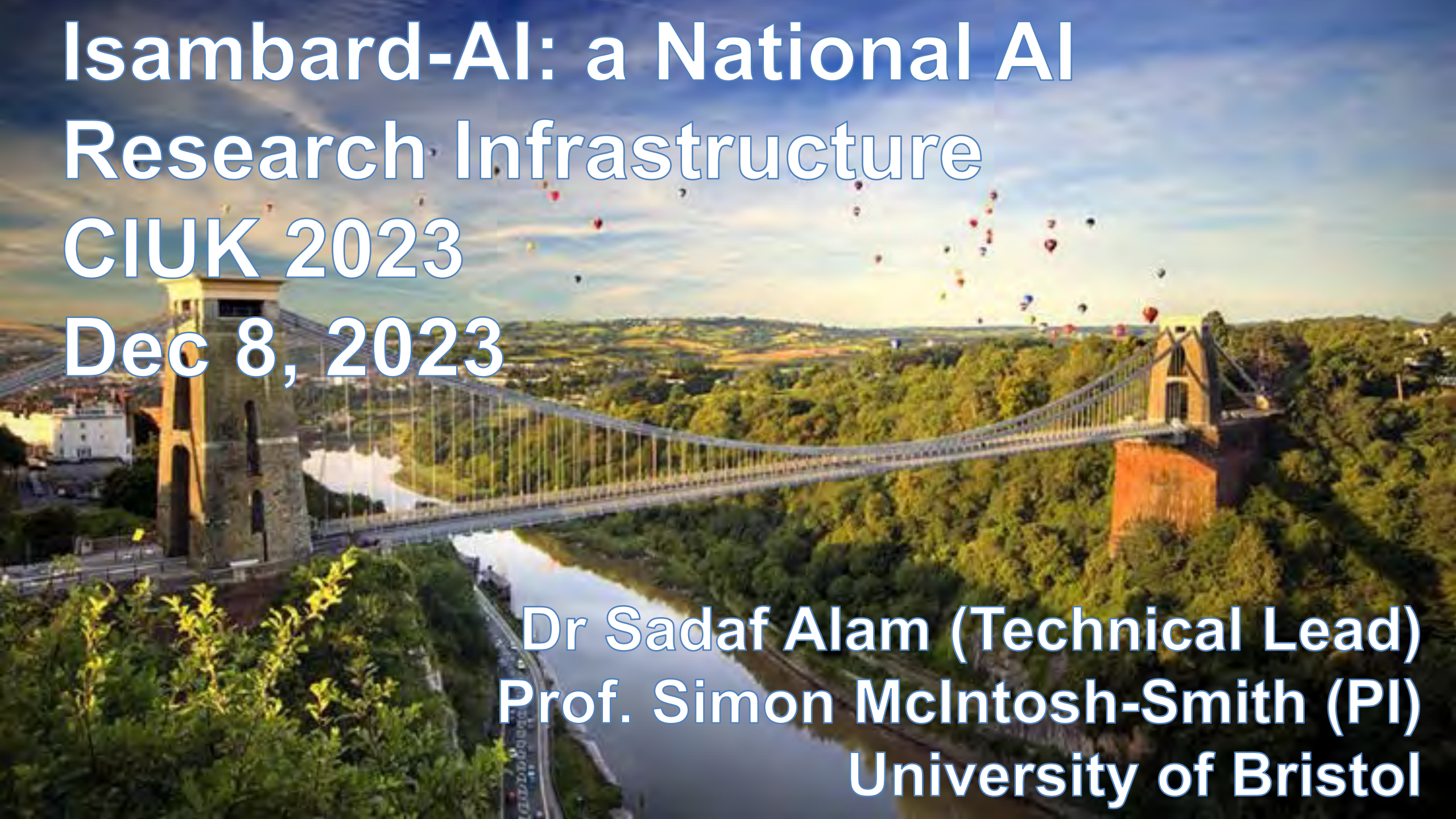
Sadaf ensures end-to-end integrity of HPC systems and storage solutions and leads strategic projects at the centre. She has held several different roles across her career including group lead of future systems, chief architect and head of operations. She is a member of ACM, ACM-W, SIGHPC and Women in HPC, and was the technical chair of the world Supercomputing conference SC22.

Sadaf was the chief architect of multiple generations of Piz Daint supercomputing platforms, which is one of Europe’s fastest and among the top 3 supercomputers in the world for many years, and also chief architect of the MeteoSwiss innovative, co-designed operational numerical weather forecasting platforms.

Isambard-AI: a National AI Research Infrastructure

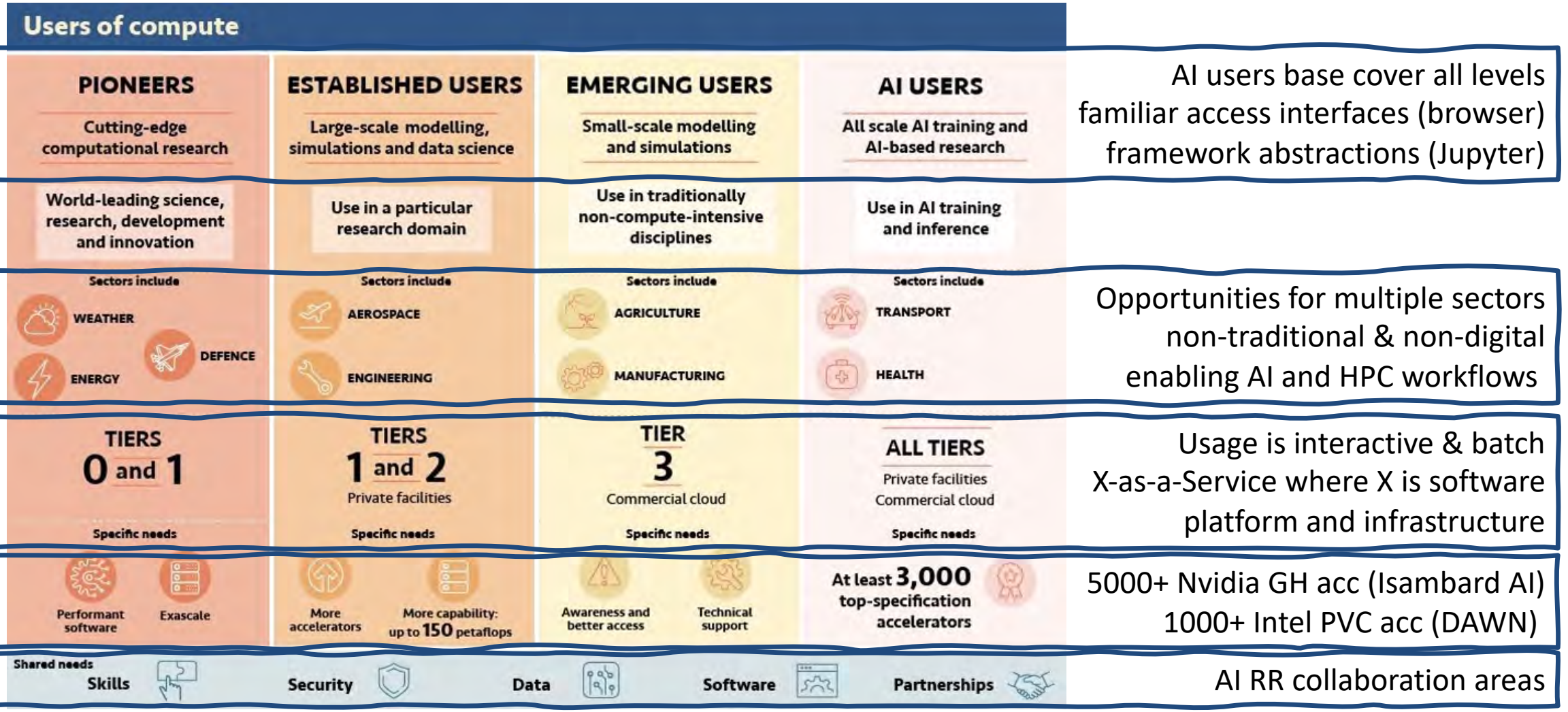
CIUK 2023

Dec 8, 2023



Dr Sadaf Alam (Technical Lead)
Prof. Simon McIntosh-Smith (PI)
University of Bristol

Design Specifications for AI Research Resource (RR)



University of BRISTOL

References: Independent Review of The Future of Compute: Final report and recommendations, March 2023; National AI Strategy - AI Action Plan, July 2022; £300 million to launch first phase of new AI Research Resource

AI RR Collaboration Workgroups (Isambard AI and DAWN)

Cybersecurity for Digital Research Infrastructure

Federated Identity and Access Management

AI and ML Environments and Frameworks

Data motion and management

User support, training and outreach

Nov 27-28 kick-off meeting between Bristol and Cambridge
Established collaborative space
Excalibur proposal for federation and linking of two AI RRs

University of BRISTOL

Anatomy of a national AI Research Resource (AIRR)

- Access methods as similar as possible to existing resources
  - E.g. cloud-style, Jupyter notebooks, as well as HPC style, e.g. ssh & batch
- Fully **federated**, with support for true **multi-tenancy**
- Highly scalable resource, providing from 1 to 1,000s of GPUs
- Flexible, fast storage solution, optimized both for high IOPs and small file accesses, as well as bulk parallel file access for BW
- System architecture optimized for a wide-range of AI workloads:
  - Training, including next-generation VLLMs
  - Inference
  - Hybrid workflows, including AI+HPC

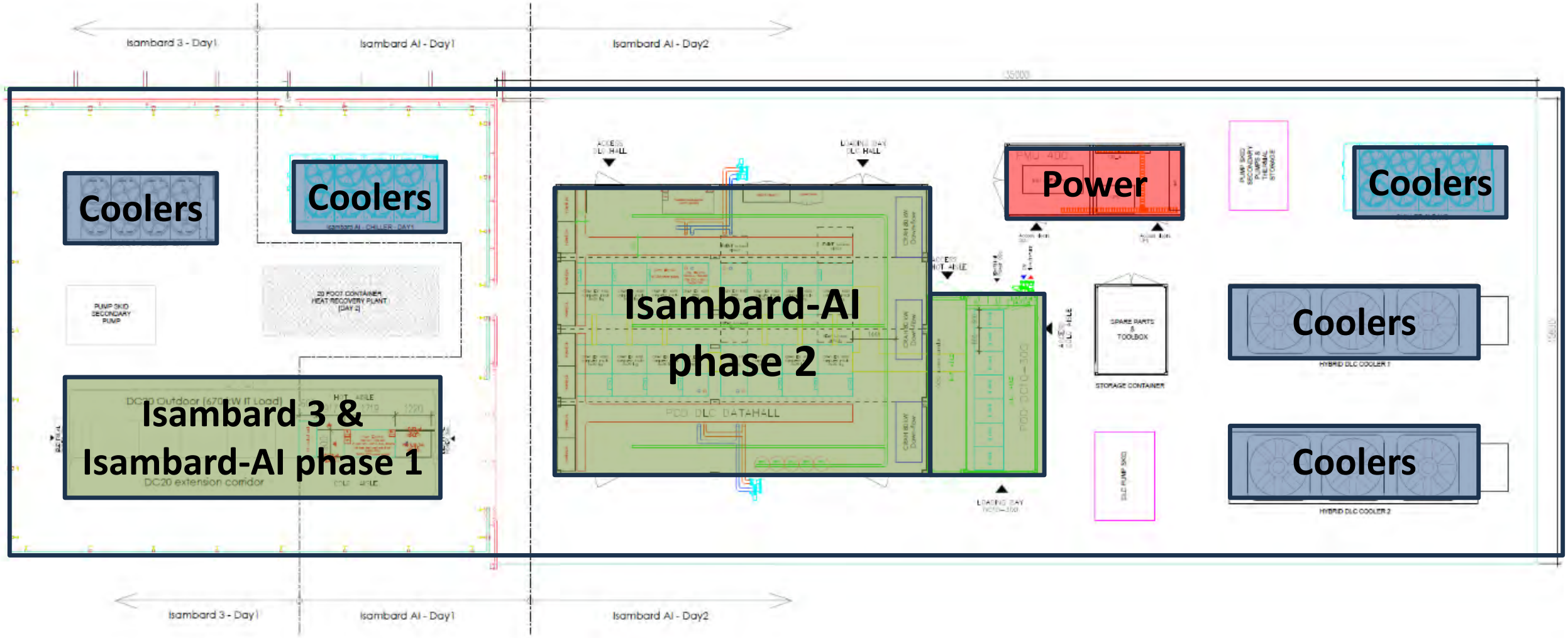
Isambard-AI: a national AI research infrastructure

- Isambard-AI will form the main part of the UK's national AI Research Resource (AIRR), over £300M investment in total
- New GPU system to be added alongside Isambard 3
- 5,448 NVIDIA GH200 Grace-Hopper GPUs
  - >21 ExaFLOPs for AI (8-bit), >200 PetaFLOPs for HPC
- Comfortably in top 5 fastest open AI systems, top 10 for HPC globally
- Large, fast storage system, all-flash (~25 PB)
- Software stack optimized for AI workflows and cloud-style usage
- ~5MW operating power, direct liquid cooling with heat reuse
- Modular Data Centre (MDC) technology for efficient deployment

University of BRISTOL

https://www.gov.uk/government/news/bristol-set-to-host-uks-most-powerful-supercomputer-to-turbocharge-ai-innovation


BELL
oakland
WC



University of BRISTOL

- An HPE EX2500 system
- We install one of these, with 168 GPUs in it, in Isambard 3 in March next year
- Early access from May
- Equivalent to about 500 NVIDIA A100 GPUs for AI
- Picture from IEEE/ACM SuperComputing in Denver, November 2023 (SC23)

University of BRISTOL

- Sadaf and Simon visited the HPE booth at SC23 last week to see the Isambard-AI hardware
- The HPE EX4000 main system
- We deploy 12 of these, each with 440 GPUs in them (5,280 in total), in the main Isambard-AI POD in Jul-Aug 2024
- Picture from IEEE/ACM SuperComputing in Denver, November 2023 (SC23)

University of BRISTOL

Summary

- The next-generation of flexible, user-oriented AI services is coming
- Optimised for LLM development, training, inference, and hybrid
- ~5,500 latest generation NVIDIA GPUs
- Fast, multi-modal storage, all solid-state
- True multi-tenancy support
- Designed to be able to evolve the software environment over time
- Ultra **energy efficient to** meet NetZero goals
- Enabling next-generation AI sciences

University of BRISTOL

Call to action

- We'll have one of the world's fastest, most advanced AI supercomputers in Bristol from next summer
- What does this make possible?
- What could you use this for which was previously unachievable?

- We'll be hiring a world-leading support team throughout 2024 to help run Isambard-AI – let us know if you're interested!

University of BRISTOL


UKRI
Science and Technology Facilities Council
COMPUTING INSIGHT UK 2023
"Productive Supercomputing"
WEDNESDAY 6 DECEMBER
Manchester Central, UK
www.ukri.org/CIUK
SYCL™
CIUK ZERO

