UK Research and Innovation

Session Overview

09:30 - 09:50 **Project Overview** (Martin Juckes)
09:50 - 10:00 **HPC-JEEP** (Alastair Basden and Andy Turner)
10:00 - 10:10 **IRISCAST** (Jonathan Hays)
10:10 - 10:20 **ENERGETIC** (Deepayan Bhowmik and Teymoor Ali)
10:20 - 10:30 **CARBON-QUANDRI** (Daniel Schien)
10:30 - 11:00 **Panel Discussion**
(Chair: Ag Stephens; Panel: Martin Juckes, Wim Vanderbauwhede, Justin O'Byrne)

NET ZERO DIGITAL RESEARCH INFRASTRUCTURE (DRI)

Roadmap for UKRI to reach Net Zero by 2040

Goal: actions to reduce the carbon emissions from data generation, analysis, storage and dissemination.

PLEASE HELP if you are the manager, supervisor or contact of a UKRI-owned/majority-funded "facility". **COMPLETE OUR SURVEY** to help us map the carbon landscape of the UKRI DRI.

UK Research and Innovation

https://bit.ly/netzerodri

https://net-zero-dri.ceda.ac.uk

support@ceda.ac.uk

UK Research and Innovation

UKRI Net Zero Digital Research Infrastructure Scoping Project

https://net-zero-dri.ceda.ac.uk/

Martin Juckes, Charlotte Pascoe, Ag Stephens, Poppy Townsend, Katie Cartmell, Jen Bulpett

CIUK, Manchester, Friday 1st December 2022

Project Ambition

- Collect **evidence to inform** UKRI Digital Research Infrastructure (DRI) Investment decisions
- Provide UKRI and their community with an **outline roadmap for achieving** carbon neutrality in their DRI by 2040 or sooner
- Enable UKRI to play a **positive and leading role** in the national and global transition to a sustainable economy

UK Research and Innovation

Who we are

Scoping project (£1.8m) – ending in Summer 2023

- Core project team – CEDA/NCAS
- Science Advisory Board (Prof Mary E Black)
- Steering Committee (Anna Angus-Smyth - NERC)
- Project partners – various universities/UKRI councils are undertaking some work

Stephen Mobbs
PI

Martin Juckes
Project Lead

Poppy Townsend
Communications Manager

Jen Bulpett
Senior Project Manager

Charlotte Pascoe
Science Officer

Ag Stephens
Technical Officer

Katie Cartmell
Project Manager

Centre for Environmental Data Analysis
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Partners

The core team is supported by partners from 20 institutions, bringing a huge range of experience.

What is the DRI?

- The UKRI DRI is the UKRI owned and majority funded Digital Research Infrastructure

- In practice, it does not matter whether a facility is 40%, 60% or 100% funded by UKRI: this project is focused on gathering evidence to support those who want to reduce the net emissions of digital research infrastructure to zero.

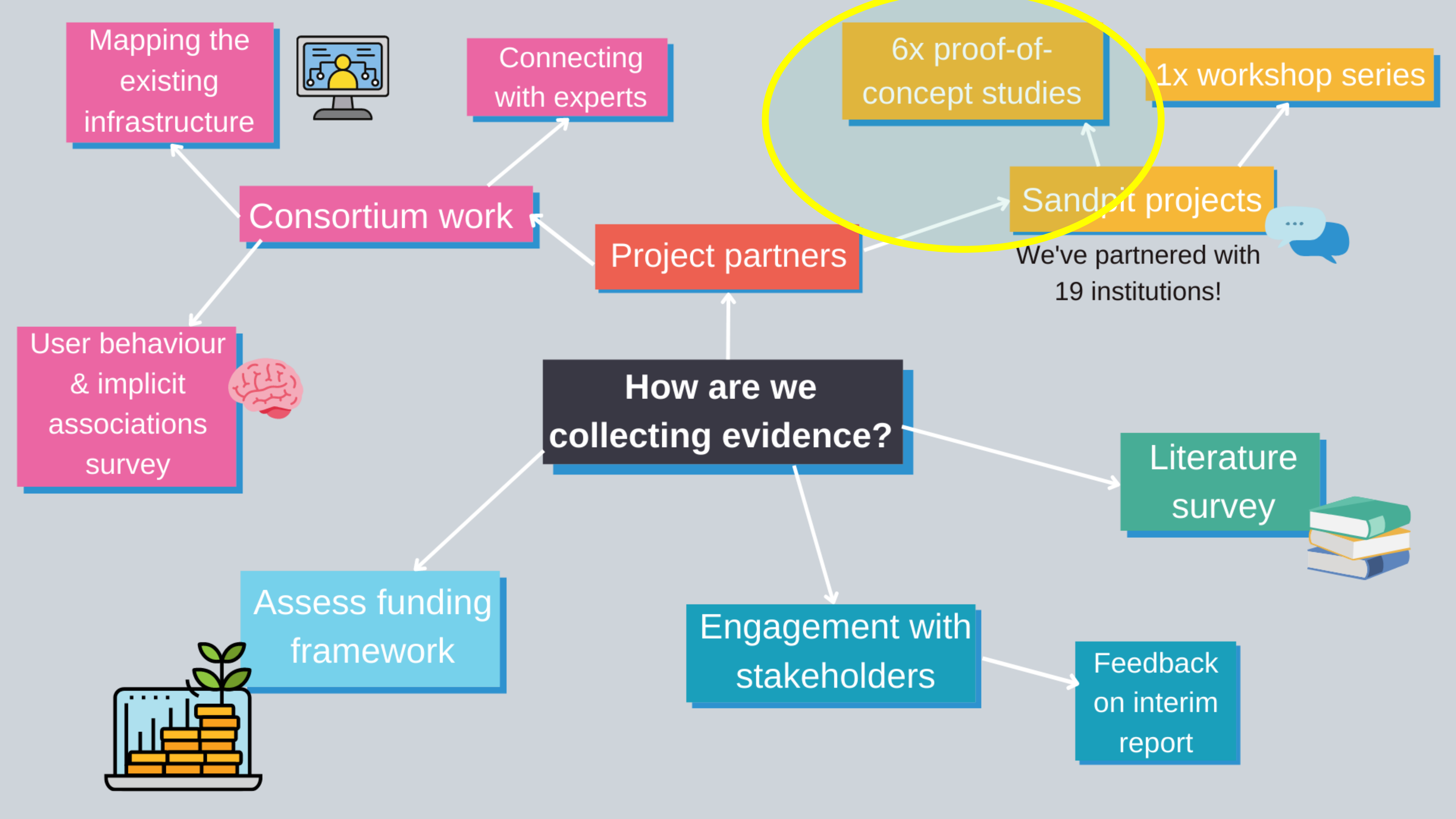
UK Research and Innovation

Some Carbon Basics

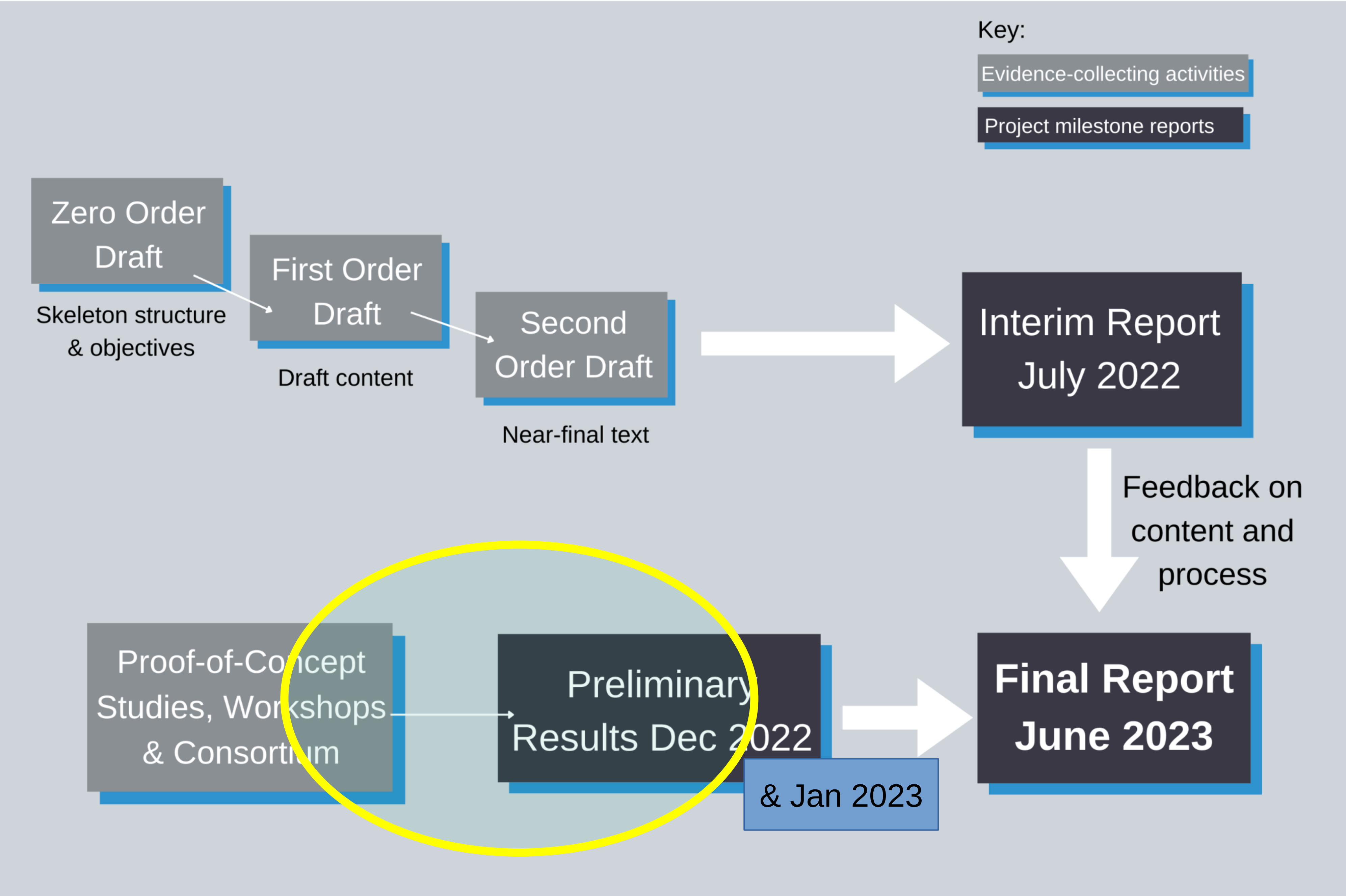
- The carbon footprint is dominated by power supply and manufacturing.
- Carbon offsets do not increase costs much, but does not work very well, if at all.
- Carbon sequestration appears to work, but could double costs. Sequestration costs could go down as technology increases or go up if/when demand outstrips supply.
- Drastic reductions in emissions are needed; different approaches will be needed for power supply and manufacturing.
- In parallel, we need to ensure that we make best possible use of resources, so we are not paying for power consumption which can be avoided.

