Sustainable Computing at Scale:
measuring and minimising emissions from software systems

Dr Eoin Woods
Artechra
July 2025


Climate change is one of the most urgent and serious problems facing humanity
Image by Markus Kammermann from Pixabay

WHAT DOES SOFTWARE HAVE TO DO WITH CLIMATE CHANGE?

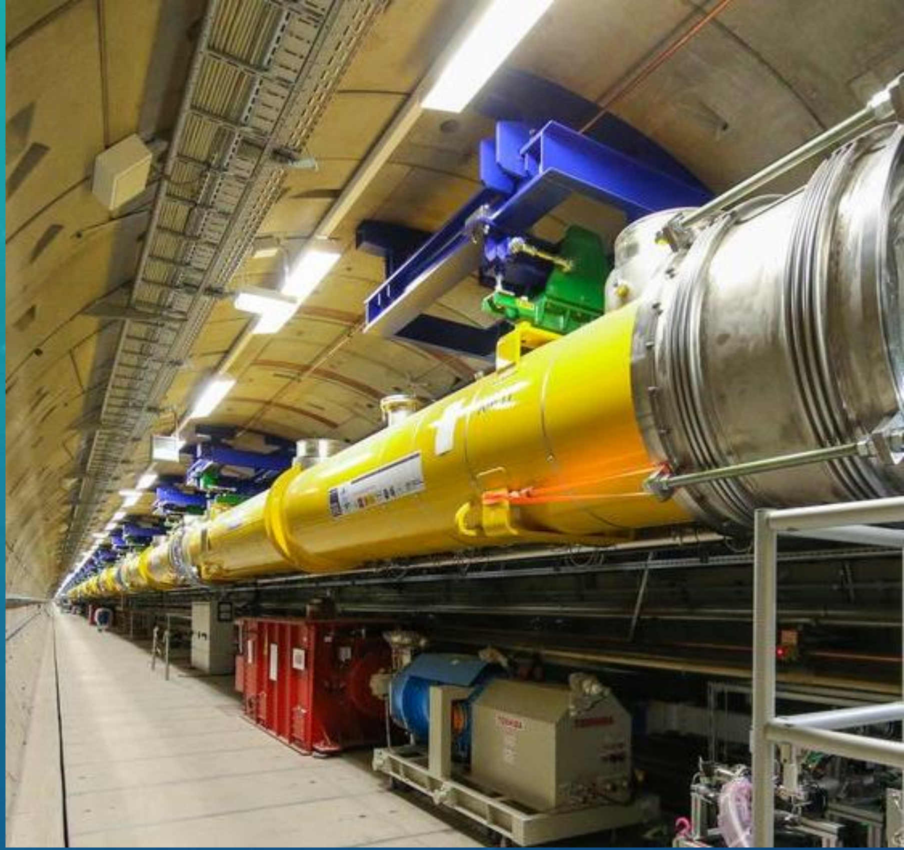
- Anything to do with measuring climate change impact is complicated, but …
- Estimates for the impact of ICT are 2-4% of global emissions
  - By comparison aviation is about 3% !
- And it is growing, at least in part due to AI
  - Data centres may use 5% of electricity by 2030
  - Microsoft's GHG emissions up 30% last year[1]

Image by István from Pixabay

1: The Verge: https://tinyurl.com/ms-emissions

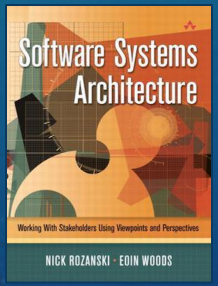
SCIENTIFIC COMPUTING CONTRIBUTES

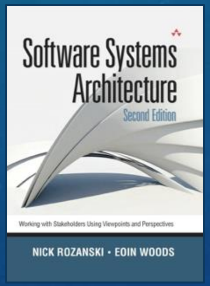
- CERN’s ATLAS experiment
  - 600,000 compute cores
  - 1,000 PB storage
  - 10 PB new data per year
  - Expecting ~2m compute cores in 2030
- Edinburgh’s ARCHER2 environment
  - 750,000 compute cores
  - 15 PB of storage

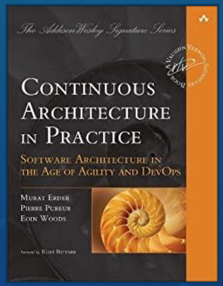


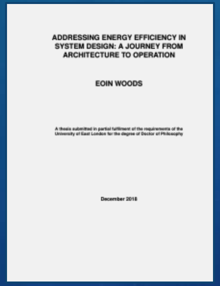
Eoin Woods

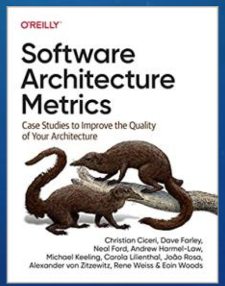
- Independent consultant
- Academic visitor in Dept of Computing, Imperial College London
- Ex-Chief Engineer at Endava based in London (2015-2025)
- 10+ years in products - Bull, Sybase, InterTrust
- 10 years in capital markets - UBS and BGI
- PhD in Software Architecture & Energy Efficiency[1] (2019)