Instructors

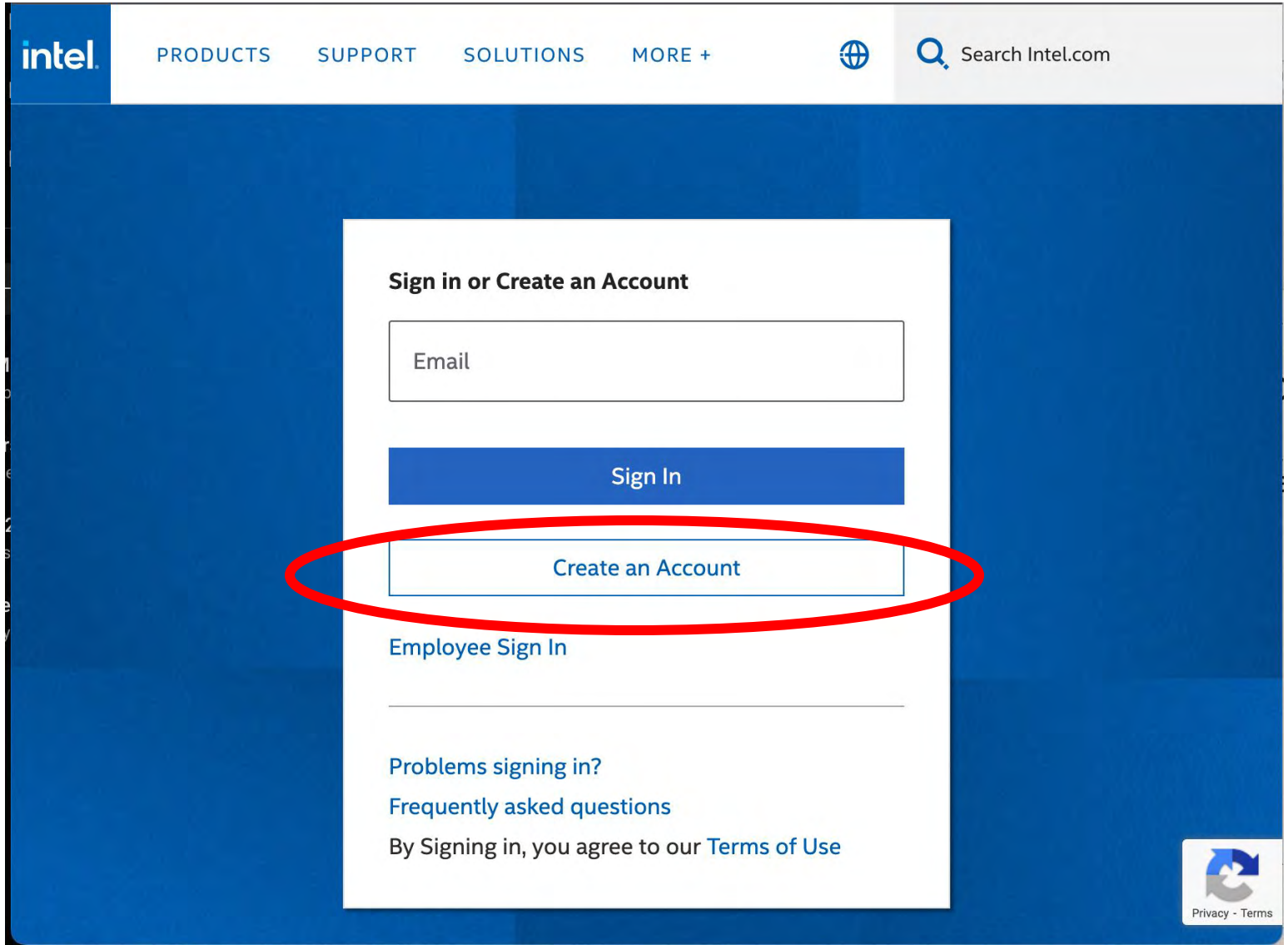
- Chris Edsall

- Fergus Baker

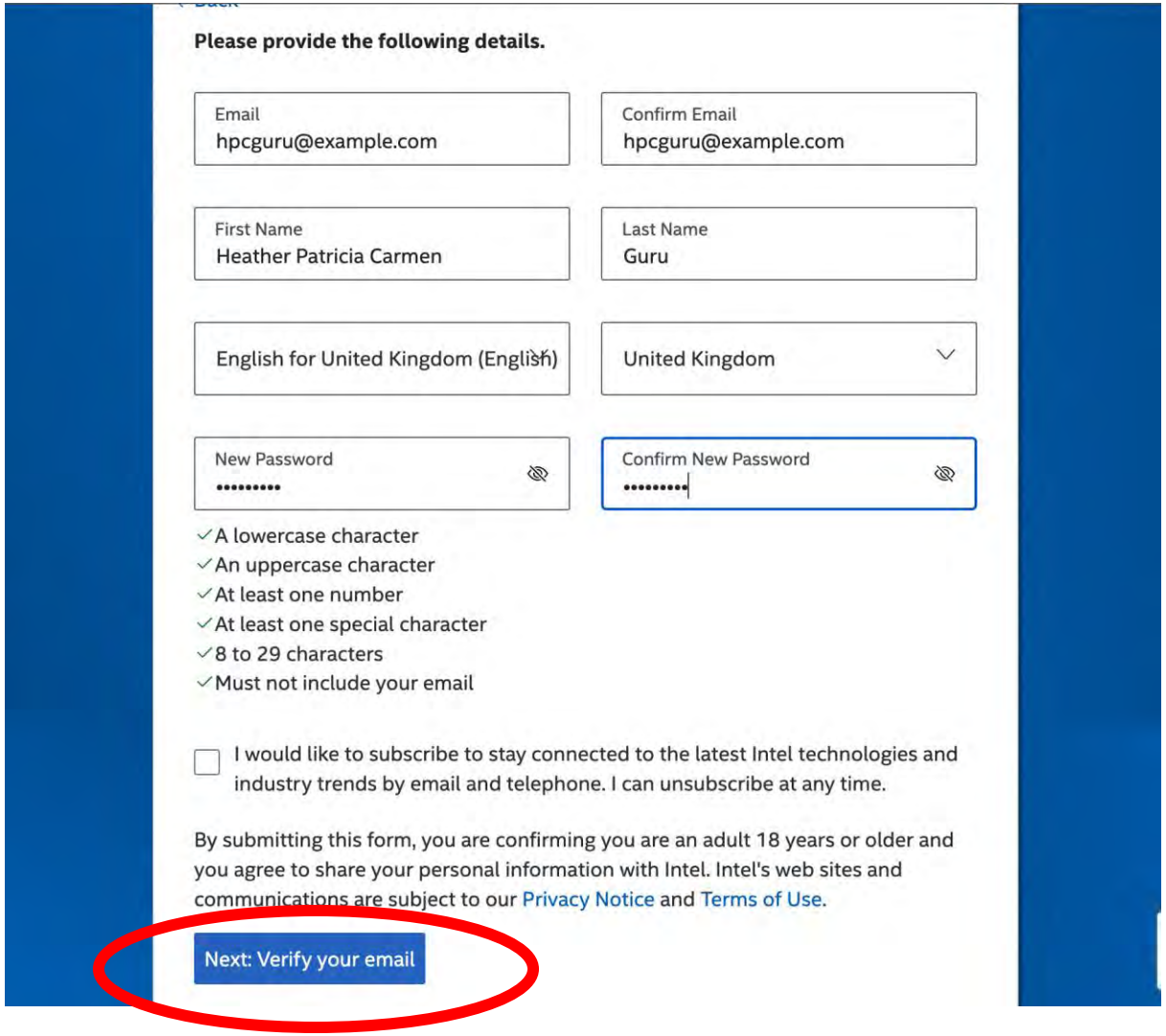
- Miren Radia

console.cloud.intel.com

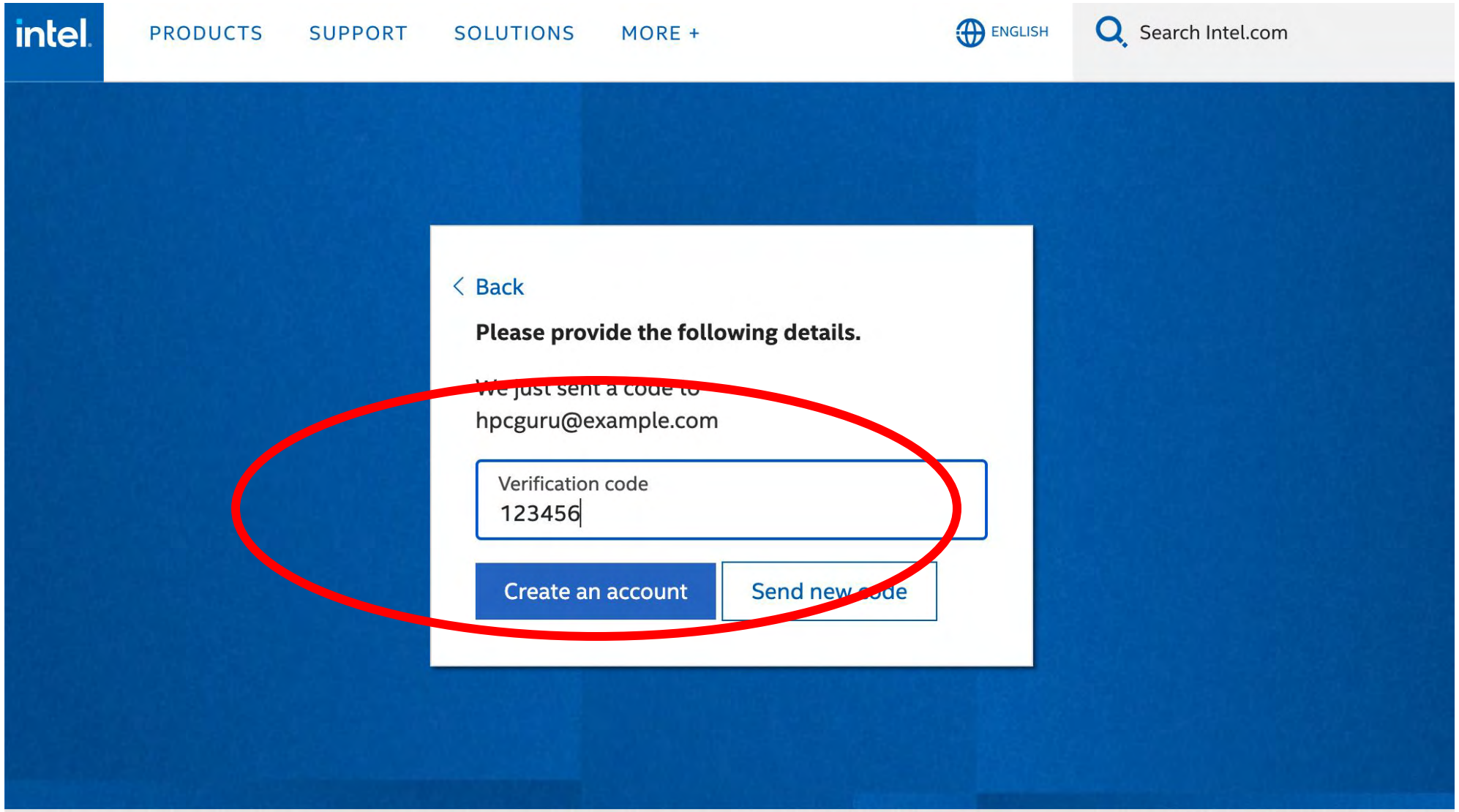
console.cloud.intel.com



console.cloud.intel.com



console.cloud.intel.com



console.cloud.intel.com

Register to access this content

Please sign up for Intel® Developer Cloud, choose a tier to get started:

Standard - Free

Explore and evaluate the latest Intel® AI products, get support from the Intel community:

- Develop AI skills.
- Access cutting edge learning resources.
- Get support from the Intel community.

Choose Standard

Premium - Pay as You Go

Get single-user access to the latest Intel products, Standard tier benefits, plus premium support:

- Early prerelease hardware.
- AI and machine learning software toolkits.
- Intel premium support.

Choose Premium

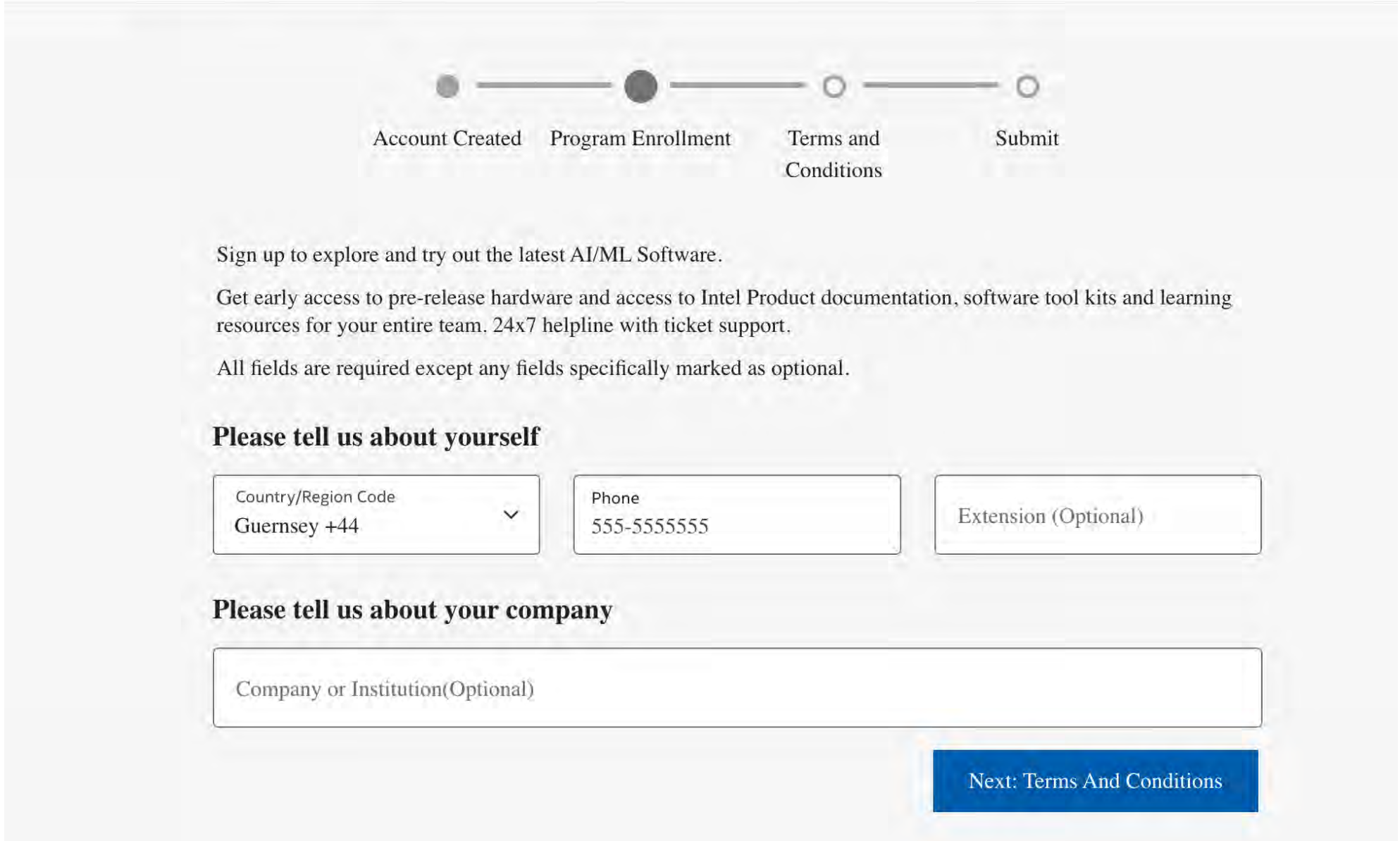
Enteprise - Team Access

In addition to Premium level benefits, an Enterprise subscription includes discounts and team access to the latest Intel products:

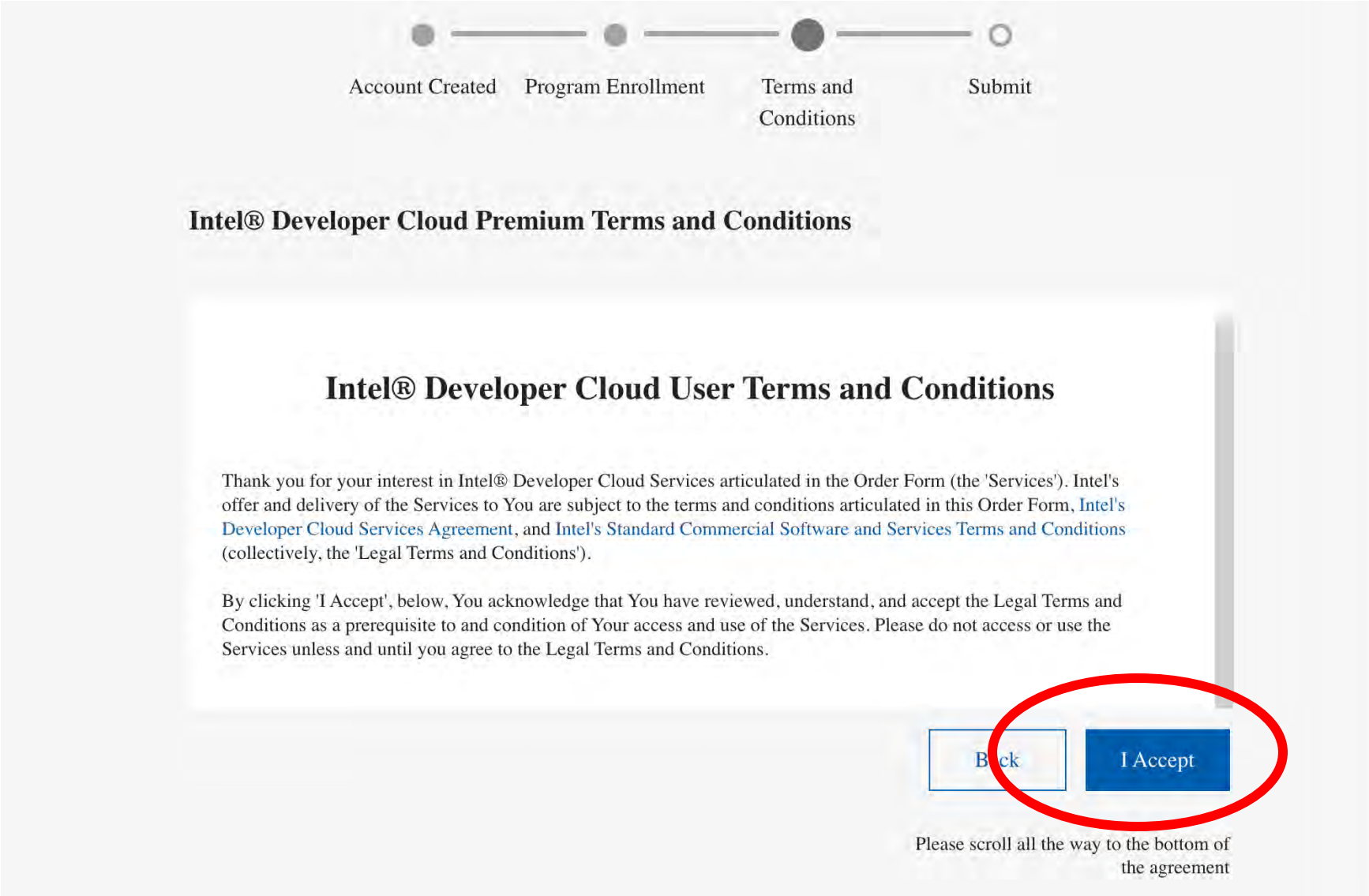
- Billed subscription for teams.
- Deploy inference on your own infrastructure.
- Use CPU, GPU, and AI accelerators.
- 24 x 7 Intel premium plus support.