Carbon Budget Challenges

- Multiple metrics: there are 3 measures of the carbon footprint of electricity supply:
  - **National Carbon Intensity** : national annual average
  - **Purchase Carbon Intensity** : reflecting financial flows
  - **Grid Carbon Intensity** : reflecting the actual local flow of power, including high use of fossil fuels at times of peak demand
- Users do not have clear information about the footprint of their work, so little incentive to improve efficiency.
- The carbon footprint of manufacturing is large, but very poorly quantified.
- **The majority of the carbon footprint is tied up in activities which are outside the direct control of the DRI stakeholders, such as institutional electricity supply and procurement rules**


How are we collecting evidence?
Project partners
Consortium work
Mapping the existing infrastructure
Connecting with experts
User behaviour & implicit associations survey
Sandpit projects
6x proof-of-concept studies
1x workshop series
We've partnered with 19 institutions!
Literature survey
Assess funding framework
Engagement with stakeholders
Feedback on interim report

Timeline for key project outputs



UK Research and Innovation

Sandpit events

Sandpit A - 9th and 11th May: community and organisational challenges
Sandpit B - 23rd and 25th May: technical and operational challenges

- Each sandpit consisted of two 3-hour online sessions (Monday and Wednesday):
  - setting the scene, meeting each other, exploring ideas.
  - forming teams, defining objectives
- Short proposals were submitted on Thursday and evaluated by a panel on Friday
- 7 projects funded, 4 on technical and operational challenges (presenting today) and 3 on community and organisational challenges.

UK Research and Innovation

HPC JEEP
Job Efficiency & Energy Usage

IRISCAST
Audit of Carbon Costs

ENERGETIC
Testing Heterogeneous Architecture

CQUANDRI
Quantifying Carbon Emissions

The UKRI Net Zero DRI Scoping project contains 9 consortium projects (right) and has funded 7 additional projects (left) through a sandpit event.

These projects will investigate a broad range of technical and social issues related to the Net Zero target.

ARCHER 2
In-depth case study

JASMIN (SUPERCOMPUTER)
In-depth case study

SUSTAINABLE COMPUTING
Providing expertise

SCAFELL PIKE (SUPERCOMPUTER)
In-depth case study

MAPPING THE DRI
Survey of UKRI DRI Estate

JADE (AI FOCUS)
In-depth case study

Machines and Workflows

People and Process

ARINZRIT
Identifying socio-technical forces

GO ZERO
A Net Zero Workshop Series

VALUE
Net Zero Decision-Making

ROADMAP DEVELOPMENT
Providing expertise

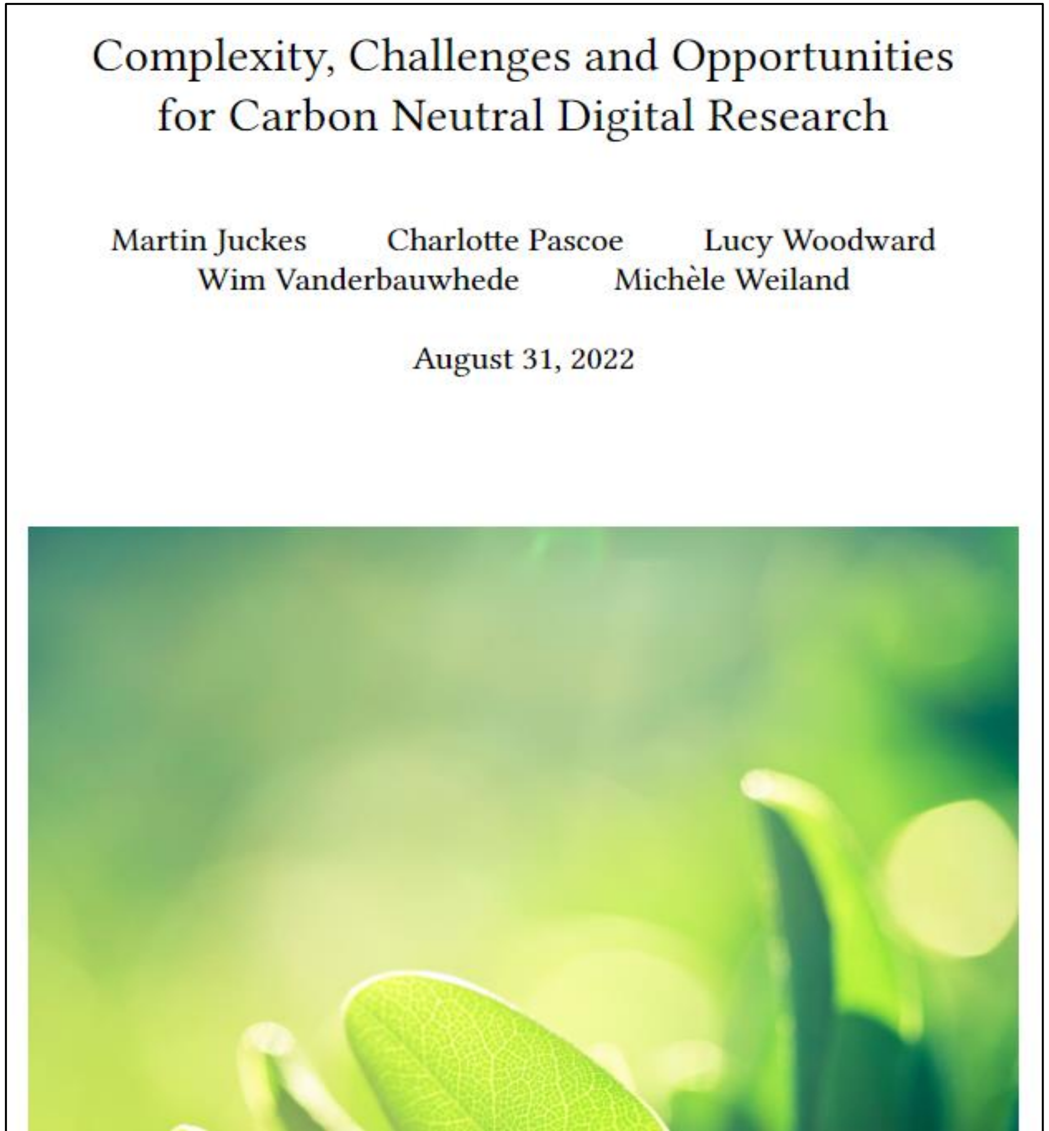
ARTIST COMMISSION
Art for collective action on sustainable DRI

DRI USER BEHAVIOUR SURVEY
Implicit Associations towards sustainable DRI

The interim report : published August 2022

- Initial findings based on a literature survey and stakeholder engagement

bit.ly/nzdri_interim


Complexity, Challenges and Opportunities for Carbon Neutral Digital Research

Martin Juckes Charlotte Pascoe Lucy Woodward Wim Vanderbauwhede Michèle Weiland

August 31, 2022

UK Research and Innovation

Selected Interim Recommendations

- Build consensus, lead by creating a space for ideas, adopt best practice
- Use multi-year contracts for electricity supply; exploit on-site generation and storage
- Use contracts and other lines of influence to reduce carbon intensity of supply chain
- Invest in people to improve efficiency of resource use
- Develop policies to ensure that efficiency leads to lower carbon footprint rather than all going to higher throughput
- Reduce emissions as much as possible and explore all options for removing carbon from the atmosphere
- **Improve quantification of the immense societal benefit delivered by the UKRI DRI in parallel to improving quantification of carbon footprint**

Coming soon

- Sandpit project final reports: January 2023
- Stakeholder workshops: February 2023
- Cross-UKRI Workshop : May 2023
  (Showcasing findings; review recommendations)
- Publication of project conclusions: June 2023

https://net-zero-dri.ceda.ac.uk/
support@ceda.ac.uk

UK Research and Innovation

Thank you

Get in touch: support@ceda.ac.uk

HPC JEEP

HPC Job Efficiency and Energy Profiling

CIUK
December 2022
Andy Turner (EPCC)
Alastair Basden (Durham/DiRAC)

epcc

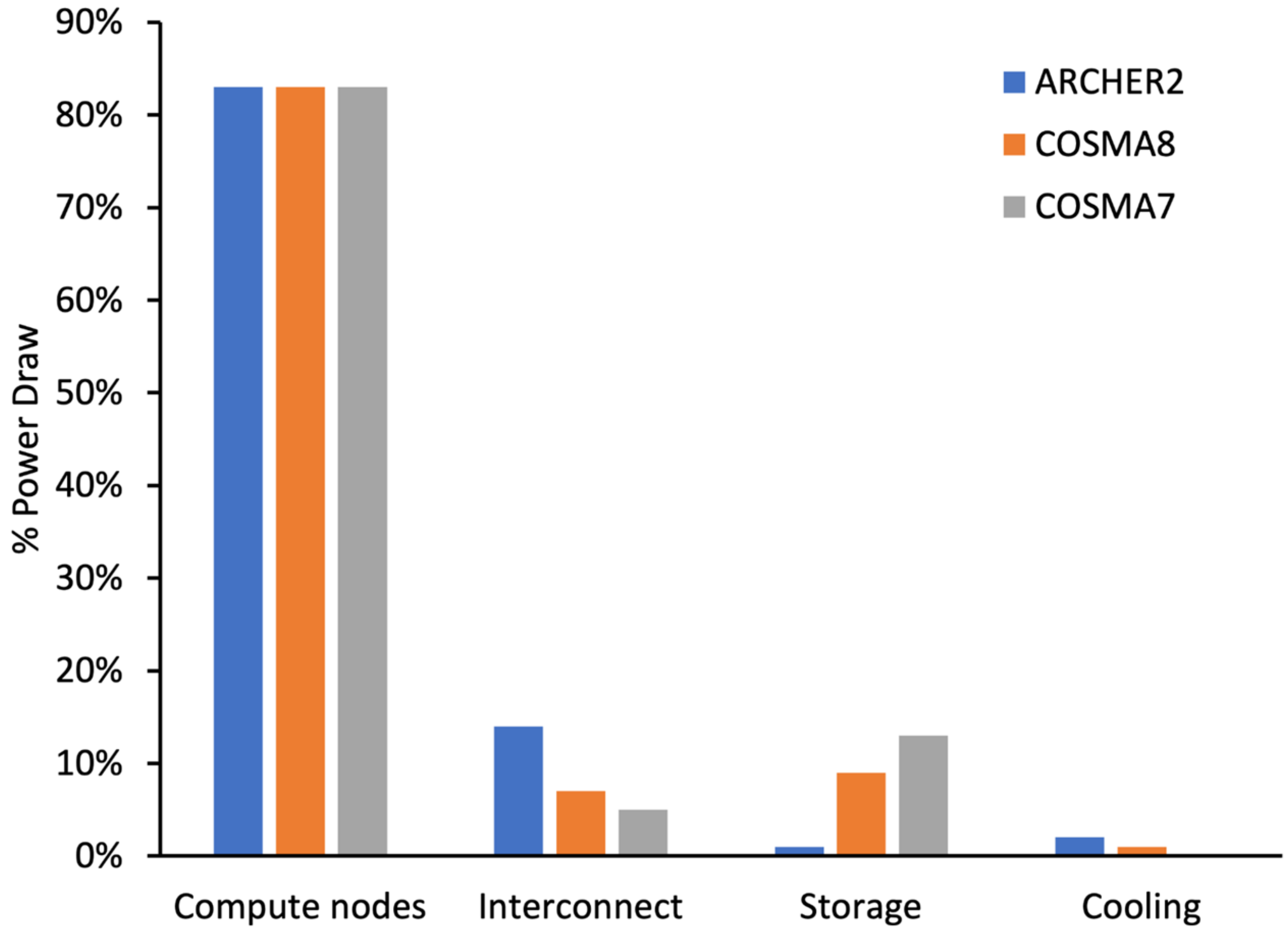
Durham University

UK Research and Innovation

HPC-JEEP scoping project aims

- Understand what per-job energy data we currently have from HPC systems and what types of analyses this data can support to help transition towards net zero
  - Report: https://doi.org/10.5281/zenodo.7128628
- Understand if the energy use data can potentially support introduction of energy-based charging
- Propose energy (and, potentially, emissions) metrics that can be provided back to HPC service stakeholders to help them transition towards net zero

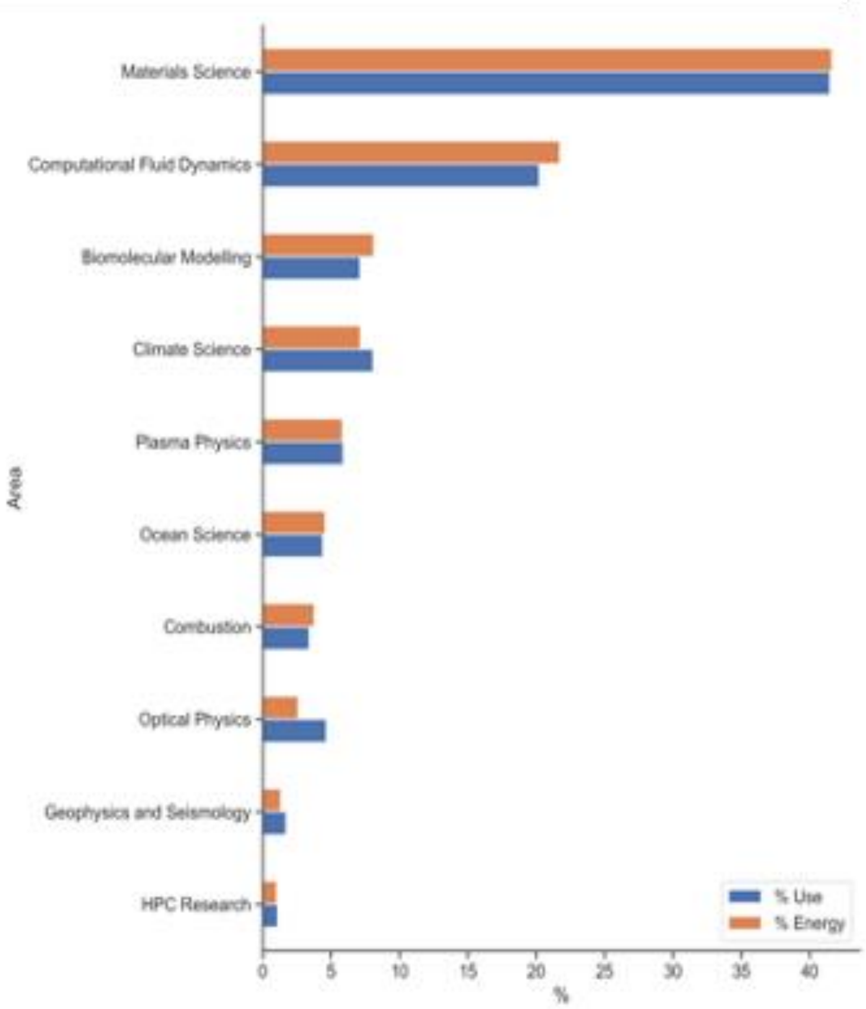
Power use on ARCHER2/COSMA systems

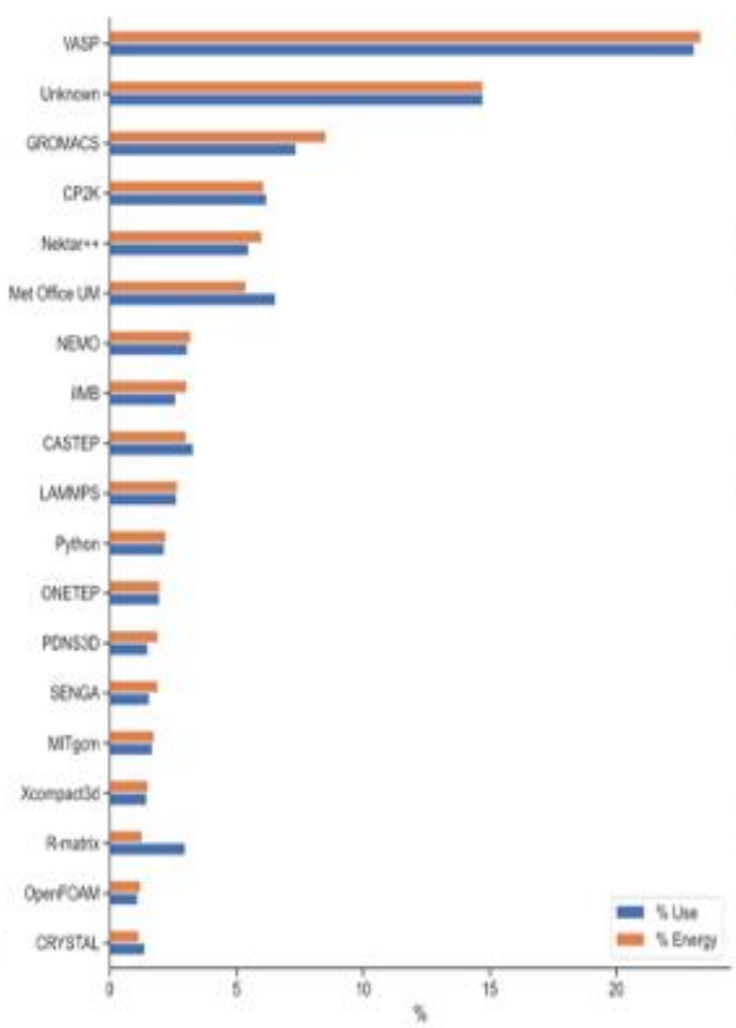


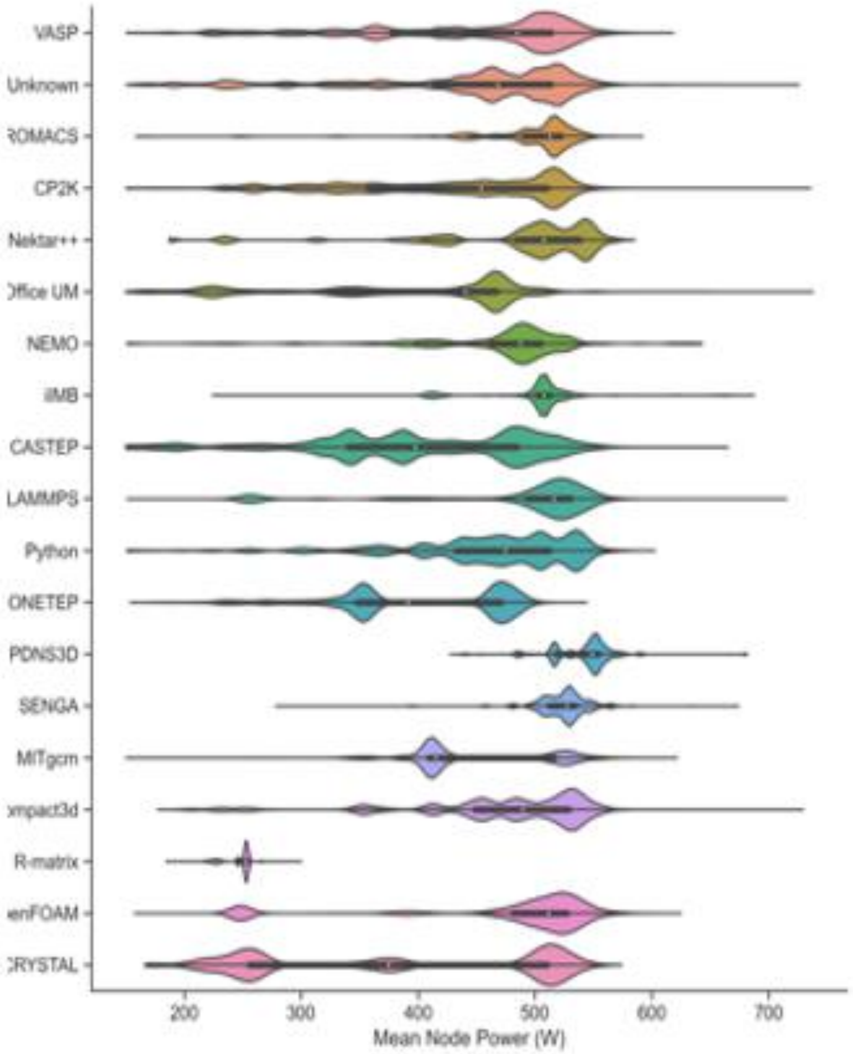
- ARCHER2 has lower storage capacity per node than COSMA
- Interconnect differences could be real or due to different vendor measurement methodologies
- Seems coincidental that the compute node numbers are all so similar
- Only includes "in-cabinet" components
    - Cooling is cabinet CDU, rather than plant rooms