Choose Enterprise

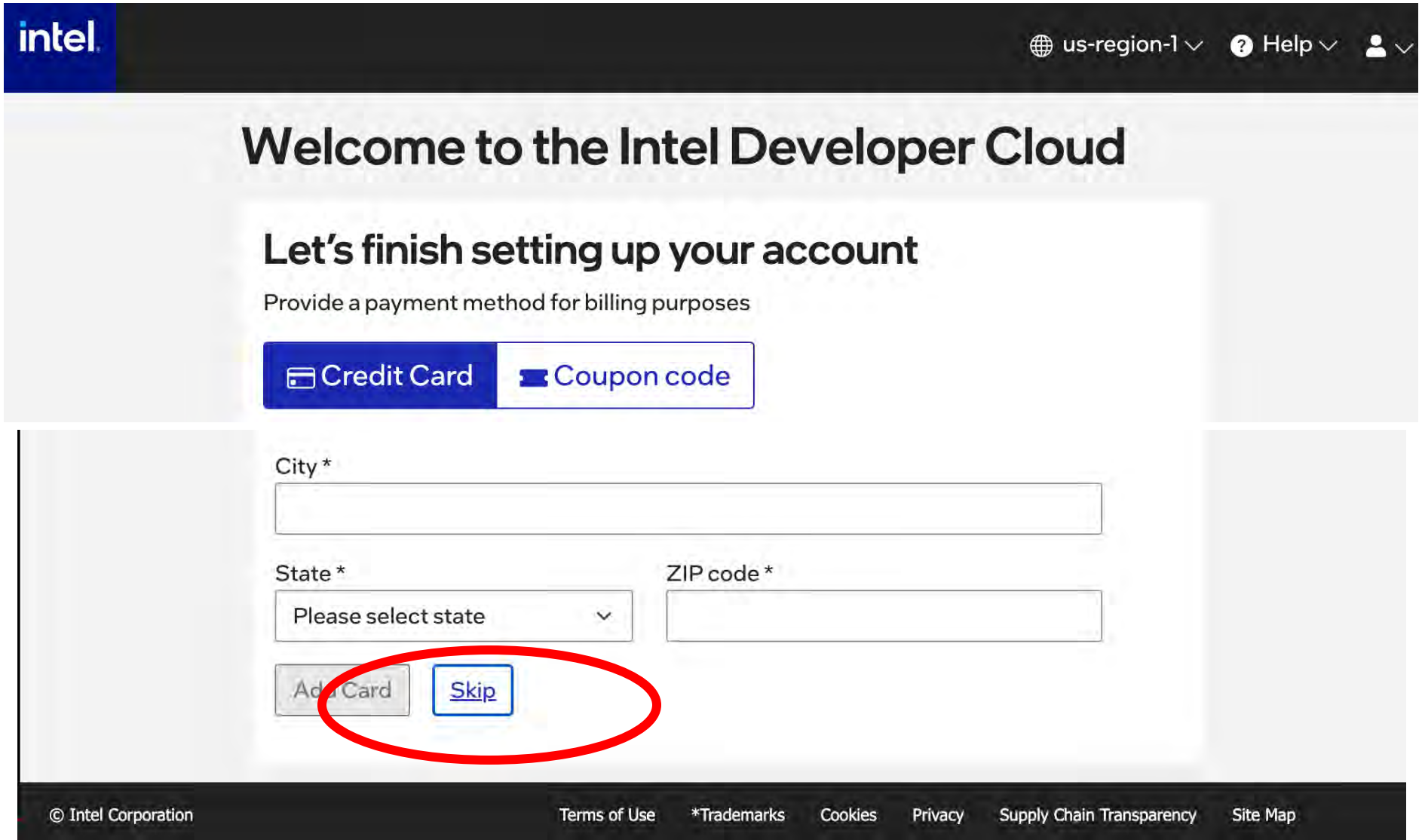
console.cloud.intel.com



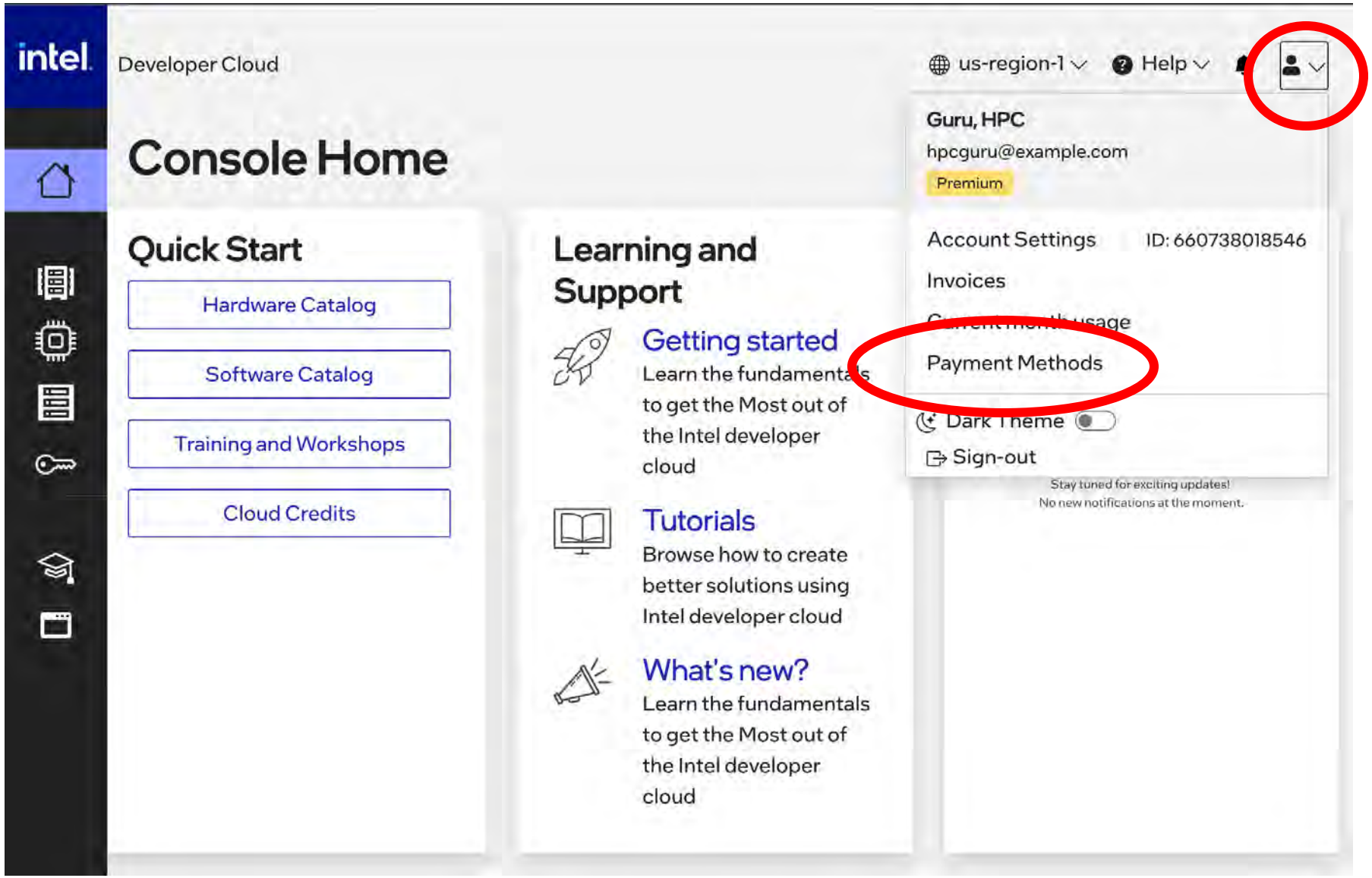
console.cloud.intel.com



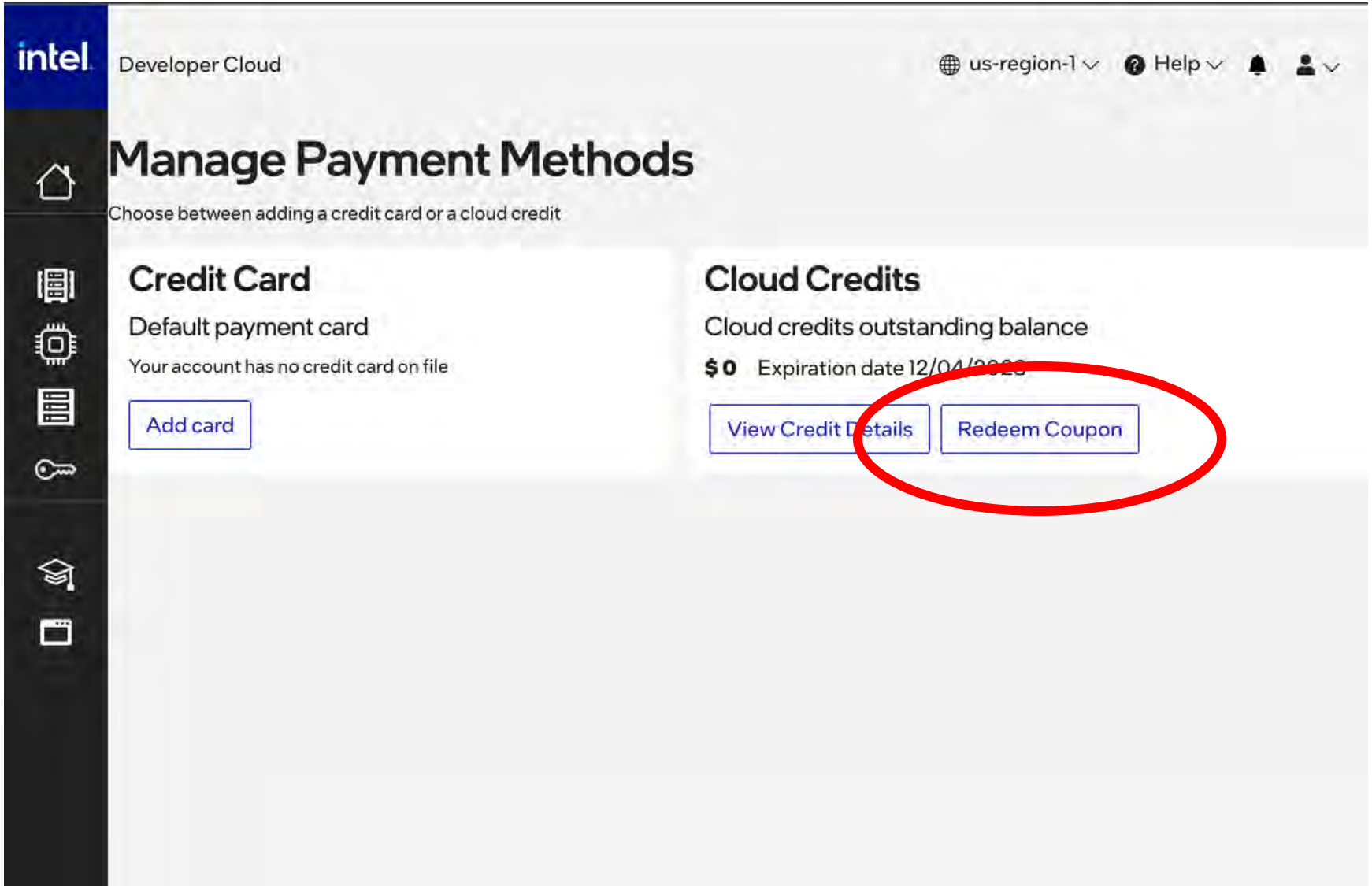
console.cloud.intel.com



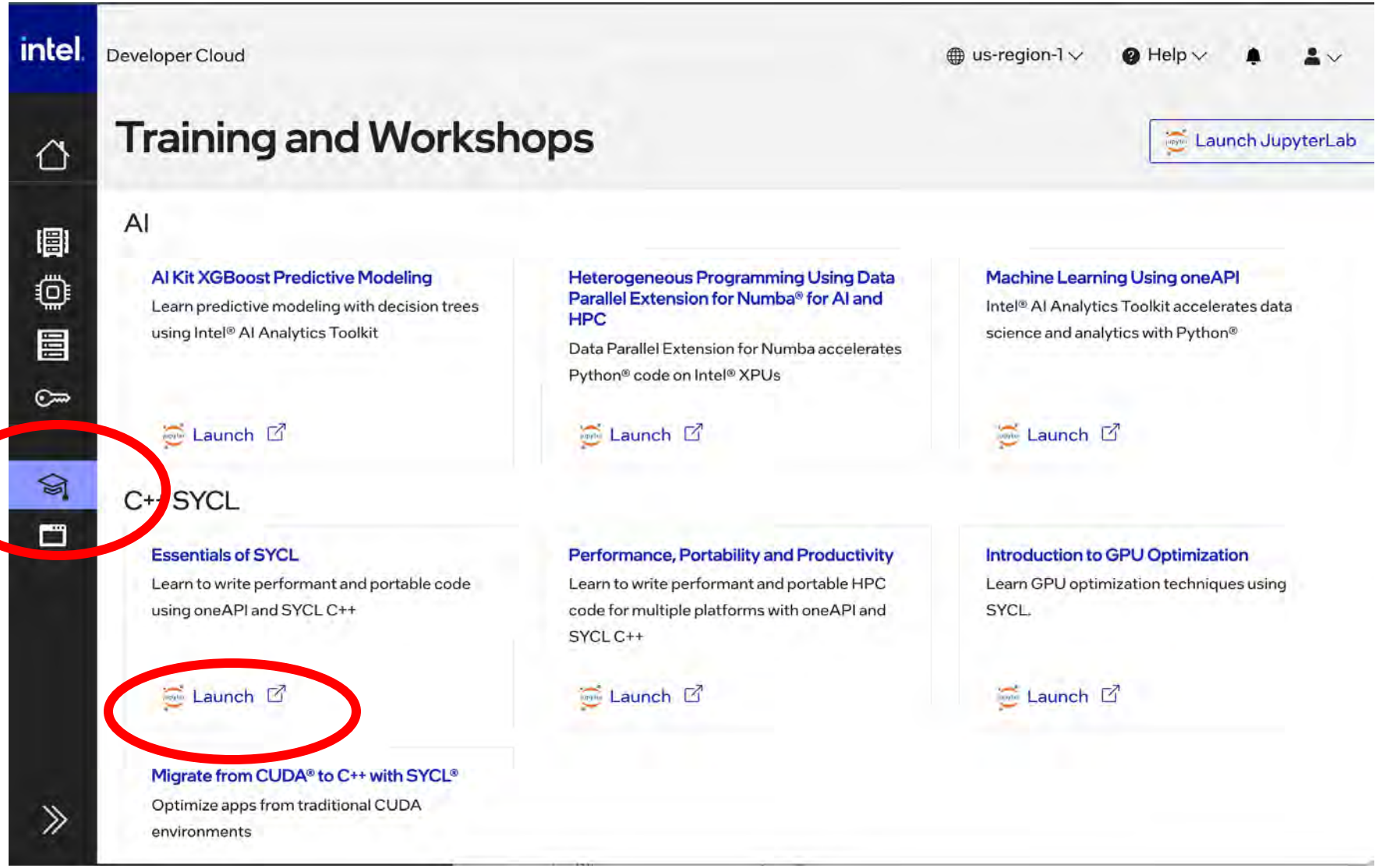
console.cloud.intel.com



console.cloud.intel.com



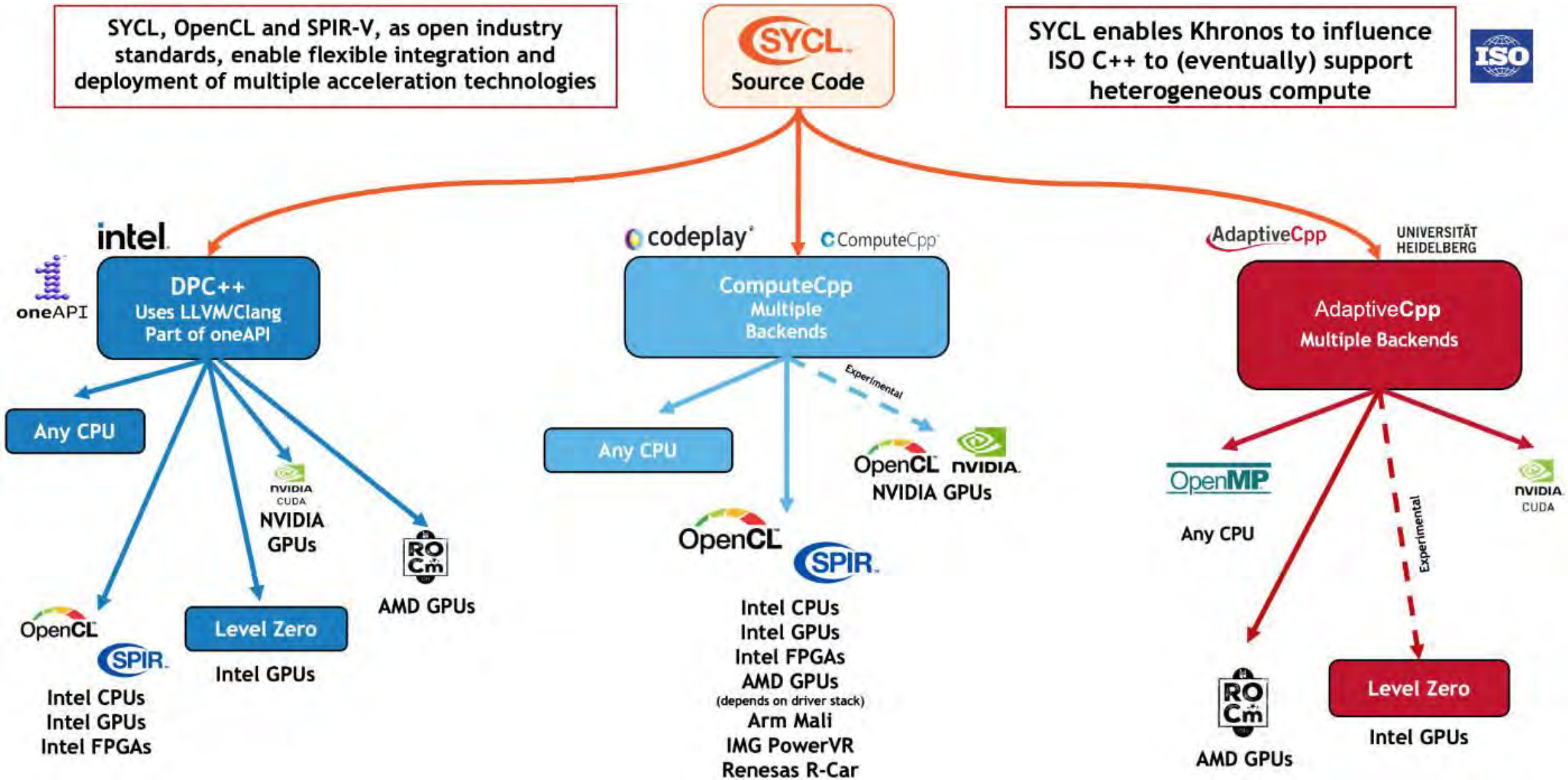
console.cloud.intel.com

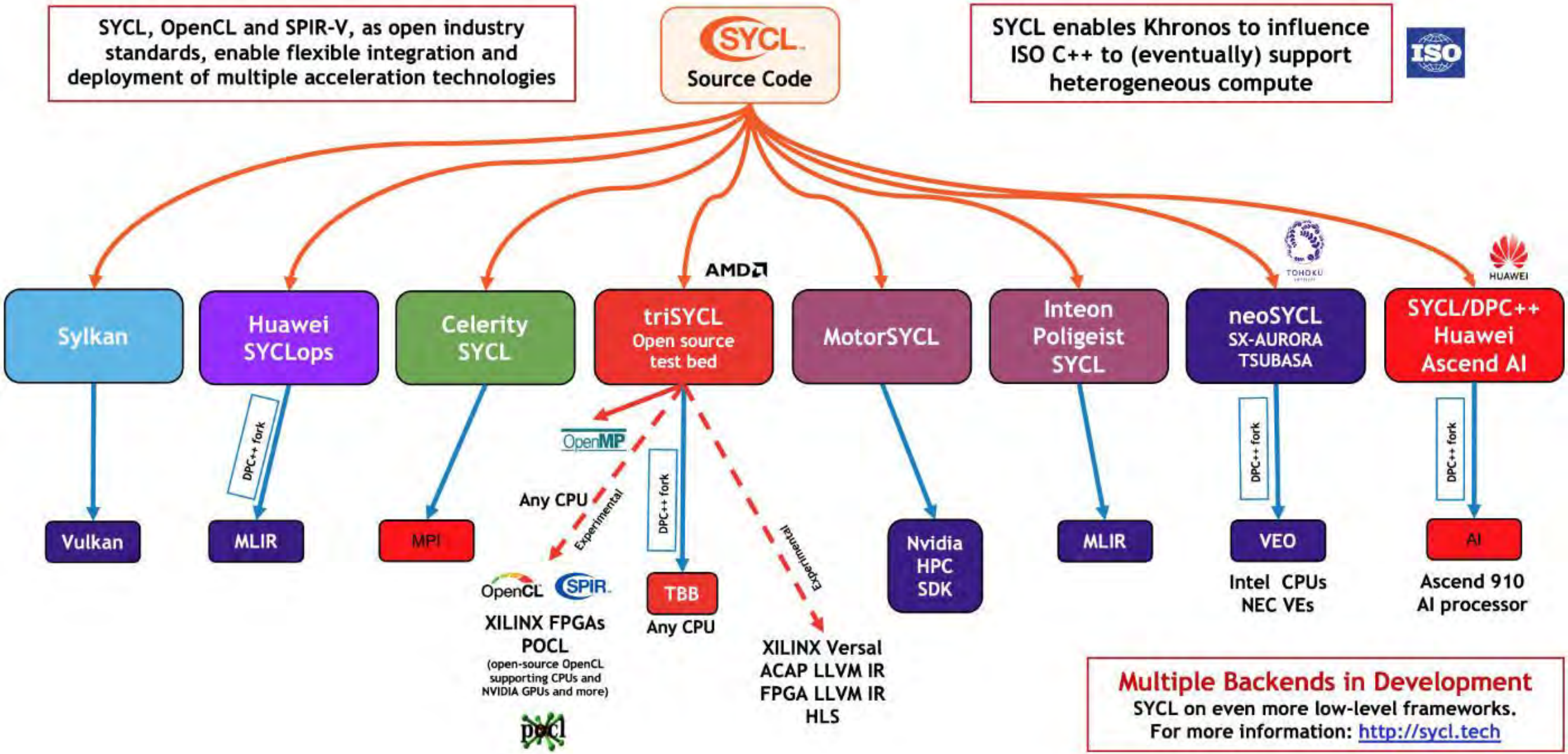



SYCL
TM

What is SYCL?

- SYCL (pronounced ‘sickle’) is a royalty-free, cross-platform abstraction layer that:
- Enables code for heterogeneous and offload processors to be written using modern ISO C++ (at least C++ 17).
- 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, OpenCL and SPIR-V, as open industry standards, enable flexible integration and deployment of multiple acceleration technologies
SYCL
Source Code
SYCL enables Khronos to influence ISO C++ to (eventually) support heterogeneous compute
ISO
intel
oneAPI
DPC++
Uses LLVM/Clang
Part of oneAPI
Any CPU
NVIDIA
CUDA
NVIDIA GPUs
AMD GPUs
OpenCL
SPIR
Intel CPUs
Intel GPUs
Intel FPGAs
Level Zero
Intel GPUs
codeplay
ComputeCpp
ComputeCpp
Multiple Backends
Experimental
Any CPU
OpenCL
NVIDIA
NVIDIA GPUs
OpenCL
SPIR
Intel CPUs
Intel GPUs
Intel FPGAs
AMD GPUs
(depends on driver stack)
Arm Mali
IMG PowerVR
Renesas R-Car
AdaptiveCpp
UNIVERSITÄT HEIDELBERG
AdaptiveCpp
Multiple Backends
OpenMP
Any CPU
NVIDIA
CUDA
Experimental
AMD GPUs
Level Zero
Intel GPUs


SYCL, OpenCL and SPIR-V, as open industry standards, enable flexible integration and deployment of multiple acceleration technologies
SYCL
Source Code
SYCL enables Khronos to influence ISO C++ to (eventually) support heterogeneous compute
ISO
Sylkan
Huawei SYCLops
Celerity SYCL
AMD
triSYCL
Open source test bed
MotorSYCL
Inteon Poligeist SYCL
TOHOKU
neoSYCL
SX-AURORA TSUBASA
HUAWEI
SYCL/DPC++
Huawei Ascend AI
DPC++ fork
DPC++ fork
DPC++ fork
OpenMP
Any CPU
Experimental
Experimental
Vulkan
MLIR
MPI
OpenCL
SPIR
XILINX FPGAs
POCL
(open-source OpenCL supporting CPUs and NVIDIA GPUs and more)
pocl
TBB
Any CPU
XILINX Versal
ACAP LLVM IR
FPGA LLVM IR
HLS
Nvidia HPC SDK
MLIR
VEO
Intel CPUs
NEC VEs
AI
Ascend 910
AI processor
Multiple Backends in Development
SYCL on even more low-level frameworks.
For more information: http://sycl.tech

CIUK 2023 Poster Competition Winner

Code coupling library for Multiphysics CFD and Solid Stress software for HPC Verification

Jessica Gould

UKRI - STFC

This project aims to couple programs for modelling of thermoelastic effects with computational fluid dynamics. It focuses on engineering the neutron and muon spallation source target, TS2 at ISIS. TS2 is made from Tungsten with a Tantalum casing, and has two water cooling channels. As the proton beam meets the nose, it experiences a periodic temperature change, causing stress throughout the target that limits its lifetime. Improvements would hope to save resources, reduce downtime and prevent radioactive waste. In modelling TS2, this project will provide an accessible code coupling library for high resolution verification of productive modelling alternatives, funded by the Ada Lovelace Centre.

Two open source software packages are the focus; Code_Saturne for CFD and Fenicsx, a PDE solver being developed for parallel computing that can be run on platforms from laptops to HPC's. Current options for coupled code models are limited by cost, licensing and resolution, and don't hold capacity for HPC. This project will build a library that will allow us to couple key engineering considerations between intricate finite element mesh's running on separate programs. This will be applied to a model of the ISIS TS2 and compared to current models and sensor data from ISIS for verification. The library will be useable in a container, meaning it can be run in a diverse range of supercomputing environments. It will be useable as a high-resolution, open source, HPC Library for Multiphysics code coupling. Importantly, it has the potential to provide quick and accurate supercomputer verification of low resolution, computationally less intensive, CFD and Solid Stress models.

The poster outlines the aims of the project and explains the challenges of modelling TS2. It shows images generated of heat transfer in turbulent flow, stresses in the dual material, considerations for parallel computing, how it can be used for productive HPC, the ease of use of the new open source software library and methods for verification of a successful project.

Science and Technology Facilities Council

Scientific Computing

scd.stfc.ac.uk

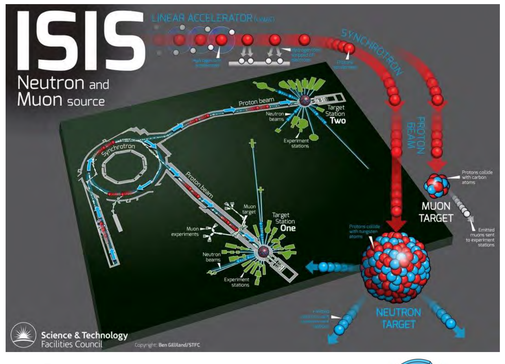
Code coupling library for Multiphysics CFD and Solid Stress software for HPC Verification

Computational Engineering Group, SCD, DL: Jessica A. Gould, Gregory Cartland-Glover

This project couples programs that model thermoelastic effects with computational fluid dynamics for engineering the neutron and muon spallation source target, TS2 at ISIS.

The ultimate aim is to develop a Multiphysics code coupling library that will be useable in a container, meaning it can be run in a diverse range of supercomputing environments. The two open-source software packages which are the focus; Code_Saturne for CFD and FEniCSx, a finite element solver, are both being developed for parallel computing that can be run on platforms from laptops to HPC's.

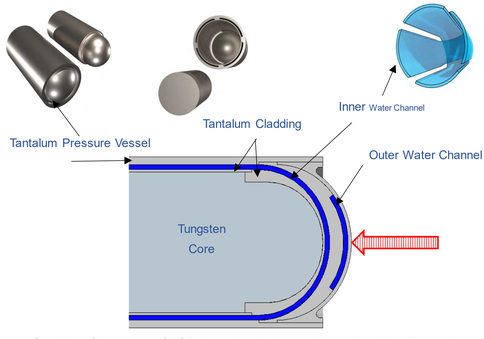
The project has an emphasis on considerations for parallel computing, how it can be used for productive HPC and the ease of use of the new open source software library.