Analysing ARCHER2 energy data

Methodology and tools at:
https://doi.org/10.5281/zenodo.7128628







Note: this is only compute node

Energy-based charging on ARCHER2

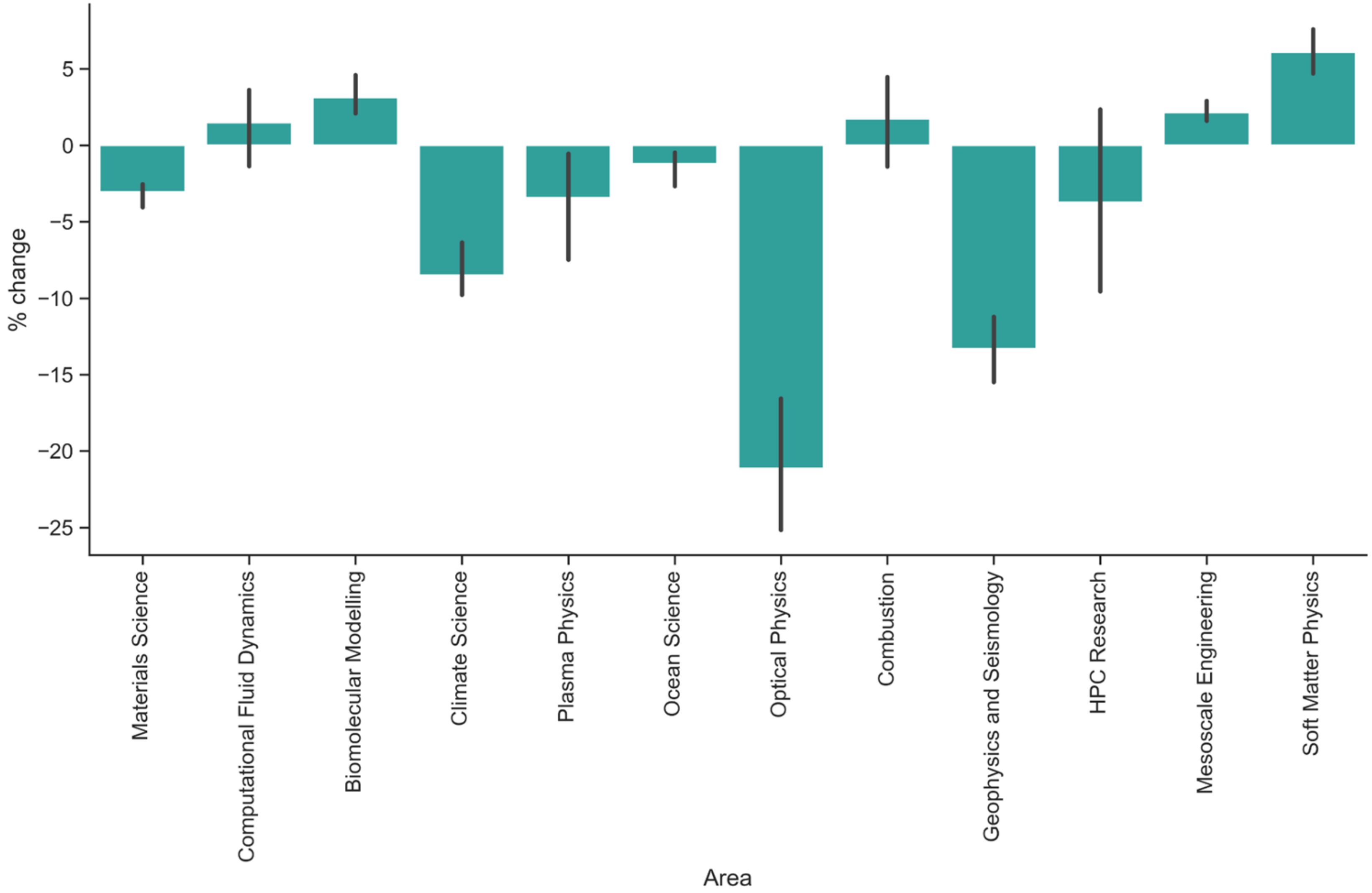
Charging based on:

50% Residency - how many nodes you have for how long

50% Energy - how much energy the job uses

Comparison is to 100% residency charge (nodeh) for 3 month period. Bars indicate range of monthly variation.

Overall reduction in total charge by 3% - corresponds to an overall 3% allocation boost unless allocations are updated.



Embodied energy - embodied CO_2

- Building an HPC system embeds CO_2 produced during production
  - HPC systems are often produced in countries with high carbon intensity
- COSMA7 compute:
  - Dell C6420 servers: 1,240 $kgCO_2$ (according to Dell)
  - In production for ~4 years (so far)
  - 452 nodes, total energy consumption ~3200 MWh (7 MWh/server)
  - Average CO_2 intensity in North East over previous 13 months ~38 gCO_2/kWh
    - 269 $kgCO_2$ per server (over the 4 year lifetime), 67 $kgCO_2$ per year:
      - 18 years for production CO_2 to equal embodied CO_2
    - Embodied % for a 4 year lifetime is ~80%
    - This will only increase as the UK national grid greens
  - Note: Calculation different depending on CO_2 intensity.
    - E.g East Midlands, 280 gCO_2/kWh: 500 $kgCO_2$/year, 2.5 year payback
    - Embodied % for a 4 year lifetime is ~40%

UK Research and Innovation

Embodied CO_2 notes

- A per-region approach is not necessarily valid
  - UK has a national grid
  - Average national CO_2 intensity over past ~3 months is ~180g / kWh
  - So, 320kg CO_2/year from COSMA7 nodes
    - 4 years operation for embodied CO_2 to equal compute production CO_2
      - Will increase in future years (assuming embodied CO_2 doesn't change)
- How long should we be running systems for?
  - 4 years means CO_2 is ~50% embodied
  - 8 years seems reasonable (though obviously, many factors)
    - 33% of CO_2 produced will be due to the embodied part
    - Probably longer than we currently do!
- Important to push suppliers for lower embodied energy

User Reporting

- Quarterly emails sent to COSMA users and project PIs
  - Total energy used by their jobs for each user
    - Compute node
    - Estimate of fraction of storage/fabric
    - Carbon intensity value over that period
    - Mass of CO_2 generated
    - Some context (flights, miles driven, household usage, etc)
  - Total energy used by each project
    - And a list of largest users
- In future, UKRI may charge by kWh rather than core-hour
  - Helps to advise on how much to apply for
  - Benefit to making codes more efficient

UK Research and Innovation

Summary

- Providing users with a summary of their compute CO_2
- Providing UKRI with recommendations for future systems
  - Both procurement and operation

UK Research and Innovation

IRISCAST: IRIS Carbon Audit Snaphot

J. Hays – IRIS Science Director
IRISCAST Project PI

CIUK 2022 – UKRI NetZero Scoping Project
2nd December 2022

UK Research and Innovation

eInfrastructure for Research and Innovation for STFC

IRIS is a cooperative community bringing together (mainly) STFC computing interests

Formed bottom up by science communities and compute providers

Works closely with STFC but run by the community

UK Research and Innovation

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

Challenges/questions

Estimating the carbon costs for scientific computing across a broad heterogeneous landscape

Identifying the key drivers

Identifying the hurdles and barriers

Communicating the costs to drive change

Working coherently across different communities

Actions and Objectives

Work together coherently across different facilities with different remits, tooling, and capabilities.

Learn by doing!

Document the gaps, the barriers and the issues, drive requirements for future work and decision making

Communicate across our communities and build a foundation for future action

Good robust decisions need good robust information

IRISCAST is a 6 month project funded within the UKRI Net Zero Scoping Project

Project Team

Alison Packer (STFC)
Anish Mudaraddi (STFC)
Derek Ross (STFC)
Dan Traynor (QMUL)
Jon Hays (QMUL)
Alex Owen (QMUL)
Dan Whitehouse (Imperial)
Adrian Jackson (Edinburgh)
Alastair Basden (Durham)
Nic Walton (Cambridge)
Alex Ogden (Cambridge)

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

Facilities

QMUL GridPP Tier 2
Imperial GridPP Tier 2
STFC SCD Cloud

STFC SCARF
DiRAC (Durham)
Cambridge IRIS
HPC/Cloud

UK Research and Innovation

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

Inventory

Data collection

Analysis

Community Engagement

UK Research and Innovation

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

DONE

Inventory
- Define the scope of the audit
- Build a comprehensive list of all equipment covered by the audit
- Needed to build carbon model including embodied costs

UK Research and Innovation

Good robust decisions need good robust information

DONE(ish)

Data Collection

- Collect data over a 24 hour period covering differing operating conditions
  - Rack, Node, and Job level logging
- Store data in central repository

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

In progress

Analysis

- Integrate the different datasets into coherent curated data set
- Refine carbon model
- Extract insights, observations, and conclusions

UK Research and Innovation

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

In progress

Community Engagement

- Talk at CIUK
- Produce draft report
- Publish curated data set and definition of the carbon modelling
- Engage with our communities through an IRIS Workshop – 6th, 7th January in Cambridge

UK Research and Innovation

IRIS-CAST – The **C**arbon costing for computing **A**udit **S**napsho**T**

Good robust decisions need good robust information

Next steps…

Inventory – DONE

Data collection – DONE(ish)

Analysis – In progress

Community Engagement – In progress

Energy-aware Heterogeneous Computing at Scale (ENERGETIC)

Teymoor Ali & Deepayan Bhowmik

Newcastle University

Motivation/Aim

- Current HPCs consist of various combinations of accelerators CPUs, GPUs and FPGAs.
- Little data on the energy efficiency of codes or algorithms across different architectures
  - No established framework or methodologies.
  - Little use of existing tools.
- Project Aim: whether the use of heterogeneous architecture could significantly reduce the energy-to-solution.