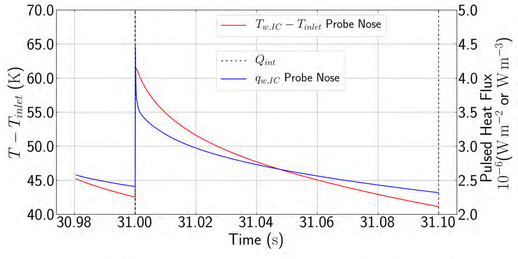
At ISIS, beams of protons are accelerated up to 80% the speed of light and ejected at target station 2 (TS2). The resulting neutron scattering allows for material properties to be measured for applications in medicine and engineering.

TS2 is made from Tungsten with a Tantalum casing,and has two water cooling channels. As the proton beam meets the nose, it experiences a periodic temperature change, causing stress throughout the target that limits its lifetime.
Improvements would hope to save resources, reduce downtime and prevent radioactive waste.

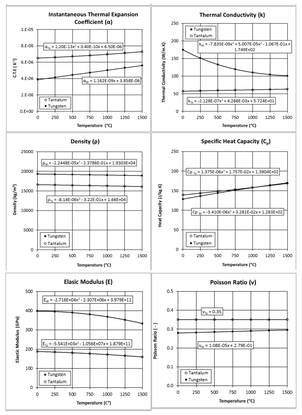
There are many contributions to take into consideration for the stress, as the two metals have different material properties meaning there is tensile stress in the cladding as it contracts more than the core as it cools leaving residual stresses throughout.



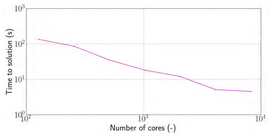
Graphic of the nose of TS2 showing the beam tip and cooling channels



Graph of EBRSM pulsed heat source ~23M cells by Greg Cartland-Glover



Challenges of the project
- Multiphysics Problem
- Stress concentrations from geometry
- Concurrent simulations
- Mesh communication
- Code coupling
- Current models
- Sensor data



Parallel Scale-up on Archer 2 with ~100M cells

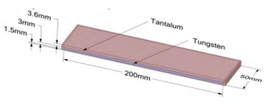
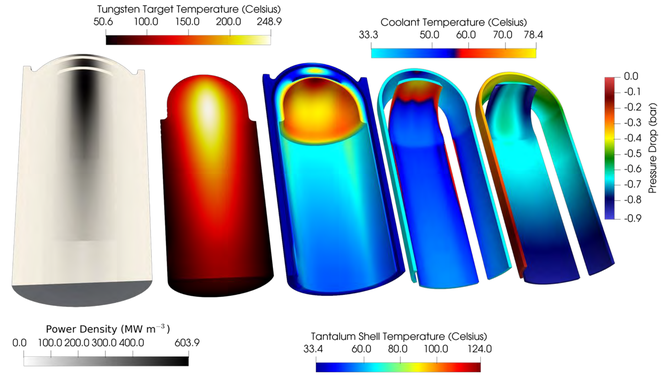
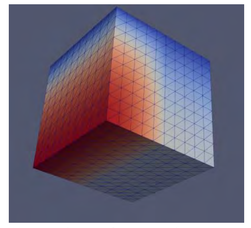


Photo and schematic of the test piece used for ENGIN-X experiemental test data

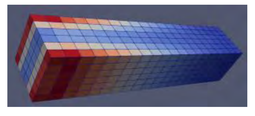


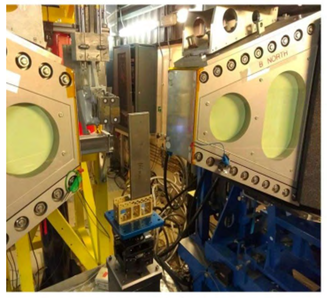
CHT with EBRSM turbulence ~23M cells (16M solid+7M fluid)

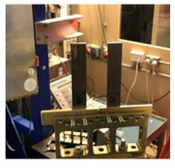
Models such as this are created using a variational formula of the heat equation for a body of dual materials, and equations governing small elastic deformations of a body. This involves writing a python script using the FEniCSx library of PDE solvers over a GMSH mesh using the Finite Element Method. There is a model in development for a fully coupled thermoelastic evolution problem, to model thermal expansion and stresses in a two-layer metal plate. For a fully coupled time-dependent thermal elastic simulation, mixed function space with displacement and temperature is essential, which has been implemented in the main branch of FEniCSx.



Visualization in Paraview of a dual material mesh using Poisson Equation (above) and modelling Von Mises stress (below) written using FEniCSx PDE solver







The ENGIN-X experiments that provide the test data on the dual material plate

Current options for coupled code models are limited by cost, licensing and resolution, and don't hold capacity for HPC. Code Saturne is a highly parallelisable software for computational fluid dynamics, which does not have a licensing limitation.

The next steps will be to set the material properties for a tantalum clad tungsten plate for validation against ENGIN-X measurements. When the code is verified, this can be applied to the geometry of the target, before being coupled to conjugate heat transfer calculations.
The final step in this will be to use the target model of thermal stresses to couple with RANS calculations of the target or the cooling channels of the target.

Modelling Pulsed Heat Source CHT of TS2
- We will model fluid flow in the channels with Code_Saturne and exchange surface temperature and its gradient with solid model of TS2 in FEniCSx to which the pulsed heat source is applied
- Have run cases using ~215 million cells on Archer 2 using 32 nodes and 16 and 32 nodes of Scarf using 23 million cells for conjugate heat transfer results
- Conjugate heat transfer with pulsed heat source
- Temperature using ~ 23 million cells (16M solid + 7M fluid)
- Use pulsed EBRSM model and exchange volume or surface temperature data with solid model of TS2 in FEniCSx

This project will build a library that will allow us to couple key engineering considerations between intricate finite element mesh's running on separate programs. This will be applied to a model of the ISIS TS2 and compared to current models and sensor data from ISIS for verification. The result is a suite of programs and libraries that can model thermal elastics coupled to conjugate heat transfer. It will be useable as a high-resolution, open source, HPC Library for Multiphysics code coupling. Importantly, it has the potential to provide accurate supercomputer benchmarking of lower resolution, computationally less intensive, CFD and Solid Stress models.

archer2 SCARF ParaView FEniCS project Gmsh code_saturne

Finding spin glass ground states using multi-stage quantum walks

Asa Hopkins

University of Strathclyde

Some current quantum devices use quantum annealing to perform quantum computation, which comes with a theoretical guarantee of success when enough time is given. This time, however, grows exponentially with problem size. An alternative model called a quantum walk does not have the guarantee of success but also avoids the exponential time needed. In practice, the success rate is exponentially small, making it no better. In this poster, a technique called multi-stage quantum walks is investigated, which can be seen as interpolating a quantum anneal by chaining together quantum walks. For the specific problem of finding spin-glass ground states, numerical results seem to suggest that multi-stage quantum walks only require polynomial time to find a solution, as surprisingly the optimal number of stages needed does not grow exponentially.