Benchmark Algorithms

- Selected HPC Challenge benchmarks
  - Single Precision General Matrix Multiplication (SGEMM)
  - 2D Fast Fourier Transform (FFT)
  - STREAM (Main Memory Bandwidth)
- And also deep learning based computer vision (CNN)
  - Still under processing

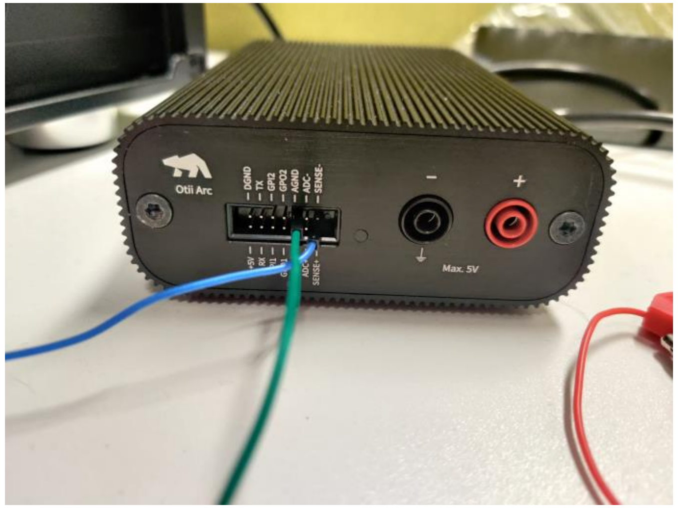
Benchmark systems

- Standalone Heterogenous System
  - CPU: i9-11900KF
  - GPU: Nvidia A2000
  - FPGA: Xilinx Alveo U50

- High Performance Clusters:
  - EPCC FPGA Test Bed
    - FPGA: Xilinx Alveo U280
  - Myriad
    - CPU: Xeon Gold 6240 CPU
    - GPU: NVIDIA A100

UK Research and Innovation

Measurement Approach - Datalogger



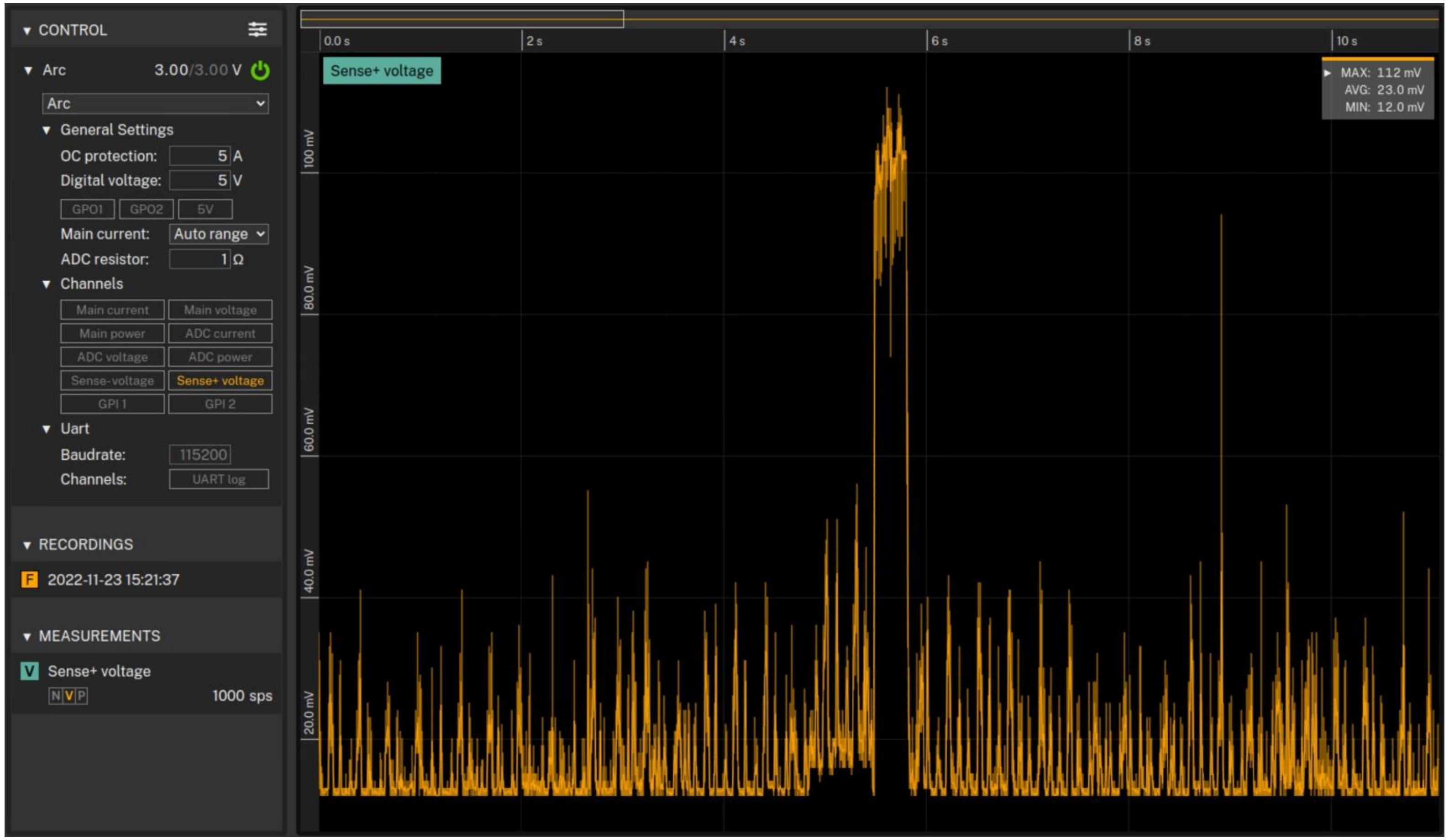
Hardware:

- Current Clamp
- Otii arc 3 datalogger
- Multimeter

Power Measurement Software:

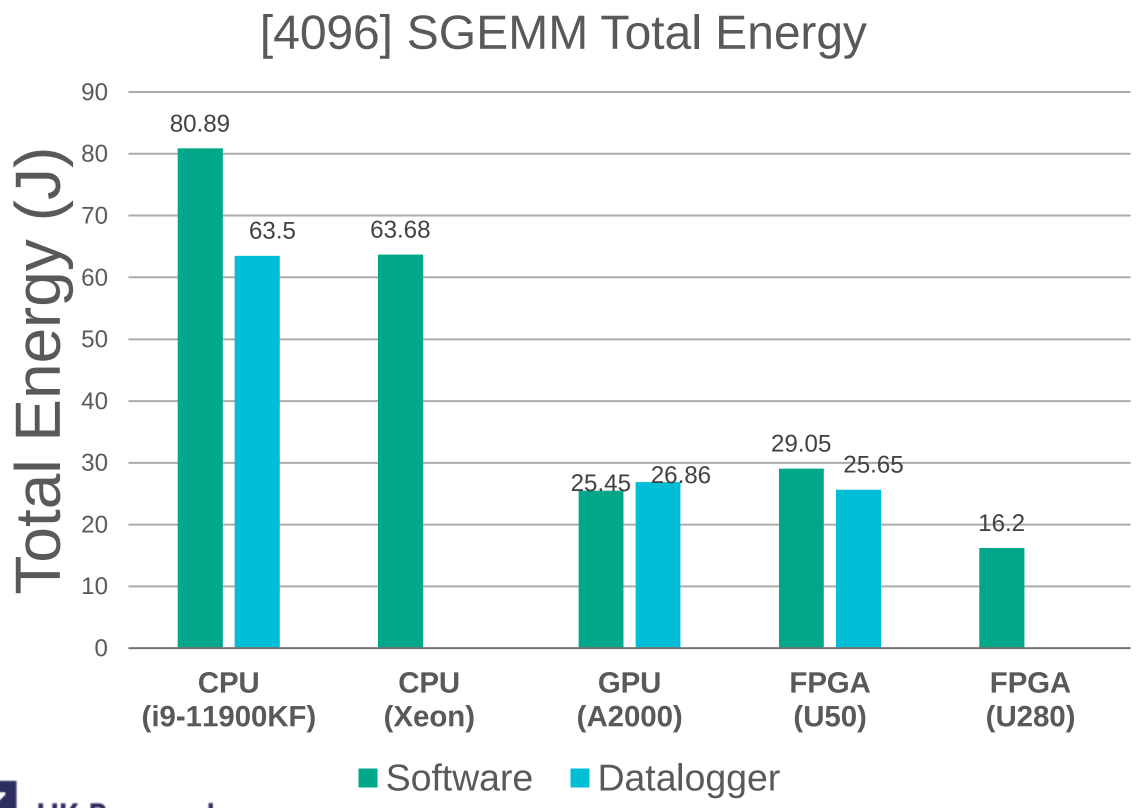
- CPU (RAPL)
- GPU (NVML)
- FPGA (XBUTIL)