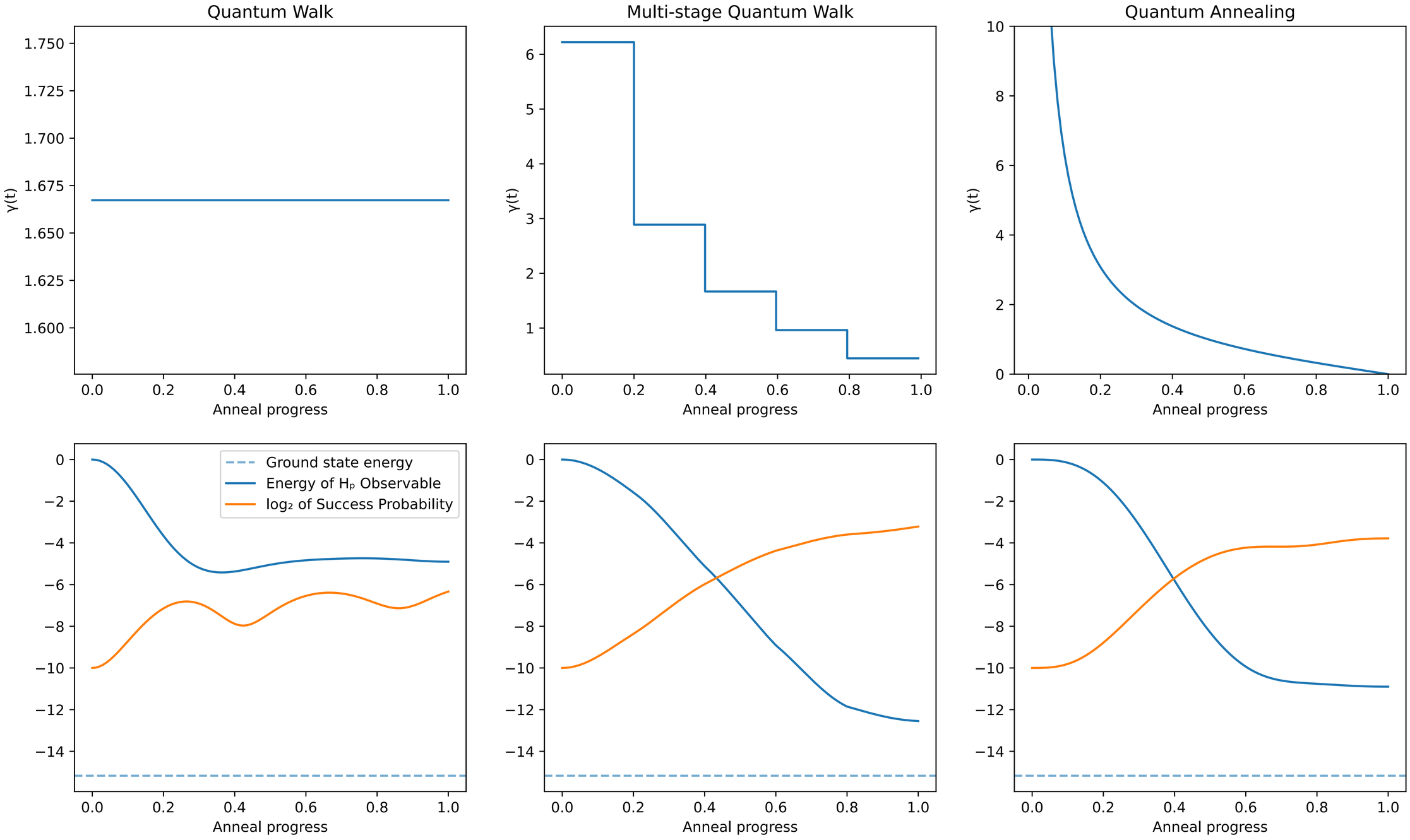
Finding spin glass ground states using multi-stage quantum walks

Asa Hopkins
asa.hopkins@strath.ac.uk
https://github.com/Asa-Hopkins/Multistage-QW

Solving problems with quantum annealing:

1. Encode the problem's solution as the ground state of a Hamiltonian $\mathbf{H}_P$
2. Prepare a register of qubits in the ground state of another Hamiltonian $\mathbf{H}_G$, which is chosen so that this is easy
3. Evolve the system under the Hamiltonian $A(t)\mathbf{H}_P + B(t)\mathbf{H}_G$
4. Measure the system, and hope it gives the optimal solution

To make things simpler, we work with the ratio $\gamma(t) = \frac{B(t)}{A(t)}$, leaving the energy scale to be defined by the hardware. Multi-stage quantum walks can then be implemented as a quantum anneal where $\gamma(t)$ is piecewise constant.

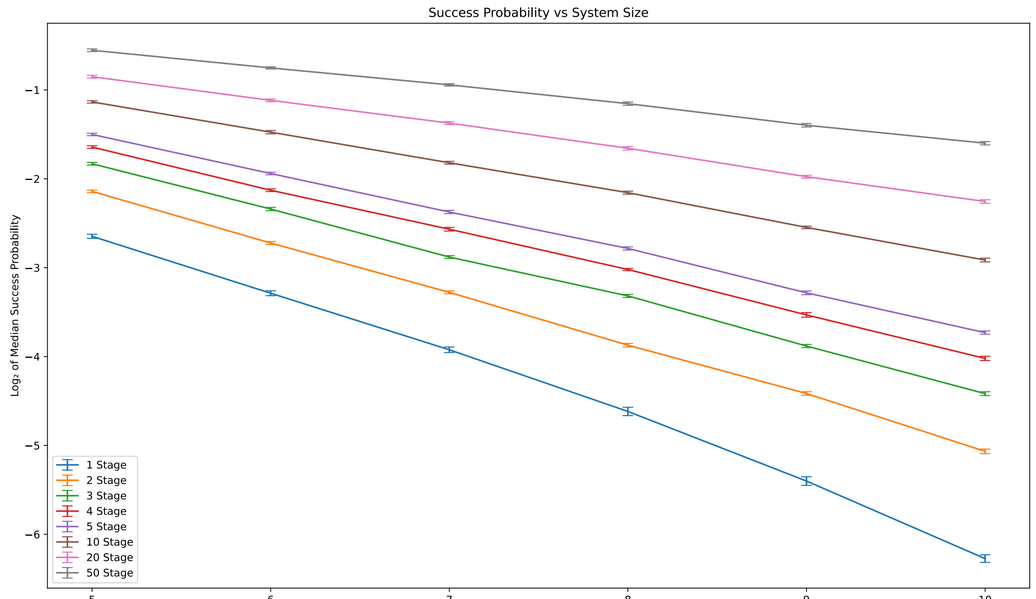


When to measure?

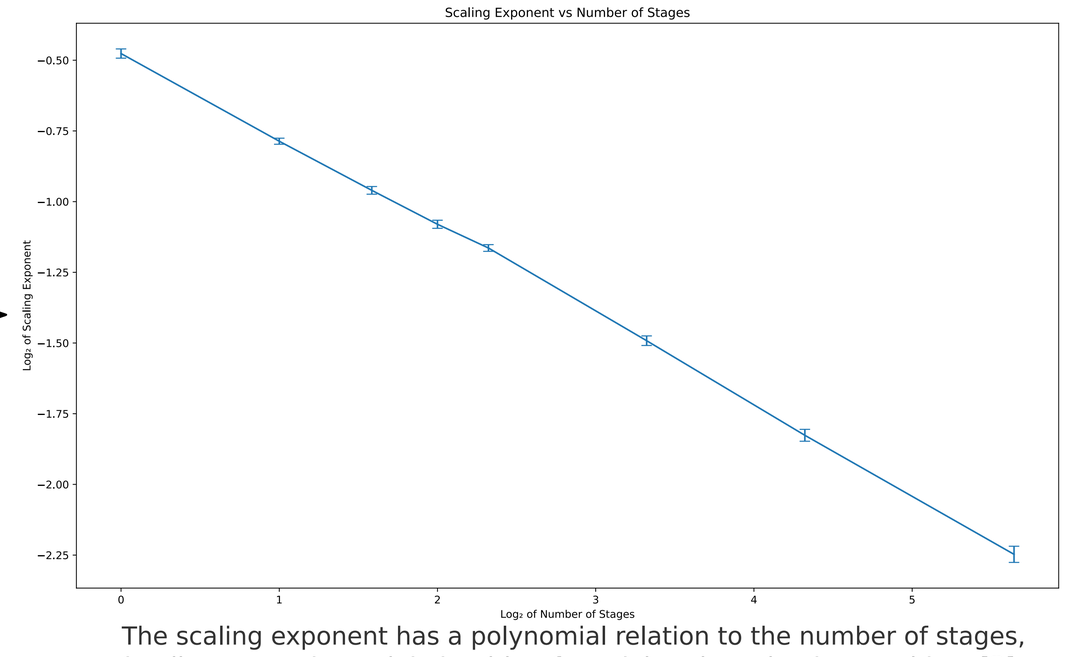
To average out the oscillating succes rate, it is necessary to sample at multiple times, but what times are best? There is a 'warm up' time where the state energy is decreasing rapidly, and a Taylor expansion shows that this decrease is initially quadratic. Since the energy always starts at 0 and can never go below the true ground state which is $O(n)$ [1], then after a time $O(\sqrt{n})$ the rapid decrease must have finished.

How to choose $\gamma(t)$?

In quantum annealing, it is known that for any given problem there is a choice of $\gamma(t)$ that solves the problem in linear time. Finding this perfect γ for each problem is sadly infeasible so heuristics are used to choose reasonable but suboptimal values. The 'infinite time' success chance suggests that we want each stage to rotate the state a fixed amount towards the end state. To achieve this we therefore 'rotate' the Hamiltonian and choose $\cos(\frac{n\pi}{2(m+1)})\mathbf{H}_G + \sin(\frac{n\pi}{2(m+1)})\mathbf{H}_P$ for the n^{th} stage out of m.



The problem becomes exponentially more difficult as more spins are added, but the difficulty grows more slowly for more stages.



The scaling exponent has a polynomial relation to the number of stages, leading to a polynomial algorithm for solving the spin glass problem [2].

PhD supervisor: Viv Kendon

[1] Adam Callison et al. *Finding spin glass ground states using quantum walks.* https://arxiv.org/abs/1903.05003
[2] Aditi Misra-Spieldenner et al. *Mean-Field Approximate Optimization Algorithm.* https://arxiv.org/abs/2303.00329

University of Strathclyde Science
UK Research and Innovation

Computational Infrared and Raman spectroscopy in ChemShell

Jingcheng Guan

University College London

To aid the interpretation of experimental vibrational spectra, we developed and implemented computational infrared and Raman facilities in the ChemShell computational chemistry environment using hybrid quantum mechanical and molecular mechanical approach. Density functional theory for the electronic structure calculations and classical forcefields for the environment are employed. Computational vibrational spectra for chemical active sites are reported using electrostatic and fully polarizable embedding to achieve more realistic vibrational signatures of material systems, including solvated molecules, proteins, zeolites and metal oxide surfaces, providing useful insight into the effect of the chemical environment on the signatures obtained from experiment. This work has been facilitated by the efficient task-farming parallelism implemented in ChemShell for high-performance computing platforms.

Computational Infrared and Raman Spectra by Hybrid QM/MM

Jingcheng Guan*, You Lu, Kakali Sen, Jamal Abdul Nasir, Andrew M. Beale, C. Richard A. Catlow, Thomas W. Keal and Alexey A. Sokol

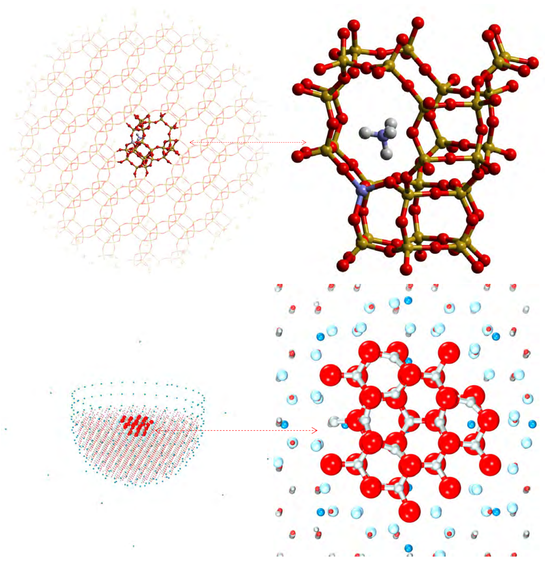
***Chemistry Department, University College London**

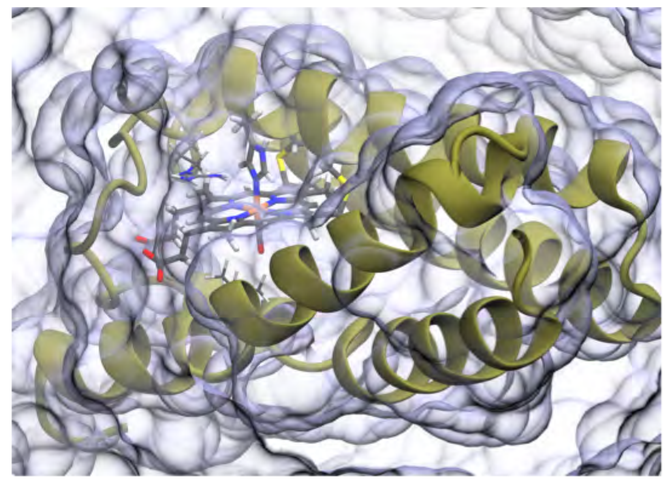
Motivation

- To better interpret experimental vibrational signatures of chemical active sites with environmental effects

Chemistry of interest

- Solvated histidine molecule
- Brønsted and Lewis acid sites in catalytic chabazite
- Hydrogenation on zinc oxide polar surfaces
- Haem protein with differently oriented ligands
- Excellent agreement with experiment

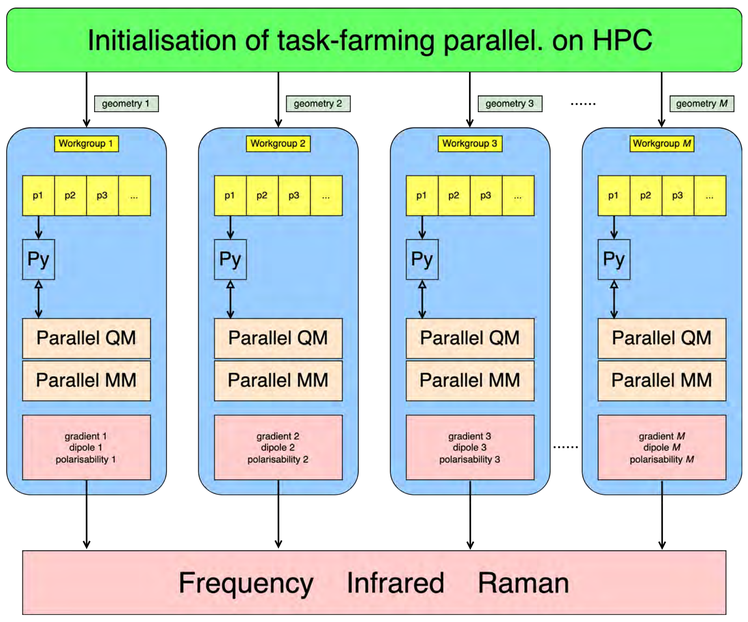




ChemShell
multiscale computational chemistry

Methodology

- Harmonic approximation of PES (normal modes)
- DFT and TDDFT
- Hybrid QM/MM
- Electrostatic & fully polarizable embedding
- CPKS: *v*-dependent polarisability

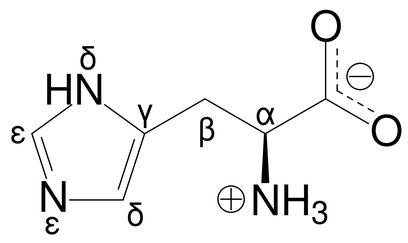


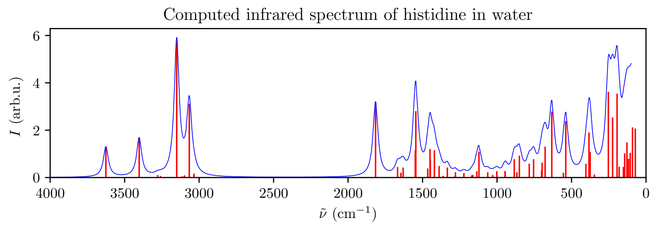
Development

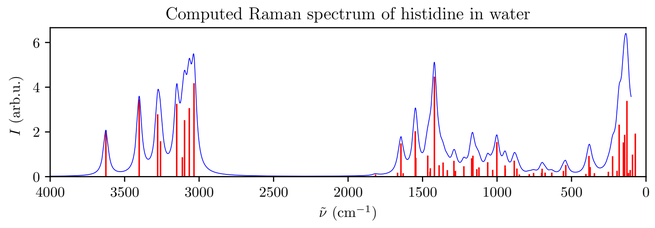
- ChemShell
- NWChem
- Taskfarming parallel
- Infrared
- Resonance and non-resonance Raman

Solvated Histidine Molecule & Haem Protein

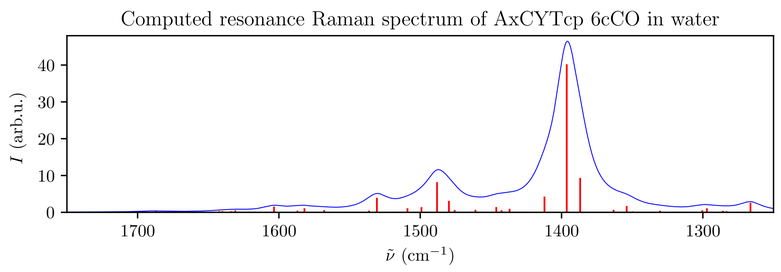
- IR and Raman of solvated histidine
- Thermodynamic snapshot obtained from MD

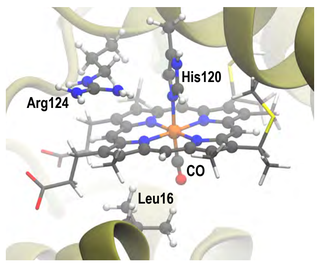






- Resonance Raman of 6cCO haem protein with histidine and arginine in water
- Electronic transition oscillators determined by TDDFT

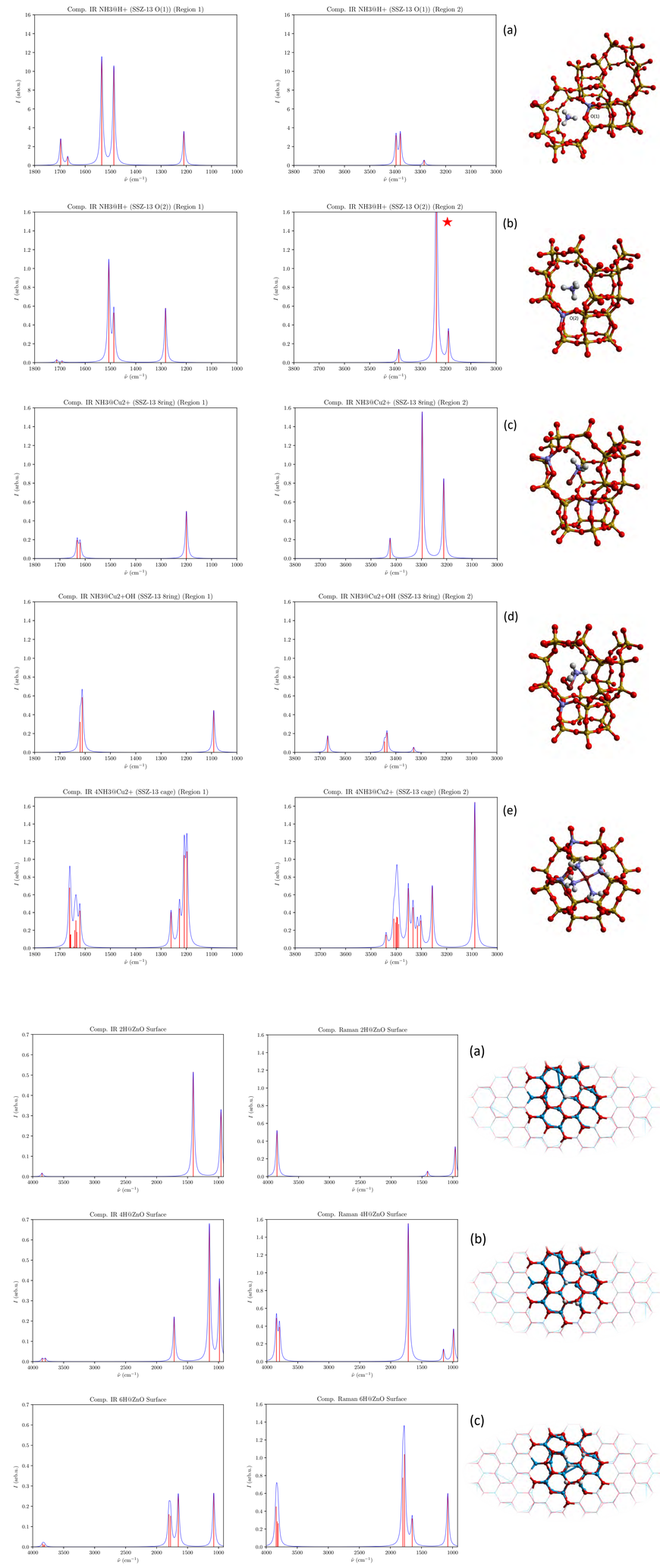




Catalytic Chabazite & Zinc Oxide Polar Surface

- NH_3 probing
- Brønsted acidity (H^+)

- Lewis acidity (Cu^{2+})
- $(NH_3)_4@Cu^{2+}$
- Type-II Hydride
- Type-I Hydrides



Contact: jingcheng.guan@ucl.ac.uk

[1] Guan, Jingcheng, et al. Philosophical Transactions of the Royal Society A 381.2250 (2023): 20220234.

[2] Abdul Nasir, Jamal, et al. Journal of the American Chemical Society (2022).

UKRI Science and Technology Facilities Council
MATERIALS AND MOLECULAR MODELLING HUB
HEC MCC UK
archer2

Lattice Boltzmann Inspired Quantum Walk for Solving PDEs

Lara Janiurek

University of Strathclyde

Efficiently solving Partial Differential Equations (PDEs) holds potential for substantial scientific advancements across various disciplines. This work aims to create a quantum algorithm involving a quantum walk for solving PDEs in a fluid dynamics context. We seek an algorithm applicable on practical quantum computers. The algorithm being currently developed is inspired by Lattice Boltzmann methods, and aims to simulate families of PDE, enabling parameter tuning to match desired PDEs.

Lattice Boltzmann Inspired Quantum Walk for Solving PDEs

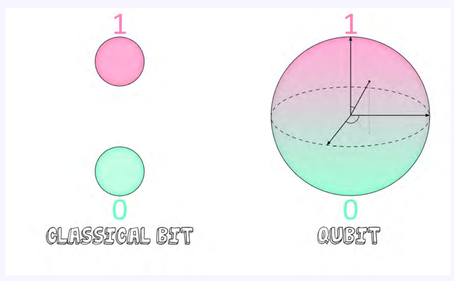
Lara Janiurek
University of Strathclyde

Problem

- Efficiently solving Partial Differential Equations (PDEs) holds potential for substantial scientific advancements across various disciplines.
- This work aims to create a quantum algorithm involving a quantum walk for solving PDEs in a fluid dynamics context. We seek an algorithm applicable on practical quantum computers.
- The algorithm being currently developed is inspired by Lattice Boltzmann methods, and aims to simulate families of PDE, enabling parameter tuning to match desired PDEs.

Quantum Computing

Quantum computing uses quantum bits (qubits) to process information, which can exist in superpositions of 0 and 1 and exhibit entanglement. Quantum gates manipulate qubits, enabling parallel processing and potential exponential speedup for certain problems [1].



Neutral Atom Arrays (Strathclyde)

Quantum SPH paper

The algorithm developed must be applicable on practical quantum devices, such as the SQuAre neutral atom array being developed at Strathclyde. Another quantum algorithm for fluid dynamics is being developed by the QEVEC team based on Smoothed Particle Hydrodynamics.

Lattice Boltzmann Methods

- We define the **Navier Stokes Equation**:

$$\rho \frac{d\vec{V}}{dt} = -\nabla p + \rho \vec{g} + \mu \nabla^2 \vec{V}$$

- Lattice Boltzmann methods computationally simulate fluid dynamics and complex phenomena. Instead of solving Navier-Stokes equations directly, they discretize space and time into a lattice [2].
- Particles collide and propagate using simplified kinetic equations.
- **Both discrete time quantum walks and Lattice Boltzmann methods employ collision and streaming steps on a lattice.**

Discrete Time Quantum Walks (DTQW)

- DTQWs embody quantum evolution in discrete intervals, enabling superposition of states, unlike classical walks.
- The process involves a quantum particle carrying a multi-state quantum system for the coin. The coin toss is effected by a unitary operator [1].

We define the **evolution of a discrete time quantum walk** as [3]:

$$|\Psi_\mathbf{t}\rangle = \mathbf{U^t} |\Psi_0\rangle \quad \text{where} \quad \mathbf{U} = \mathbf{S}(\mathbf{C} \otimes \mathbf{I})$$

The system evolves in discrete time steps. Each step involves a quantum coin operation ($\mathbf{C}$) and a conditional shift ($\mathbf{S}$) based on the coin's outcome. An example of the 1D DTQW shift and coin operators are:

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} cos(\theta) & sin(\theta) \\ sin(\theta) & -cos(\theta) \end{bmatrix}$$

$$\mathbf{S} |\uparrow\rangle \langle\uparrow| \otimes \sum_i |i+1\rangle \langle i| + |\downarrow\rangle \langle\downarrow| \otimes \sum_i |i-1\rangle \langle i|$$

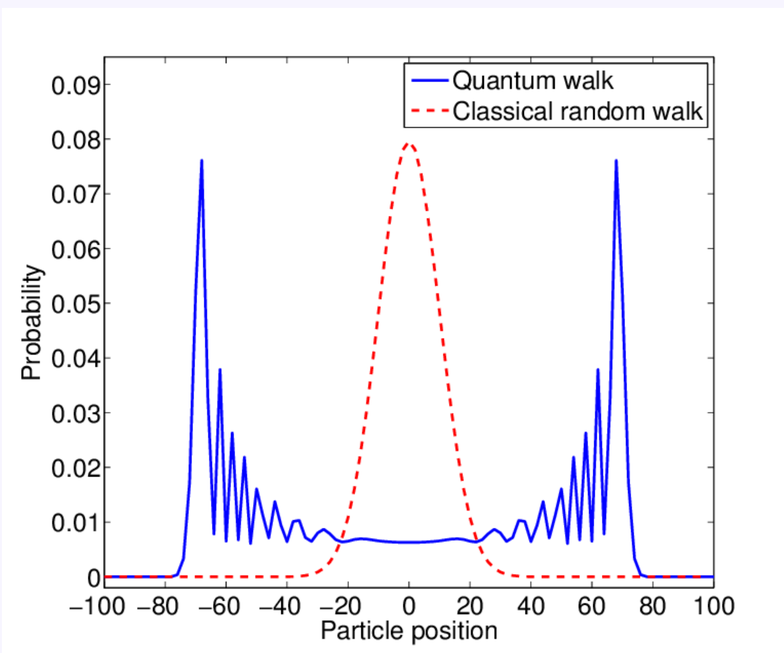


Figure 1: D2Q9 Lattice

- DTQWs discretely approximate continuum equations. E.g, the continuum limit of a simple 1D walk is the 1D Dirac equation.
- The coin operator's parameters determine the system's dynamics and the resulting family of differential equations in the continuum.
- Tunable coin parameters offer degrees of freedom to customise desired partial differential equations, and introducing non-linearity is essential for simulating higher-order differential equations [3].

Lattice Boltzmann Inspired Quantum Walk

- A quantum walk may be performed on the same D2Q9 lattice used in Lattice Boltzmann, shown in Figure 2 [2].
- A D2Q9 walk involves a 9-dimensional coin operator.
- The quantum coin operator determines the family of simulated PDEs. Parameters in the coin facilitate user-adjustment for correct PDEs in the continuum limit.

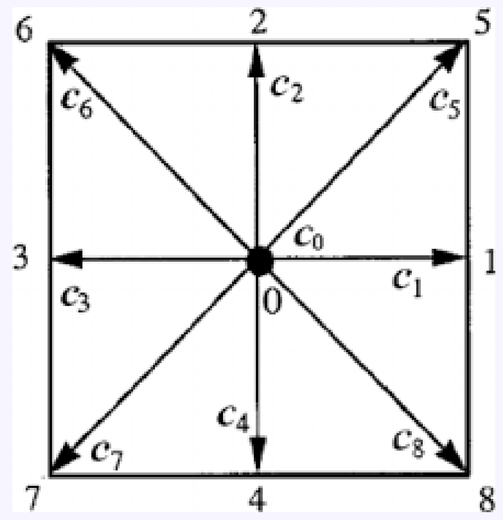


Figure 2: D2Q9 Lattice

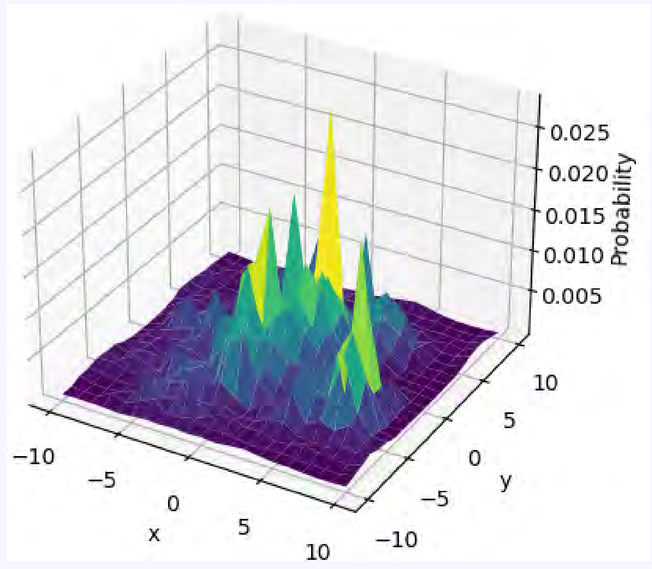


Figure 3: Simple D2Q9 walk with entangled coins

- We've successfully parameterised a 9-dimensional coin with Lattice Boltzmann dynamics. However, the non-linear relaxation term in Lattice Boltzmann must be integrated into the quantum walk.
- This non-linearity generates higher-order continuum equations, essential for a non-linear PDE solver. Efficiently incorporating this non-linearity while maintaining accurate dynamics is currently being explored.

References

[1] Lovett, N.B., Cooper, S., Everitt, M., Trevers, M. and Kendon, V., 2010. Universal quantum computation using the discrete-time quantum walk. Physical Review A, 81(4), p.042330.
[2] Krüger, T., Kusumaatmaja, H., Kuzmin, A., Shardt, O., Silva, G. and Viggen, E.M., 2017. The lattice Boltzmann method. Springer International Publishing, 10(978-3), pp.4-15.
[3] Mlodinow, L. and Brun, T.A., 2018. Discrete spacetime, quantum walks, and relativistic wave equations. Physical Review A, 97(4), p.042131.

Modelling of Muon Experiments with Muon Galaxy

Nalin Gupta & Maitrayee Singh

UKRI – STFC

In this poster we present the Muon Spectroscopy Computational Project (MSCP)- an initiative of the Scientific Computing Department (SCD) to develop sustainable and easy to use software for the interpretation of muon experiments. Particularly, we focus on an example of the application of pymuon-suite - one of the software tools being developed and maintained by MSCP as an implementation in Muon Galaxy - for the prediction of muon stopping sites in crystalline materials LiGaX4 (X = Cl, Br, I) and LiCoO2, which are being studied for their potential as battery materials.

Science and Technology Facilities Council

Scientific Computing

scd.stfc.ac.uk

Modelling of Muon Experiments in Battery Materials with Muon Galaxy

Nalin Gupta, Maitrayee Singh, Patrick Austin, Leandro Liborio
Science and Technology Facilities Council

Muons and the Muon Spectroscopy Computational Project

Muons are sub-atomic particles that are generated at target station 1 in ISIS, at STFC. Muons are 100% spin polarised particles with spin ½, which can be thought of as either light protons or very heavy electrons.

Contrary to what happens in a neutron or x-ray experiment, muons are not diffracted by the sample: they are implanted into the sample, and knowing their implantation site is crucial for the interpretation of muon experiments.

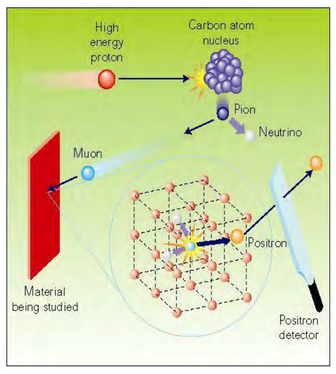


Figure 1: A schematic description of Muon experiment.

The MSCP[2] develops software tools for tackling computational challenges in muon spectroscopy. These tools can be used for:

- Identify the **muon stopping site(s)** in a crystalline system
- Simulate the **spin dynamics** of a system containing a muon, electrons, and atomic nuclei, with various experimental setups and couplings
- Fit a spin dynamics simulation to experimental data

These tools can be chained together to form an analysis pipelines known as workflows. We release the tools as Python packages: **pymuon_suite** and **muspinsim,** which can be installed using pip or conda, or run online using **Muon Galaxy**.

Muon Galaxy

Galaxy is an open-source web platform for data intensive research. It allows users to **run complex workflows** and visualise results without any programming experience.

Analysis in Galaxy is **easily reproducible**, as a consistent computational environment is used for every job, and all data files are stored in a 'history'. Data and workflows can be **easily shared** with other researchers or made public; publishing a workflow alongside a paper is a great way to ensure reproducibility.

Galaxy is well established in the biology community, but we have launched an instance specifically for muon science – known as **Muon Galaxy**[3]. We are making all our tools available in Muon Galaxy, along with tutorials, example workflows, and visualisations.

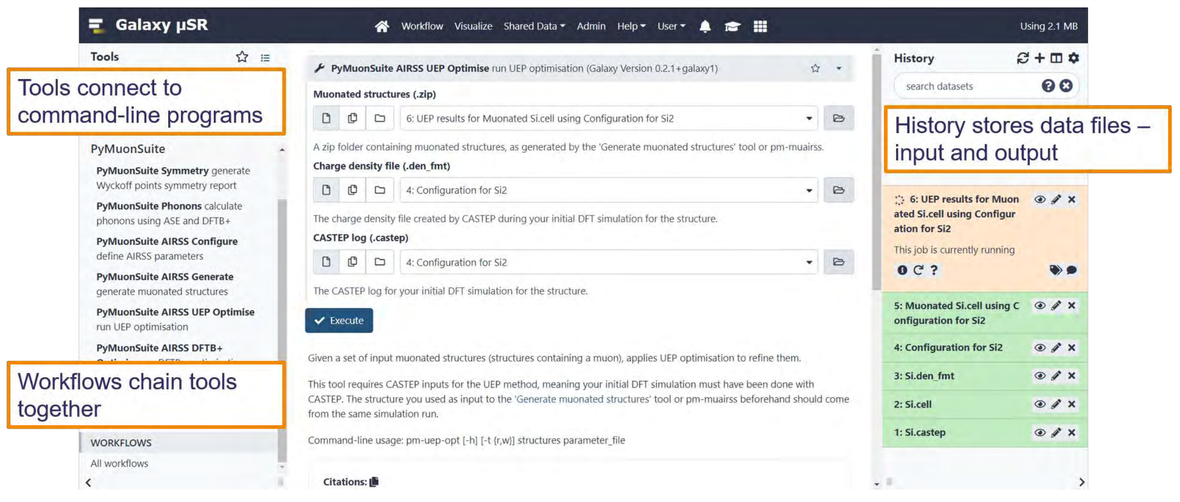


Figure 2: The muon galaxy interface with a description of the features provided.

Lithium-ion Battery Materials

Lithium-ion batteries consist of negative and positive electrodes composed by two lithium insertion materials.

During the charging process, lithium ions are inserted into the solid matrix that forms the positive and negative electrodes without destructing their core structures.

During the discharge process, lithium ions are extracted from the electrodes. As this happens, electrons are simultaneously extracted from one electrode and injected into another electrode, storing and delivering electrical energy[1].

Figure 3 exemplifies how a Li-battery operates. In spite of a long research history on Li-ion batteries, the value of the Li diffusion coefficient, D_{Li}, is still difficult to determine with any reliability.

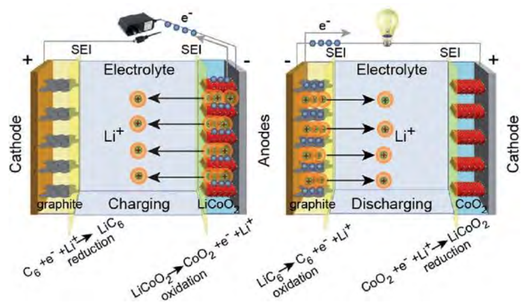


Figure 3: A schematic illustration on a lithium-ion battery consisting of two lithium insertion electrodes.

Why muons?

Knowing the muon stopping site helps to determine whether the muon is either likely to diffuse with Li when T is increased or to remain at its stopping site. If the muon site is stable with respect to Li diffusion, the Li diffusion coefficient, D_{Li}, is given by:

$$D_{Li} = \sum_{i=1}^{n} \frac{1}{N_i} Z_{\nu i} \, s_i^2 \, \boldsymbol{\nu}$$

N_i is the number of Li sites in a given diffusion path; $Z_{\nu i}$ is the Li-vacancy fraction; s_i the jump distance; and $\boldsymbol{\nu}$ is the magnetic field fluctuation rate at the muon site, which is measured in a muon experiment.