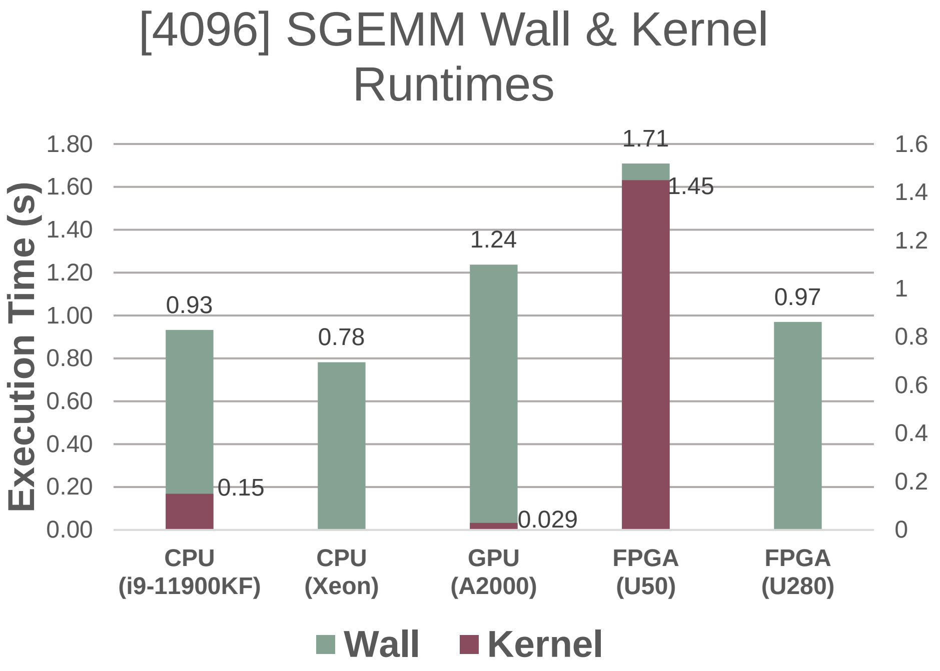
Datalogger Software



UK Research and Innovation

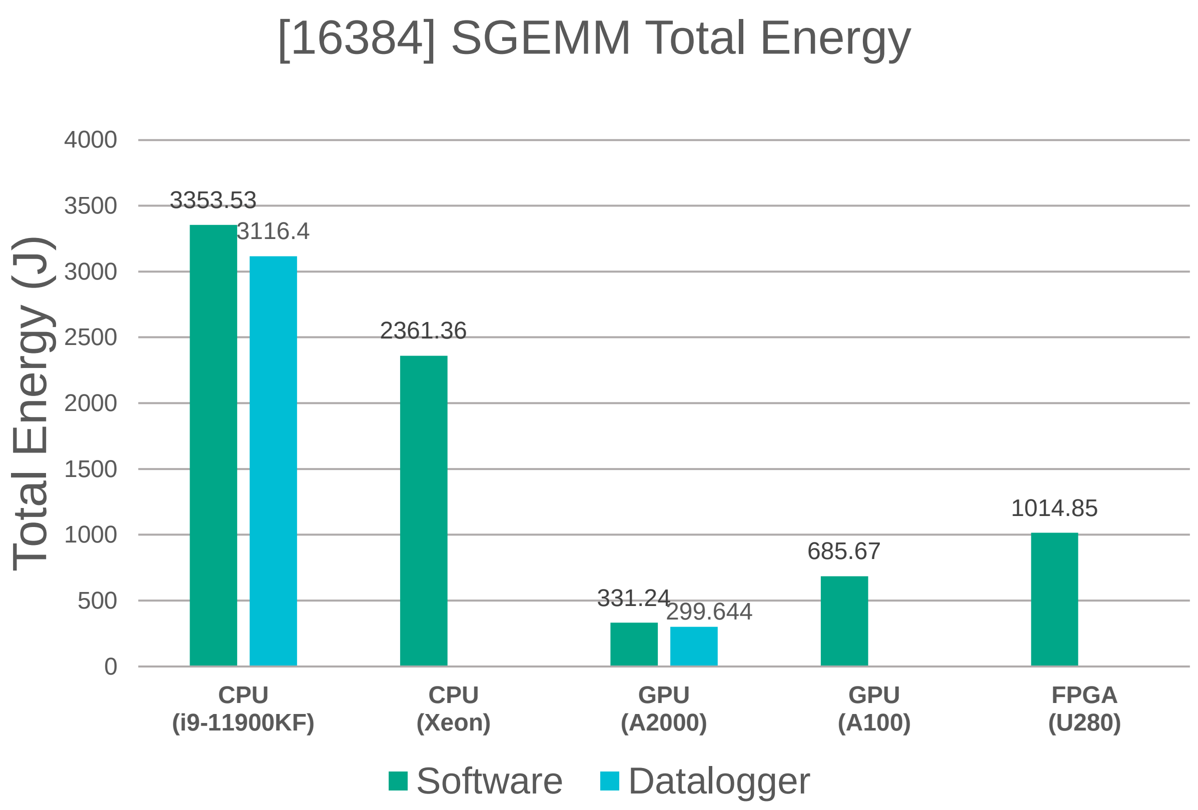
Results: SGEMM [4096]

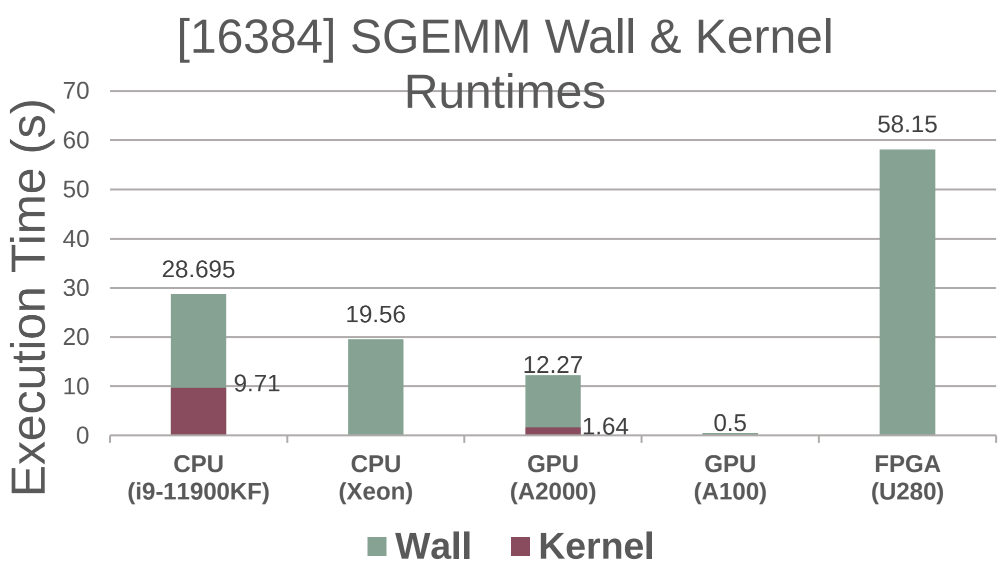




UK Research and Innovation

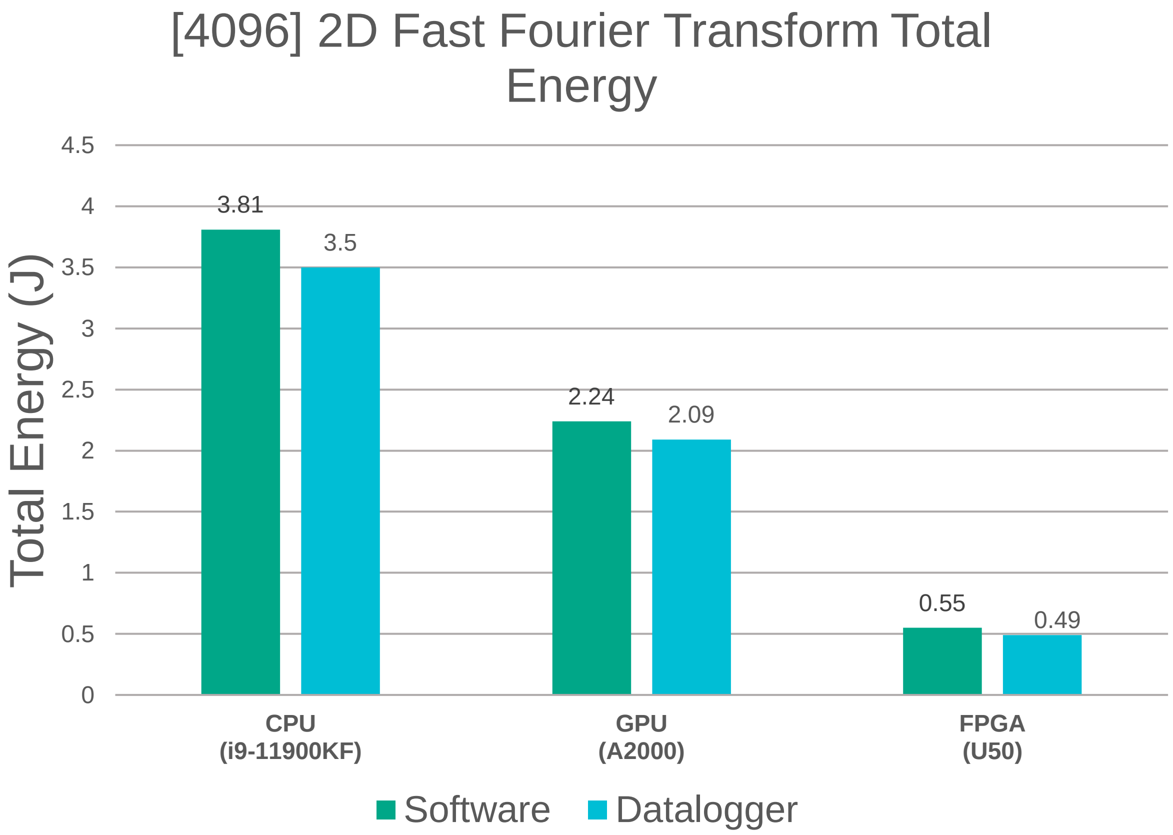
Results: SGEMM[16384]