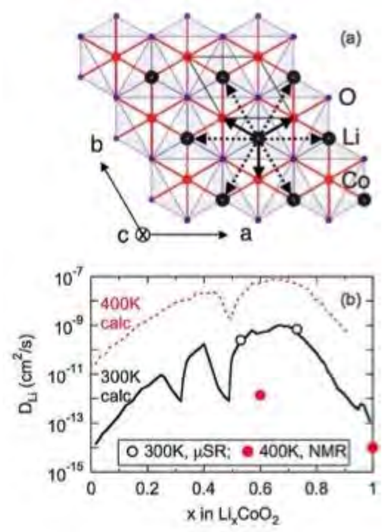
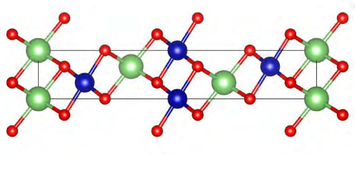


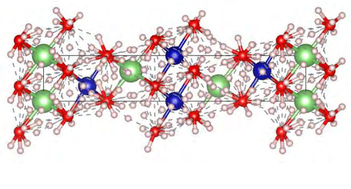
Figure 4: (a) Li diffusion paths. (b) Comparison of D_{Li} values measured with NMR and muons experiments and estimated with computer simulations[4].

The Workflow to obtain Muon stopping sites

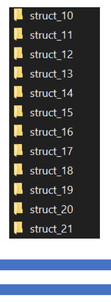
DFT pure host



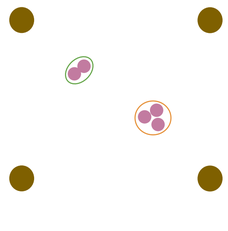
Create randomly populated muonated structures



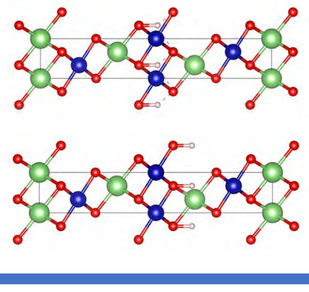
Relax electrostatic forces on muons



Cluster relaxed structures

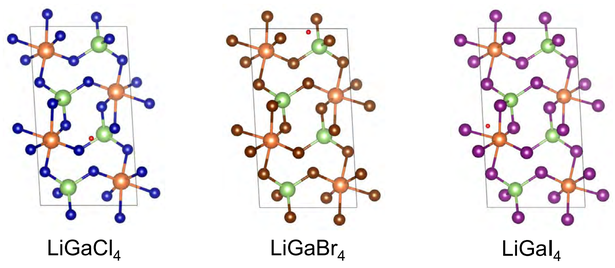


μ^+ stopping sites



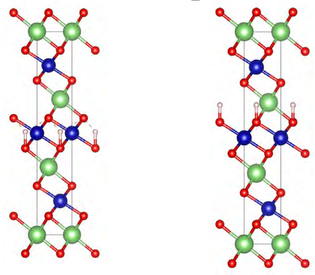
Using the UEP method[5] with the above workflow in Muon Galaxy, the most stable stopping sites for $LiGaX_4$ and $LiCoO_2$ are shown in the figures below:

$LiGaX_4$ (X=Br, Cl, I)



$LiGaCl_4$ $LiGaBr_4$ $LiGaI_4$

$LiCoO_2$



The sites obtained correlate to sites that remain unchanged when the Lithium ions diffuse through the material during the battery's charge and discharge process, allowing for accurate determination of the diffusion coefficient value.

References

[1] Journal of Power Sources 174 (2007) 449–456

[2] https://muon-spectroscopy-computational-project.github.io/

[3] https://muongalaxy.stfc.ac.uk/

[4] PRL 103, 147601 (2009)

[5] J. Chem. Phys. 28 July 2020; 153 (4): 044111

MSCP

Software Sustainability Institute

Galaxy PROJECT

ada lovelace centre

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Using HPC to significantly increase bioimaging workflow performance

Victor Ionescu

University of Liverpool

Data processing and analysis has been identified as the area which typically consumes the most time for bioimaging researchers and requires the most user involvement (Schmidt, C., et al. 2022). In the fields of electron and light microscopy this can be in part due to the large volume of data produced. In recent years several tools have become available that leverage machine learning which despite the high computational requirements of these tools and high volume of data, many are still used through a GUI on a single workstation leaving very little headroom for automatisation and improvements in runtime. We believe many of these workflows can be adapted for use on a HPC cluster. However, there is a significant barrier to entry, requiring skills typically alien to bioimaging researchers such as experience using Linux CLI. We look to improve accessibility to HPC by providing users with a familiar GUI and user configurable workflows which can scale across a HPC cluster.

Using HPC to significantly increase bioimaging workflow performance

Victor Ionescu, Marie Held, Tony McCabe, Marco Marcello, Tobias Zech, Andrew Collins

Abstract

Data processing and analysis has been identified as the area which typically consumes the most time for bioimaging researchers and requires the most user involvement (Schmidt, C., et al. 2022). In the fields of electron and light microscopy this can be in part due to the large volume of data produced. In recent years several tools have become available that leverage machine learning. Despite the high computational requirements of these tools and high volume of data, many are still used through a GUI on a single workstation leaving very little headroom for automatisation and improvements in runtime. We believe many of these tools can be modified to run on HPC cluster. By providing users with our web interface (**SLURMGUI**), we give researchers a user friendly portal to HPC without the requirement of prior experience. The purpose of this poster is to showcase the substantial performance gains achievable by harnessing the power of HPC cluster in a bioimaging context and how this can be brought to the desktop of a researcher without a background in HPC.

Figure 1. Most Time Consuming Steps in Bioimaging 2021 NFDI4BIOIMAGE community survey
Participants in the bioimaging 2021 NFDI4BIOIMAGE community survey chose which step is the most time-consuming. Diagram adapted from (Schmidt et al., 2022)

Methods

There are a number of computational methods which are critical to reducing researcher involvement in multiple aspects of analysis. Many of these methods can either be run natively on a HPC cluster or can be modified to do so. All of these tools were containerized using Apptainer (formerly known as Singularity) allowing for enhanced reproducibility and portability. This allowed us to easily run our workflows over multiple clusters and ensured that as many variables were controlled as possible.

SLURMGUI

We provided our ends users with a GUI web interface in order to increase HPC accessibility for researchers. This facilitated simple SLURM job submission and management without any prior experience using SLURM. On top of this we integrated a Javascript version of the Fiji image processing tool (Wei Ouyang et al., YW 2021). This ensured that results can be viewed and additional image processing can be added manually with a tool which researchers are already comfortable working with.

Figure 2. Workflow Steps

The instruments are connected via fiber to our storage nodes which use GlusterFS to distribute load and replicate data over multiple nodes. This ensures rapid data transfer allowing for rapid throughput and improved analysis speeds. Following this, a job is run by researchers via our web interface SLURMGUI. This starts the parent job which then prepares the image to distribute processing to each child job ensuring maximum efficiency. Finally child jobs return data and statistical analysis / follow up jobs begin.

DECONWOLF

Deconwolf (DW) is an open-source software developed to enable high-performance deconvolution of widefield fluorescence microscopy image stacks and large tissue scans. (Wernersson et al., 2022) It's designed to process large image datasets, such as those generated by whole-slide imaging microscopes, on a laptop computer .

N2V

N2V (Noise2Void) is a Tensorflow based CNN for denoising images, particularly useful in the field of bioimaging (Krull et al., 2019). It's an innovative approach as it maintains a high level of accuracy while not requiring pairs of noisy and corresponding noise-free images for training.

CELLPOSE

Cellpose is a machine learning algorithm which automates cellular segmentation (Stringer et al., 2021). The convolutional neural network (CNN) design is built on the PyTorch framework and this makes Cellpose a powerful tool due to its flexibility and accuracy.

CLEMReg

CLEMReg automates the alignment of correlative light and volume electron microscopy (vCLEM) datasets, which is a process that traditionally requires expert knowledge and manual labor (Krentzel et al., 2023). This automation leverages machine learning using the Empanada MitoNet model from the PyTorch ecosystem.

Results

All of the tools were benchmarked first on a typical researcher workstation used for analysis after data has been collected from the instruments. This was a 4 core 32 GB RAM desktop running Ubuntu 20.04 LTS. Following this the benchmark for each tool was run again on the 1144 core 7098 GB RAM LIV-SRF cluster (**Figure 3.**). When running on the cluster the multi-framed images were split following a scaling analysis benchmark to ascertain what the optimum point was for resource allocation.

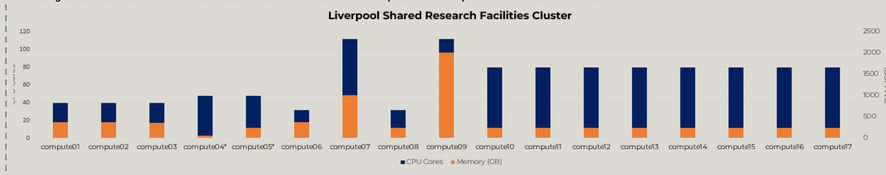


Figure 3. Liverpool Shared Research Facilities Cluster Specifications (Nodes with * are equipped with a Tesla P40 GPU)

Deconwolf

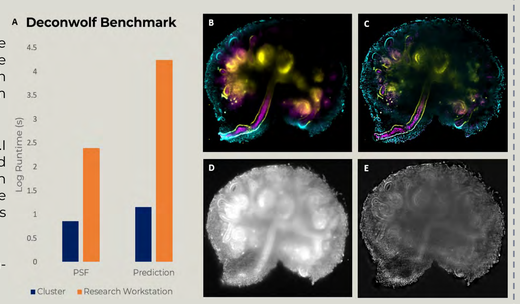
Deconwolf required very little modification in-order to get it running on the cluster. The benchmark input TIFF was split into individual frames using a C script and the LibTIFF library, this was first tested using FIJI however was too slow. Multithreading was tested however we found the bottleneck to be IO speed. The results showed a significant improvement when scaling up using the LIV-SRF cluster. While the PSF (Point spread function) generation time showed **5x** improvement in run time, the prediction element of the workflow saw the a **70x** improvement due to the ability to run each frame from the TIFF in parallel (**Figure 4**).

Figure 4. Deconwolf Benchmark and Data

A) The benchmark shows a significant performance increase when running on the Cluster. Splitting the sample image while also allocating 32 cores to each child job resulted in a decrease in total run time from 17300 seconds to 246 seconds.

B / C) Input image which was taken on a Ziess Z.I Lightsheet microscope, the channel which was used for benchmarking was dyed with AF647 and has an emission of 668 nm. Full deconwolf output image shows significantly clearer structures with far less noise.

D / E) Rod structure is much clearer following deconvolution (**D**).



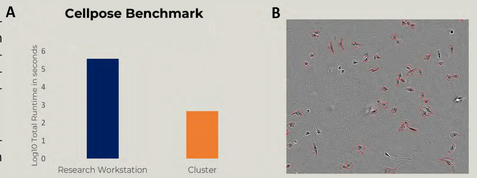
CELLPOSE

Cellpose required multiple scaling analysis to be performed before running on the cluster. This was due to its ability to run on a GPU however our conclusions were that the performance gained from utilizing the GPUs did not contend with the performance gain from distributing across all 1000 CPU cores. The dataset used for benchmarking was 7560 TIFF images, processing this manually on the researcher workstation was completely unfeasible and as such was scripted. The performance gained from running the workflow on the cluster was substantial with around a **640x** improvement in run time (**Figure 5**).

Figure 5. Cellpose Benchmark and Data

A) Cellpose benchmarking showing the performance gain achieved when running on the cluster. This was one of the more impressive results showcasing with an improvement in runtime of several orders of magnitude

B) Cellpose output image shows clearly segmented cells (outlined in red) with a high degree of accuracy.



N2V

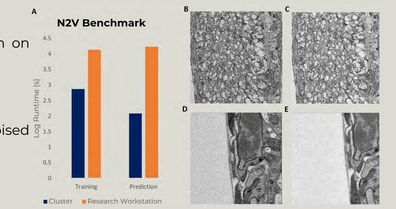
N2V was run using an adapted version of the scripts provided with the tool. The training was run with 10 epochs and 400 steps per epoch. There was a substantial improvement of a few orders of magnitude in both the training and prediction aspects of the workflow. The ability to process each subframe of the TIFF individually allowed massive parallelization of the large dataset resulting in **34x** improvement in total runtime.

Figure 6. N2V Benchmark and Data

A) N2V benchmark was significantly faster when run on the LIV-SRF Cluster.

B / C) Full input (**B**) and denoised (**C**) image.

D / E) Detail of noisy input image (**D**). Detail of denoised images showing much clearer structures (**E**)



CLEMReg

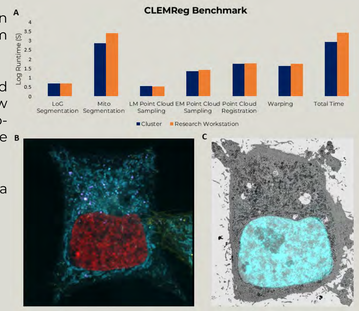
The CLEMReg benchmark was adapted from the napari-clemreg tool. We modified the Napari plugin with a CLI interface allowing us to run the workflow headless on the cluster. When tested our CLI tool yielded a much lower runtime compared to running the Napari plugin which allowed us to allocate more resources to the image processing. The segmentation step contributed significantly to the runtime discrepancy between the research workstation and cluster due to the lack of GPU on the researcher workstation.

Figure 7. CLEMReg Benchmark and Data

A) CLEMReg benchmark was significantly faster when run on the LIV-SRF Cluster. All image data used in this figure is from the original plugin publication (Krentzel et al., 2023).

B) Light microscopy image stained with 4 dyes (Red (Hoechst), Cyan (Mitotracker), Magenta (Lysotracker), Yellow (TGN46)), the image is split by colour channel and the Mitotracker channel is used to create a segmentation mask of the mitochondria across the Z-Stack.

C) Electron microscopy image with warped mitochondria overlay (cyan).



Conclusions

In the field of bioimage analysis a typical researcher workstation is no longer able to contend with the immense volume of data produced. The ever decreasing cost of server hardware makes HPC increasingly accessible and leveraging this compute power in a bioimaging setting is key to processing big data with minimal user involvement. The substantial performance gains showcased in our selected workflows highlights how HPC can expedite the processing of these big datasets. With our SLURMGUI web interface we enabled users with no experience using HPC computing to access these improved workflows and reap the rewards of decreased runtime and automatisation. Moving forward, it will be crucial to continue exploring and optimizing the use of HPC in bioimaging to further unlock its full potential.

References

KRENTZEL, D., ELPHICK, M., DOMART, M.-C., PEDDIE, C. J., LAINE, R. F., HENRIQUES, R., COLLINSON, L. M. & JONES, M. L. 2023. CLEM-Reg: An automated point cloud based registration algorithm for correlative light and volume electron microscopy. *bioRxiv*, 2023.05. 11.540445.

KRULL, A., BUCHHOLZ, T.-O. & JUG, F. Noise2void-learning denoising from single noisy images. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019. 2129-2137.

SCHMIDT, C., HANNE, J., MOORE, J., MEESTERS, C., FERRANDO-MAY, E. & WEIDTKAMP-PETERS, S. 2022. Research data management for bioimaging: the 2021 NFDI4BIOIMAGE community survey. *F1000Research*, 11.

STRINGER, C., WANG, T., MICHAELOS, M. & PACHITARIU, M. 2021. Cellpose: a generalist algorithm for cellular segmentation. *Nature methods*, 18, 100-106.

WERNERSSON, E., GELALI, E., GIRELLI, G., WANG, S., CASTILLO, D., LANGSETH, C. M., NGUYEN, H., CHATTORAJ, S., CASALS, A. M. & LUNDBERG, E. 2022. Deconwolf enables high-performance deconvolution of widefield fluorescence microscopy images.

WEI OUYANG, YW 2021. 'imjoy-team/imagej.js: ImageJ.JS v0.4.0'. Zenodo. doi: 10.5281/zenodo.4944985.

GitHub

Collaboration between CBF and CCI
Liverpool Shared Research Facilities

THE ORIGINAL REDBRICK

Enabling Heterogeneous Parallelism and Portability of Multiscale Universal Interface (MUI) library using SYCL

Mayank Kuma

UKRI-STFC

The poster outlines a successful initial implementation of SYCL to accelerate aspects of the MUI code coupling library targeting heterogeneous systems, underscoring the advantages of using this standard in the context of a complex coupled scientific computing workflow. We aim to provide insights into the practical aspects of developing and optimising SYCL-accelerated code and showcase the potential for wider adoption of this programming model in scientific and engineering applications. We demonstrate the practical application of the SYCL-based MUI library through case studies and benchmarks, showcasing its potential to enhance the parallel performance for coupling various scientific and engineering solvers. The poster serves as a resource for researchers and developers seeking to harness the power of SYCL to tackle the challenges of code coupling in parallel computing environments and heterogeneous systems.