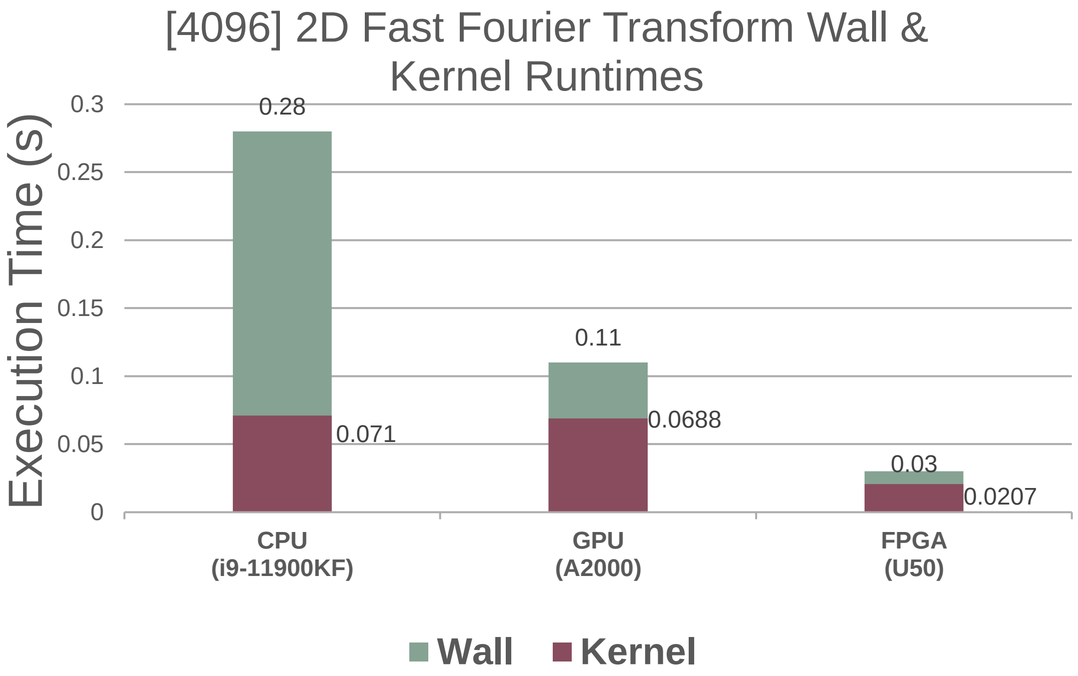




UK Research and Innovation

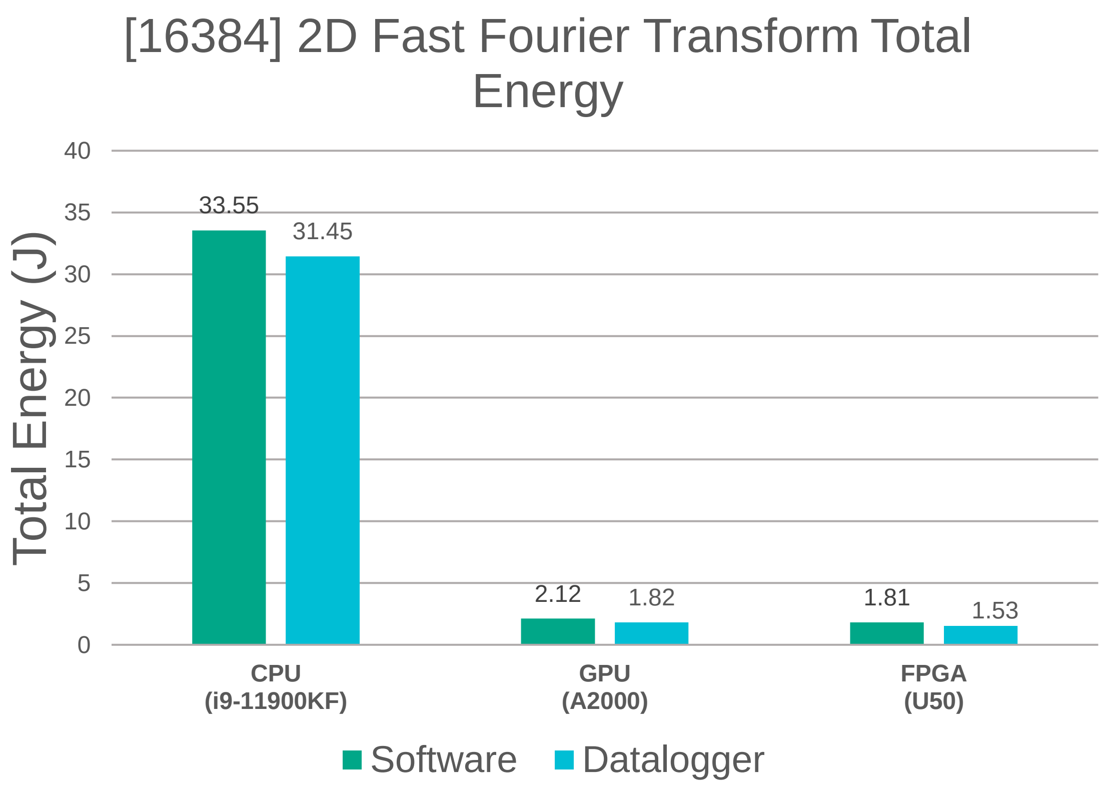
Results:2D FFT [4096]

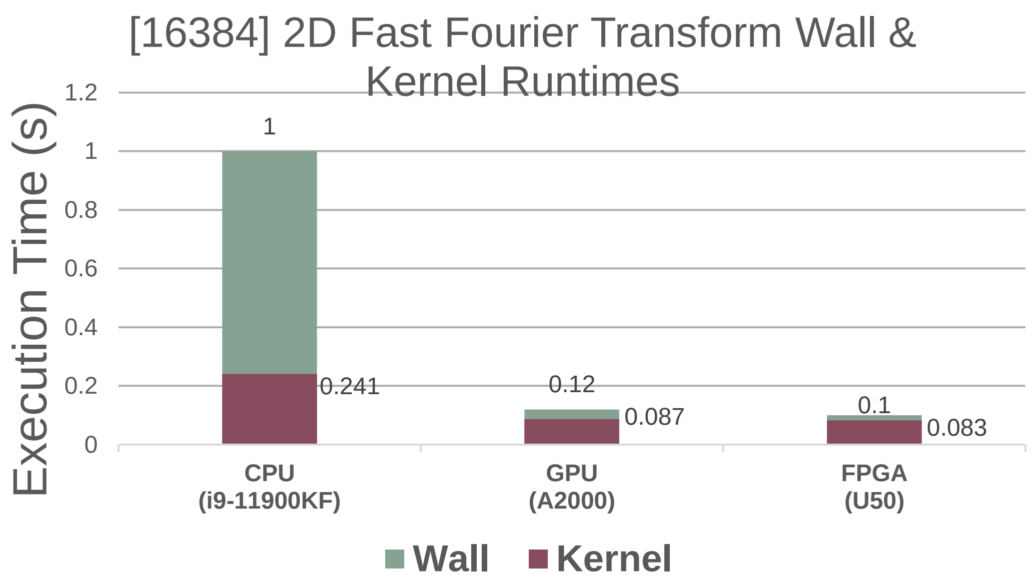




UK Research and Innovation

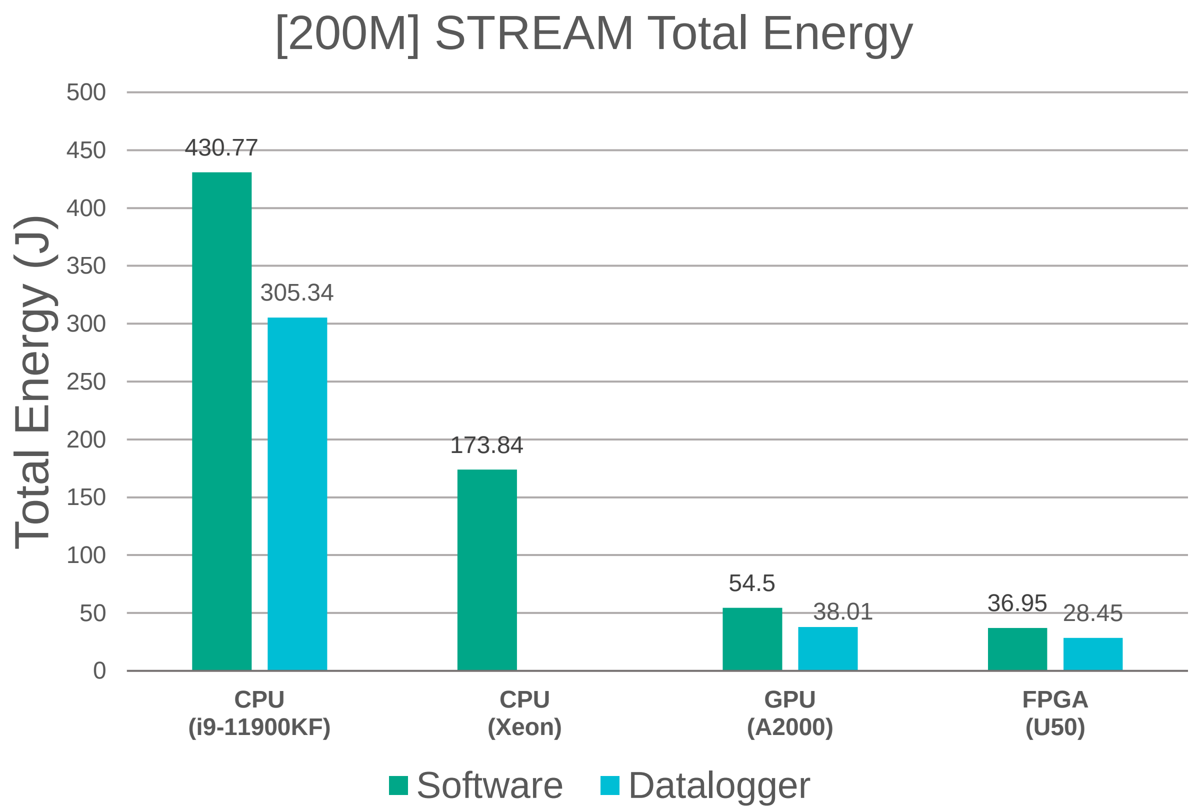
Results:2D FFT [16384]

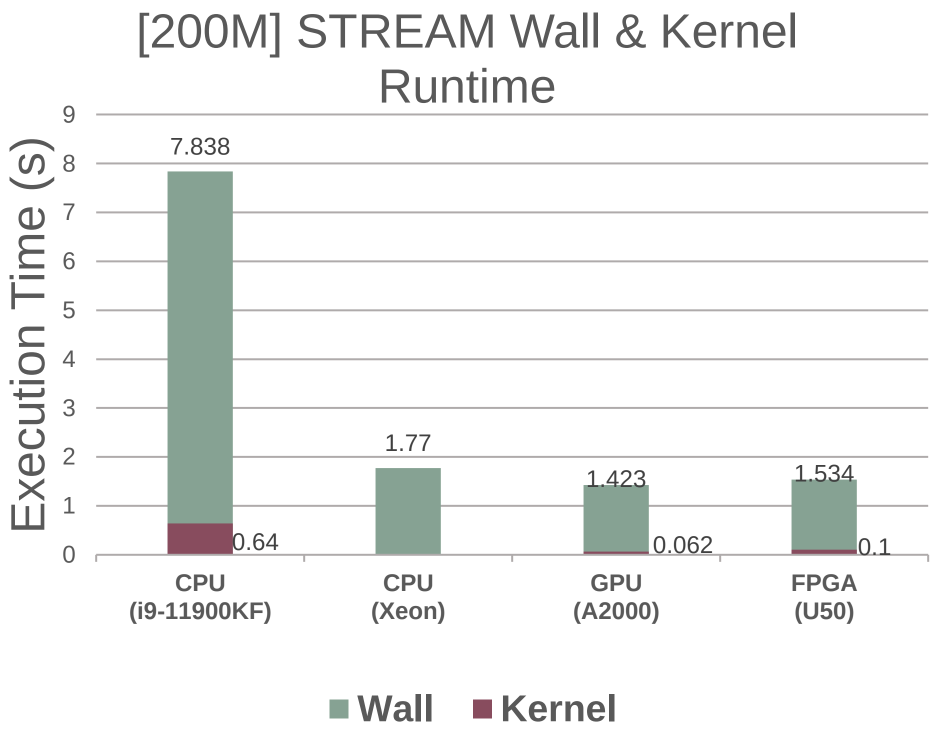




UK Research and Innovation

Results: STREAM





UK Research and Innovation

Conclusions

- Particular Algorithms are more energy efficient on one architecture over the another,
  - exploiting heterogeneity might be an answer to lower energy to solution.