Science and Technology Facilities Council

Scientific Computing

Enabling Heterogeneous Parallelism and Portability of Multiscale Universal Interface (MUI) library using SYCL

Mayank Kumar*[a], Wendi Liu[a], Matthias Kirchhart[b], Stephen M. Longshaw[a], Omar Mahfoze[a], Jayaraman Mahalingam[b]

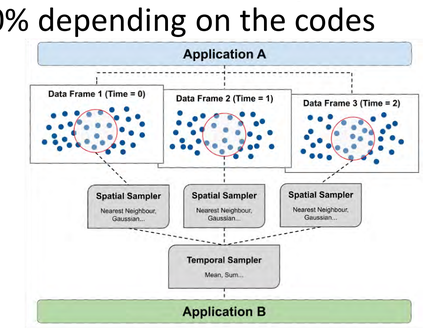
[a] UKRI STFC Scientific Computing, Daresbury Laboratory, Warrington, WA4 4AD

[b] Intel Corporation

Introduction

- A concurrent interface for coupling heterogeneous solvers[1,2]
- A header-only code-coupling library written in C++
- 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

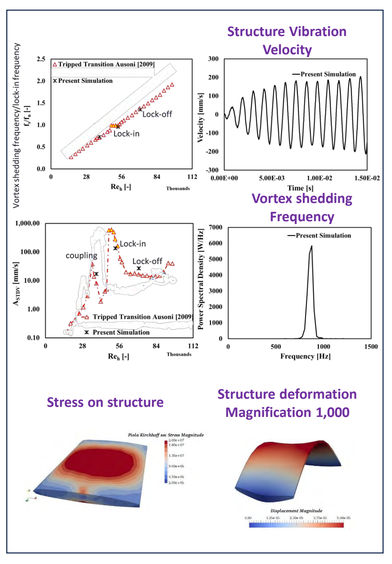
https://github.com/MxUI



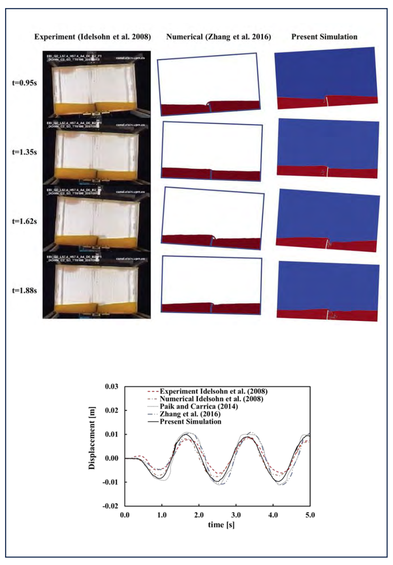
- Spatial Samplers
  - Exact
  - Linear
  - Quintic
  - RBF
- Temporal Samplers
  - Exact
  - Mean
  - Sum
- Coupling Algorithms
  - Fixed Relaxation
  - Aitken

Applications of MUI

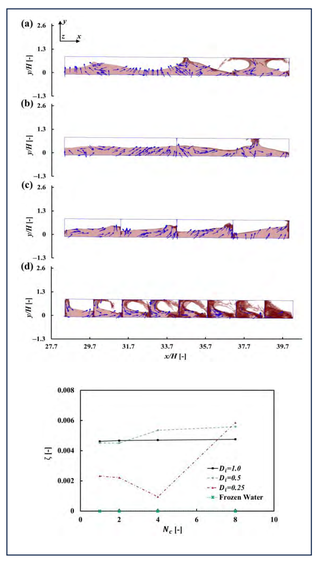
Vortex-induced vibration of the trailing edge of a hydrofoil[3]



2-D roll tank with flexible beam[4]



Slosh induced damping[5]



Radial Basis Function

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

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

- In MUI context (FSI Applications):

$$s_i = \sum_{j=1}^{M} H_{i,j}\, \alpha_j \quad \textbf{OR} \quad \boldsymbol{s = H\,\alpha}$$

$$\boldsymbol{H = A_{as}\, C_{ss}^{-1}}$$

$$\begin{aligned} \boldsymbol{H^T} &= (\boldsymbol{A_{as}\, C_{ss}^{-1}})^T \\ &= (\boldsymbol{C_{ss}^{-1}})^T\, \boldsymbol{A_{as}^T} \\ &= \boldsymbol{C_{ss}^{-1}\, A_{as}^T} \end{aligned}$$

$$\boldsymbol{C_{ss}\, H^T = A_{as}^T}$$

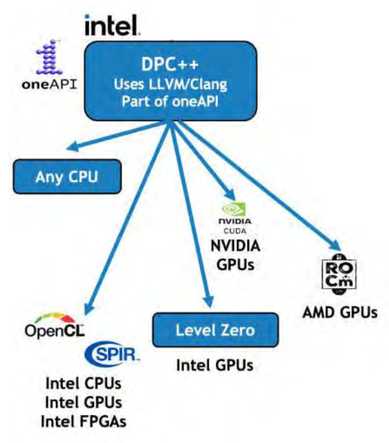
- Solvers
  - Conjugate Gradient
  - BiCGStab
  - Gauss Elimination
- Preconditioner
  - Diagonal
  - Incomplete Cholesky
  - Incomplete LU
  - SSOR

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 Implementation

Hotspots in MUI

- Generation of $\boldsymbol{C_{ss}}$ and $\boldsymbol{A_{as}}$ matrix
- Solution of $\boldsymbol{H^T}$

CG Algorithm[6]

$r_0 = b - Ax_0$

$p_0 = r_0$

$i=0$

while (r_i > tol)

{

$y_i = Ap_i$ — Matrix vector product

$\alpha = r_i.r_i \,/\, p_i.y_i$ — Inner Product

$x_{i+1} = x_i + \alpha\, p_i$ — Vector update

$r_{i+1} = r_i - \alpha\, y_i$

$\beta = r_{i+1} \,.\, r_{i+1} / r_i \,.\, r_i$

$p_{i+1} = \beta\, p_i + r_i$

$i = i+1$

}

Memory Allocations

SYCL USM

Allows reading and writing data using conventional pointers

Memory allocations can be explicit (device,host) or managed by SYCL runtime (shared)

Matrix vector product

```
template<typename ITYPE, typename VTYPE>
void sparse_matrix<ITYPE,VTYPE>::sycl_multiply_mat_vec(sycl::queue defaultQueue,
VTYPE *res_vec, VTYPE *mat_value, VTYPE *vec_value, ITYPE *res_col, ITYPE *res_row,
ITYPE *mat_column, ITYPE *mat_row,  ITYPE size_row)
{
    size_t rows = size_row;
    auto cag = [&](sycl::handler &ga)
    {
        res_row[0] = 0;
        ga.parallel_for(sycl::range(rows),[=](sycl::id<1>idx)
        {
            auto startIdx = mat_row[idx];
            auto endIdx = mat_row[idx+1];
            auto col_idx = 0;
            res_vec[idx] = 0.;
            for (int i = startIdx; i < endIdx; i++)
            {
                col_idx = mat_column[i];
                res_vec[idx] += mat_value[i]*vec_value[col_idx];
            }
            res_col[idx] = 0;
            res_row[idx+1] = idx;
        });
    };
    defaultQueue.submit(cag).wait();
}
```

SYCL Buffers

Container for data accessed by both host and device

Memory managed by SYCL runtime using accessors

Inner Product

```
template<typename ITYPE, typename VTYPE>
VTYPE sparse_matrix<ITYPE,VTYPE>::sycl_dotp_vec_vec(sycl::queue defaultQueue, VTYPE
*vec1_value, VTYPE *vec2_value,  ITYPE size_row)
{
    size_t rows = size_row;
    VTYPE *prod;
    prod = (VTYPE *)malloc(1);
    prod[0] = 0.;
    VTYPE *dotp;
    dotp = sycl::malloc_device<VTYPE>(1,defaultQueue);
    defaultQueue.memcpy(dotp, prod(sizeof(VTYPE))).wait();
    auto chg = [&](sycl::handler &hc)
    {
        hc.parallel_for(sycl::range(rows),[=](sycl::id<1> idx)
        {
            auto product = 0.;
            product = vec1_value[idx] * vec2_value[idx];
            auto v = sycl::atomic_ref<
                        VTYPE, sycl::memory_order::relaxed,
                        sycl::memory_scope::device,
                        sycl::access::address_space::global_space>(*dotp);
            v.fetch_add(product);
        });
    };
    defaultQueue.submit(chg).wait();

    defaultQueue.memcpy(prod,dotp,(sizeof(VTYPE))).wait();
    std::cout<<" Dot product value : "<< prod[0] <<std::endl;
    return (prod[0]);
}
```

Vector update

```
template<typename ITYPE, typename VTYPE>
void sparse_matrix<ITYPE,VTYPE>::sycl_set_y_axpby(sycl::queue
defaultQueue, VTYPE *Xvector, VTYPE *Yvector, VTYPE alpha, VTYPE
beta, size_t vec_size)
{
    size_t rows = vec_size;
    auto cag = [&](sycl::handler &ga)
    {
        auto Afac = alpha;
        auto Bfac = beta;
        ga.parallel_for(sycl::range(rows),[=](sycl::id<1>idx)
        {
            Yvector[idx] = Afac * Xvector[idx] + Bfac * Yvector[idx];
        });
    };
    defaultQueue.submit(cag).wait();
}
```

Diagonal Preconditioner

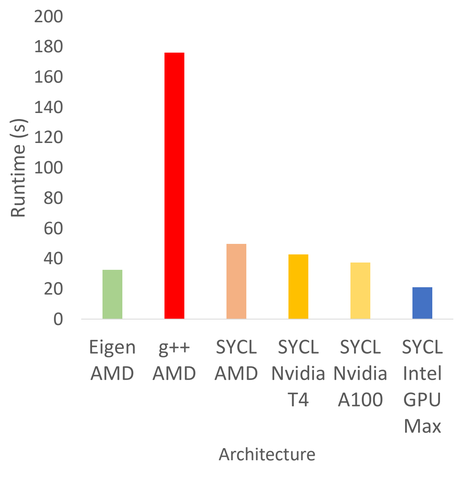
```
template<typename ITYPE, typename VTYPE>
void sparse_matrix<ITYPE,VTYPE>::sycl_populate_diag_vec(sycl::queue
defaultQueue, VTYPE *mat_value, VTYPE *vec_value, ITYPE *mat_row,
ITYPE *mat_col_idx, size_t size_row)
{
    size_t rows = size_row;
    auto cag = [&](sycl::handler &ga){
        ga.parallel_for(sycl::range(rows),[=](sycl::id<1>idx){
            auto startIdx = mat_row[idx];
            auto endIdx   = mat_row[idx+1];
            auto count = 0;
            for (int i=startIdx; i<endIdx;i++){
                if (mat_col_idx[i] == idx){
                    if (std::abs(mat_value[i]) >=
std::numeric_limits<VTYPE>::min()){
                    vec_value[idx] = 1.0/mat_value[i];}
                    else{
                        vec_value[idx] = 1.0;}
                    count++;
                    break;}}
            if (count == 0){
                vec_value[idx] = 1.0;
            }
        });
    };
    defaultQueue.submit(cag).wait();
}
```

Results

SYCL CG implementation is tested for solving a 576 x 576 matrix .

Performance compared with benchmark linear solver Eigen on AMD 32 core processor.

Same code executes on different architecture using SYCL USM shared memory.



Future work

Optimization of the linear solver.

Implementation of $\boldsymbol{C_{ss}}$ and $\boldsymbol{A_{as}}$ matrix generation using SYCL.

References

[1] Tang, Y.H., Kudo, S., Bian, X., Li, Z. and Karniadakis, G.E., 2015. Multiscale universal interface: a concurrent framework for coupling heterogeneous solvers. Journal of Computational Physics, 297, pp.13-31.

[2] Tang, Y.H., Kudo, S., Longshaw, S.M., Liu, W., Skillen, A., Mahfoze, O., Bian, X., Li, Z., Karniadakis, G.E., Nash, R.W., Richardson, C., Manson-Sawko, R., 2023. The Multiscale Universal Interface 2.0, https://doi.org/10.5281/zenodo.10054600

[3] W. Liu, W. Wang, A. Skillen, S. Longshaw, C. Moulinec, and D. Emerson. A parallel partitioned approach on fluid-structure interaction simulation using the multiscale universal interface coupling library. In 14th WCCMECCOMAS Congress, vol. 1400. 2021.

[4] Liu W, Longshaw SM, Skillen A, Emerson DR, Valente C, Gambioli F. A High-Performance Open-Source Solution for Multiphase Fluid-Structure Interaction. International Journal of Offshore and Polar Engineering. 2022 Mar 1;32(01):24-30.

[5] Liu W, Mahfoze OA, Longshaw SM, Skillen A, Emerson DR. Simulating Slosh Induced Damping, with Application to Aircraft Wing-like Structures. Applied Sciences. 2022; 12(17):8481. https://doi.org/10.3390/app12178481.

[6] Baratta, I., Richardson, C., & Wells, G. (2022). Performance analysis of matrix-free conjugate gradient kernels using SYCL. Apollo - University of Cambridge Repository. https://doi.org/10.17863/CAM.83190

Acknowledgements

This work was supported through the Intel Academic Program for oneAPI Centre’s of Excellence (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/academic-program/centers-of-excellence.html)

CIUK 2023 Breakout Sessions

ExCALIBUR Project Overview and Update: Organiser Phil Hasnip (University of York)

An overview of the Exascale Computing ALgorithms & Infrastructures Benefiting UK Research (ExCALIBUR) project. ExCALIBUR is a UK research programme that aims to deliver the next generation of high-performance simulation software for the highest-priority fields in UK research. It started in October 2019 and will run through until March 2025, redesigning high priority computer codes and algorithms to meet the demands of both advancing technology and UK research. Come along and hear abou the project and its progress so far.

Technical / SysAdmin Meetup: Organiser Simon Atack (Bristol Unversity)

A meeting for techncial staff and systems administrators to discuss the latest topics in the field and share information and ideas.

This was an open discussion forum for technical staff discuss topics of interest to them.

Discussions/Topics include: Slurm, Easybuild/Spack, GUIs, Openstack, future of xCat , Centos 7 Replacement

The beginning started with a short session about attributes needed for Federated Access, using myaccessid & puhuri.

Storage Scale User Group: Organiser Laurence Horrocks-Barlow (OCF)

Annual CIUK Storage Scale (GPFS/Spectrum Scale) usergroup.This breakout session is a community lead and influenced platform aimed towards providing Scale Storage updates user feedback and user case studies to the research, educational and scientific communities.Previous years have included product road maps, development q&a sessions and deep dives into the a range of use cases.The session aims to provide a informational and feedback platform for all users (new or experienced) of the product whether through direct or OEM licencing and appliances.In 2022 the current chair announced they were stepping down into a more assistive position due to work requirements, since this time OCF Limited have been assisting from an organisational point of view working with the community and Storage Scale providers to organise community events.

Examples, Challenges and Opportunities for Industrial use of High-Performance Computing and Scalable Artificial Intelligence: Organiser Richard Anderson (STFC Hartree Centre)

At the STFC Hartree Centre we specialise in developing and applying HPC and scalable AI related solutions to industry to increase productivity, to understand the behaviour of products (and systems) and to create solutions for the future. In this CIUK breakout session we aim to inform participants of the potential that can be achieved through the use of HPC and AI in different industry sectors. We will discuss case studies detailing the use of HPC and

AI for sectors such as materials, automotive, aerospace, energy and life sciences, for example. We will invite representatives of other centres such as DiRAC, University of Cambridge, EPCC, and University of Bristol to outline also their approaches to working with industry to give CIUK participants an overview of the potential for the uptake of HPC and AI in industry across the UK. We will describe the challenges brought to us, the solution's we applied and key learnings resulting from the completed projects. Following this, we will host discussion on the perceived challenges, opportunities, and routes to adoption associated with using HPC and AI for industry application. We aim to provide an environment to allow industry participants (or other likeminded and curious individuals) to ask questions about the work carried out to inform their own journey with these methods. Our goal is to enable participants to begin to understand what is possible and to develop a confidence that enables them to engage HPC and AI related activities for the future. Further to building the knowledge of industrial application, we will also discuss the barriers to adoption and implementation into a production environment. We would also like to obtain participant feedback from the session from those who have used HPC and AI in their own work and to provide a report back to the CIUK organisers who can use this information as a basis for informing and enabling improvement of the UK HPC infrastructure and support.

A potential agenda may look like the following:
* Introduction to the breakout (our aims goals and house rules)
* Industry sector case studies
* Challenges in applying HPC and AI to industry challenges and how to overcome (at least some) of these
* Opportunities in using HPC and AI
* General discussions and feedback from participants who have engaged with HPC and AI

CoSeC Annual Conference 2023

As part of the new CIUK Day Zero, the Computational Science Centre for Research Communities (CoSeC) ran its third 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 http://www.cosec.stfc.ac.uk

The day saw speakers present work from across the sciences on topics ranging from designing digital wave flumes through to understanding issues relating to software packaging and deployment. As part of the programme, a talk was given by invited speaker Prof. David Emerson from STFC on the history of the Collaborative Computational Projects (CCPs) and related CoSeC activities, which was then complimented by a presentation on the current and future progression of CoSeC. Each session was followed by a mini panel discussion involving each speaker and every session saw an engaged and lively conversation.

The day was concluded with a community feedback session chaired by Dr Martin Turner of the University of Manchester which looked at different aspects of the CoSeC landscape in the context of an evolving UK landscape around computational science. As a closing point, the director of STFC's Scientific Computing Department, Tom Griffin, took the opportunity to announce the appointment of STFC's Dr Stephen Longshaw as the new director of the CoSeC programme.

The event was a huge success, building well on the previous years, community feedback was to re-run the event again in 2024 as part of CIUK so please join us again in Manchester for the next event.

CIUK 2023 Women in HPC Breakfast

Introduction: The CIUK'23 Women in HPC Breakfast provided a platform for insights into the history of women in engineering and valuable contributions from senior female leaders and allies across the UK and Europe. The event, supported by UKRI STFC and Alces Flight, fostered engaging discussions on HPC over breakfast in Manchester.

Keynote Address: Teresa Schofield, CIUK 2023 WHPC Distinguished Speaker Teresa Schofield, our distinguished speaker, shared her journey as an engineer in semiconductors, offering candid reflections on the challenges and triumphs in her career. While her career was not directly within the field of HPC, Teresa's insights resonated with attendees, emphasizing the importance of workforce inclusion and investment in the field. Her remarks underscored the significance of resilience and commitment in navigating the complexities of engineering and HPC.

Lyceum Panel Discussion: The breakfast session continued with a dynamic panel discussion, the Lyceum, featuring Melyssa Fratkin (TACC), Rosemary Francis (Altair), and Emma Barnes (University of York). Attendees engaged in conversations ranging from mentorship to the role of a PhD in career progression, fostering an environment of shared learning and collaboration.

Announcement: 'Move the Needle' Project The highlight of the breakfast was the unveiling of 'Move the Needle'—a project aimed at tracking goals for workforce inclusion in HPC. This is a project that will be running throughout 2024, with a call for global participation and continued engagement in promoting diversity and inclusion in the field.

Conclusion: The CIUK'23 Women in HPC Breakfast served as a platform for meaningful dialogue, inspiration, and commitment to advancing diversity and inclusion in HPC. As we look towards the future, we remain dedicated to driving positive change and empowering all individuals in the HPC community.

Proceedings compiled by Cristin Merritt - Alces Flight


COMPUTING
INSIGHT UK 2024
UKRI
Science and
Technology
Facilities Council
SAVE
THE
DATE
5-6 December 2024
Manchester Central, UK
SEE YOU IN 2024
www.ukri.org/CIUK