- Significant time is spent optimising FPGA ports compared to both CPU and GPU.

- Greater In-depth architecture knowledge needed for FPGA's over CPU/GPU.

UK Research and Innovation

Hardware Clock Details

	SGEMM	FFT	STREAM
CPU: i9-11900KF	3.50 GHz	3.50 GHz	3.50 GHz
CPU: Xeon Gold 6240	2.60 GHz	2.60 GHz	2.60 GHz
GPU: Nvidia A2000	1200 Mhz	1200 Mhz	1200 Mhz
GPU: Nvidia A100	1095 Mhz	1095 Mhz	1095 Mhz
FPGA: Xilinx Alveo U50	Data: 300 Mhz	Data: 300 Mhz	Data: 300 Mhz
FPGA: Alveo U280	Kernel: 300 Mhz	Kernel: 300 Mhz	Kernel: 300 Mhz

UK Research and Innovation

University of BRISTOL

Carbon QuanDRI

Daniel Schien, University of Bristol
Noa Zilberman, University of Oxford
David Greenwood, Newcastle University
Alastair Dewhurst, STFC

UK Research and Innovation

https://net-zero-dri.ceda.ac.uk/cquandri/

/ˈkwɒnd(ə)ri/ - a state of perplexity or uncertainty over what to do in a difficult situation.

NetZero DRI *Services*

- Goal: Carbon-intelligent provision from a service-based perspective
- Why: New management capabilities for NetZero goals
  - Exploit dynamic marginal variability of grid carbon intensity (spatial, temporal and volume)
  - Enable efficiency within providers and consumers through transparency on a service level
- Challenge: Metrics and Methods for service-based assessments are currently missing
  - What flow do we need to measure (physical flow data, I.e. data, electricity) service flow (jobs, API requests, sessions) and value (financial)
  - And how does it translate to carbon over *various time scales* and LCA life cycle phases (data from procurement, operation, decommission)

UK Research and Innovation

Evidence for UKRI Net-Zero Strategy

- Case Study - HTC Compute Service
  - Metrics: Compute, Net and Disk I/O, Archival Storage Volume
- Model of Site Electricity Footprint based on HTC Metrics
  - Cooling
  - Storage
  - Network
  - Compute
- Carbon Footprint
- Marginal Carbon Intensity Model to Evaluate Carbon Reductions from Demand Response Mechanisms

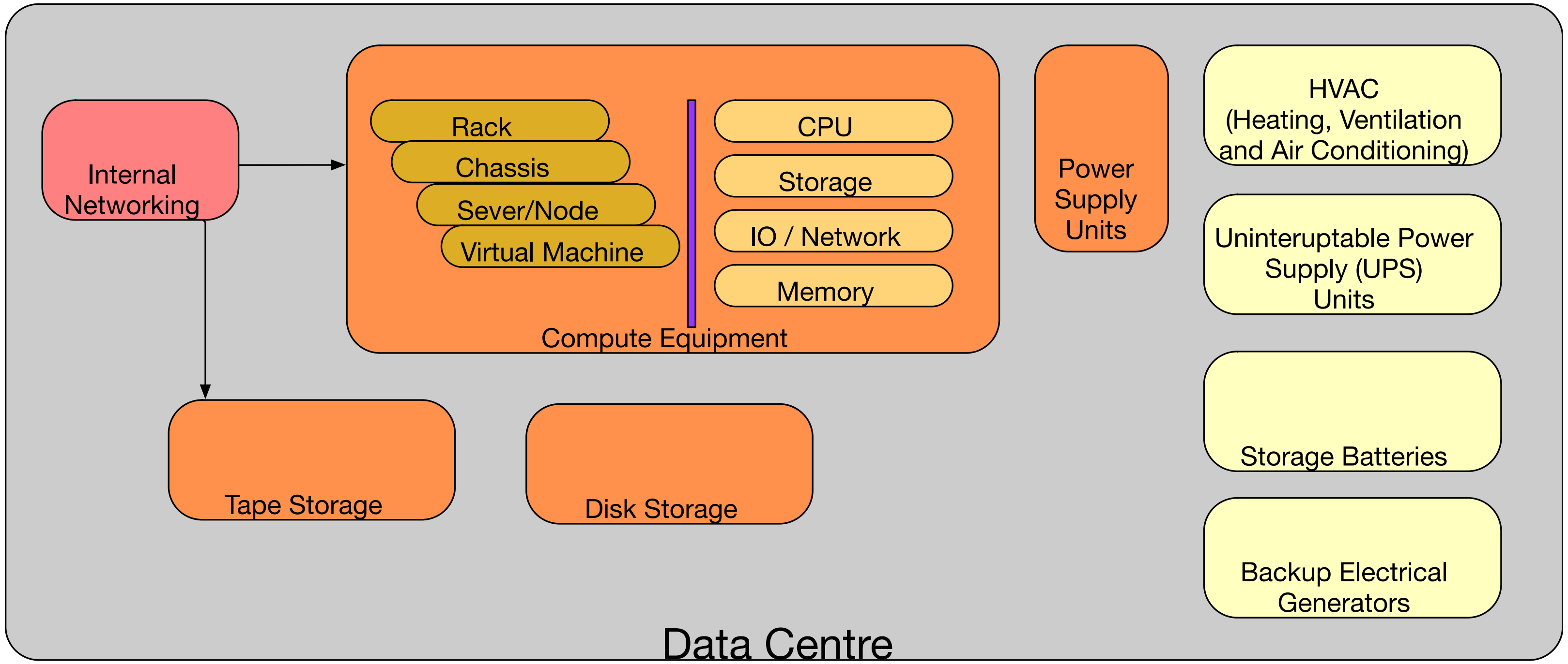
Approach

- Measurement trace that captures the sustainability aspects in the operation of a DRI service
- Location-based Marginal Carbon intensity model
- Combine to Carbon model of compute to run scenarios

UK Research and Innovation

Measurement Trace

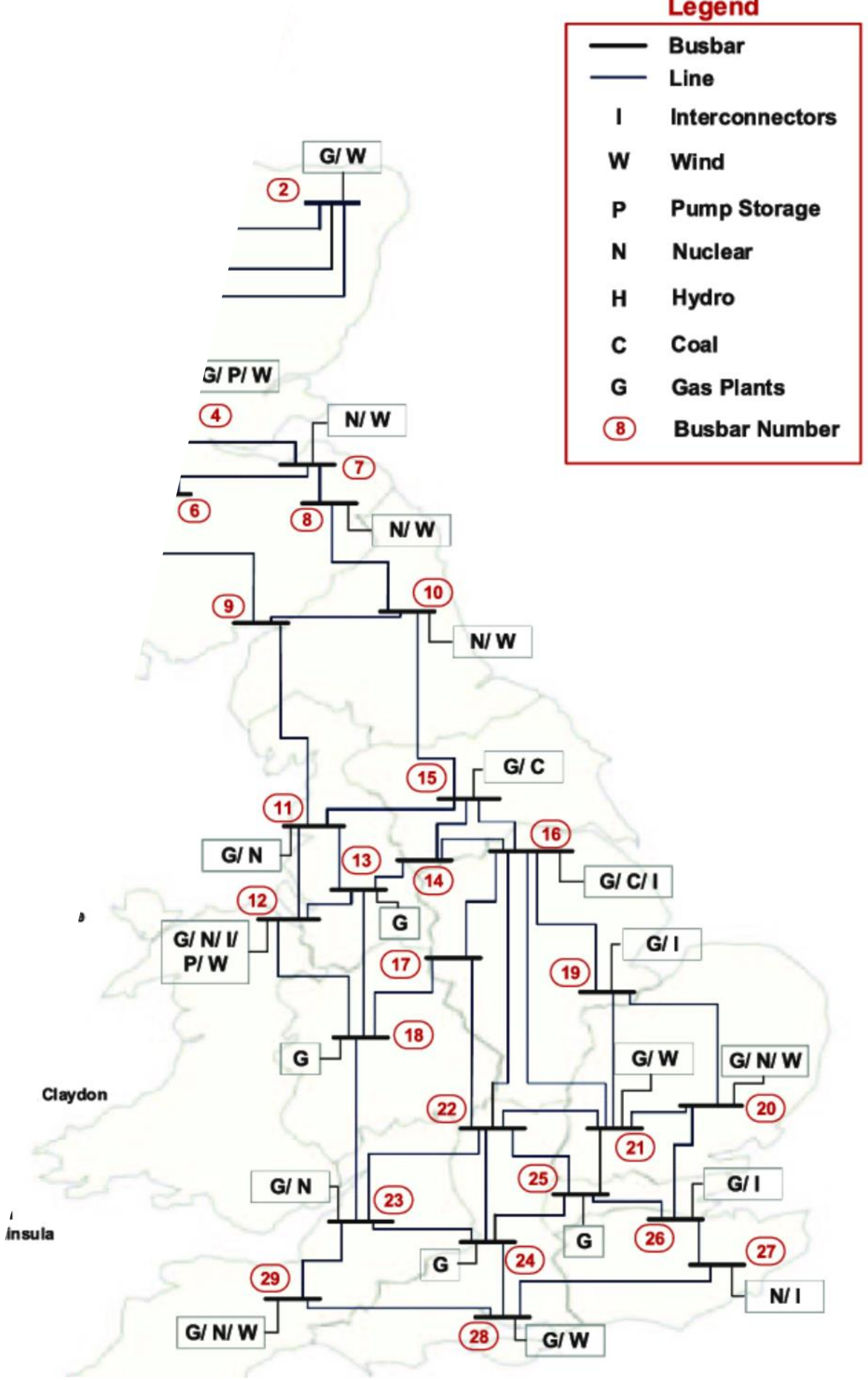
- Collecting, Site, Rack and Node power consumption
- Correlate with service data (jobs)



UK Research and Innovation

Energy Network Carbon Intensity

- 29 bus representation of the GB transmission system
- Optimal power flow model to estimate generation mix at each busbar
- Calculate
    - System Average Carbon Intensity
    - Average Nodal Carbon Intensity
    - Locational Marginal Carbon Intensity



UK Research and Innovation

Carbon Model

- Enable carbon budgets for users
- Trade-off between embodied and use phase carbon
- Evaluate benefit from creating flexible capacity
  - Batteries, overcapacity

Thank You

UK Research and Innovation

Panel Discussion

Chair: Ag Stephens (STFC CEDA)

Panel: Martin Juckes (STFC CEDA)

Wim Vanderbauwhede (Uni of Glasgow)

Justin O'Byrne (UKRI)

UK Research and Innovation