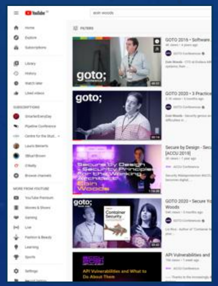












[1] https://repository.uel.ac.uk/item/8459v

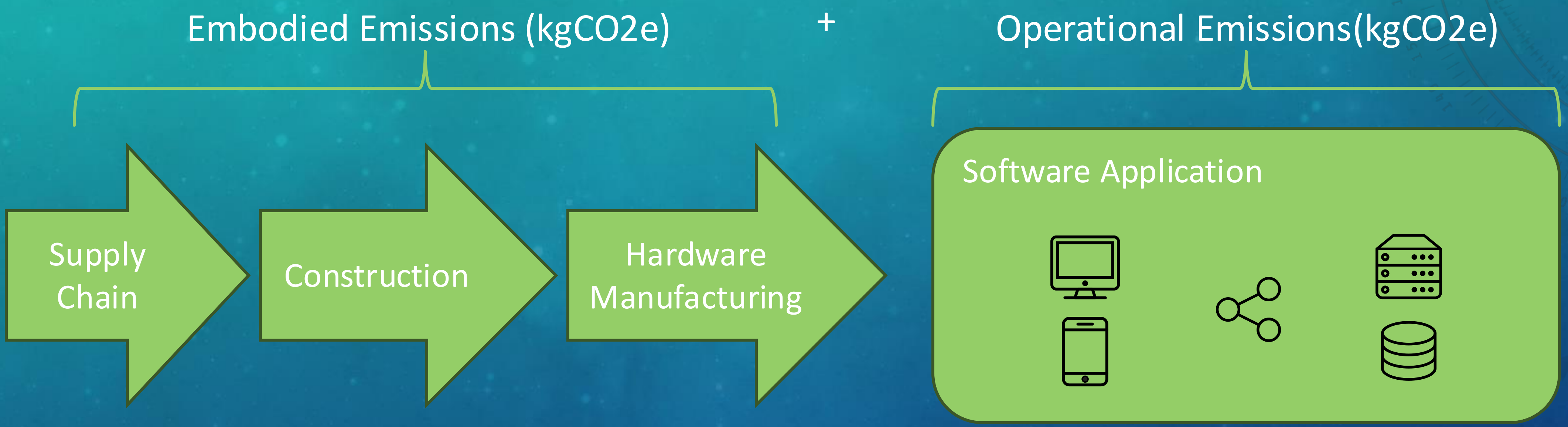
Agenda

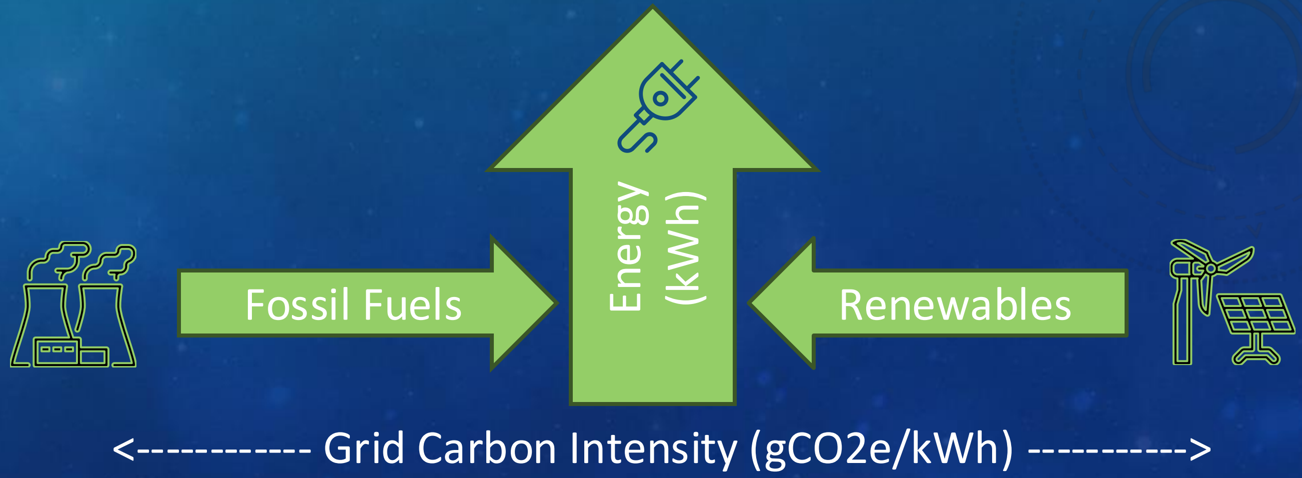
- Green Software Fundamentals
- Some Principles and Tactics
- Examples from Practice
- Beginning Your Journey

GREEN SOFTWARE FUNDAMENTALS

(Credit to James Costerton for the graphical concept)

GREENHOUSE GASES ("CARBON") AND SOFTWARE





Embodied emissions are the emissions from creating the data centre and hardware

Operational emissions are the emissions from the energy required to run the system

OPERATIONAL EMISSIONS

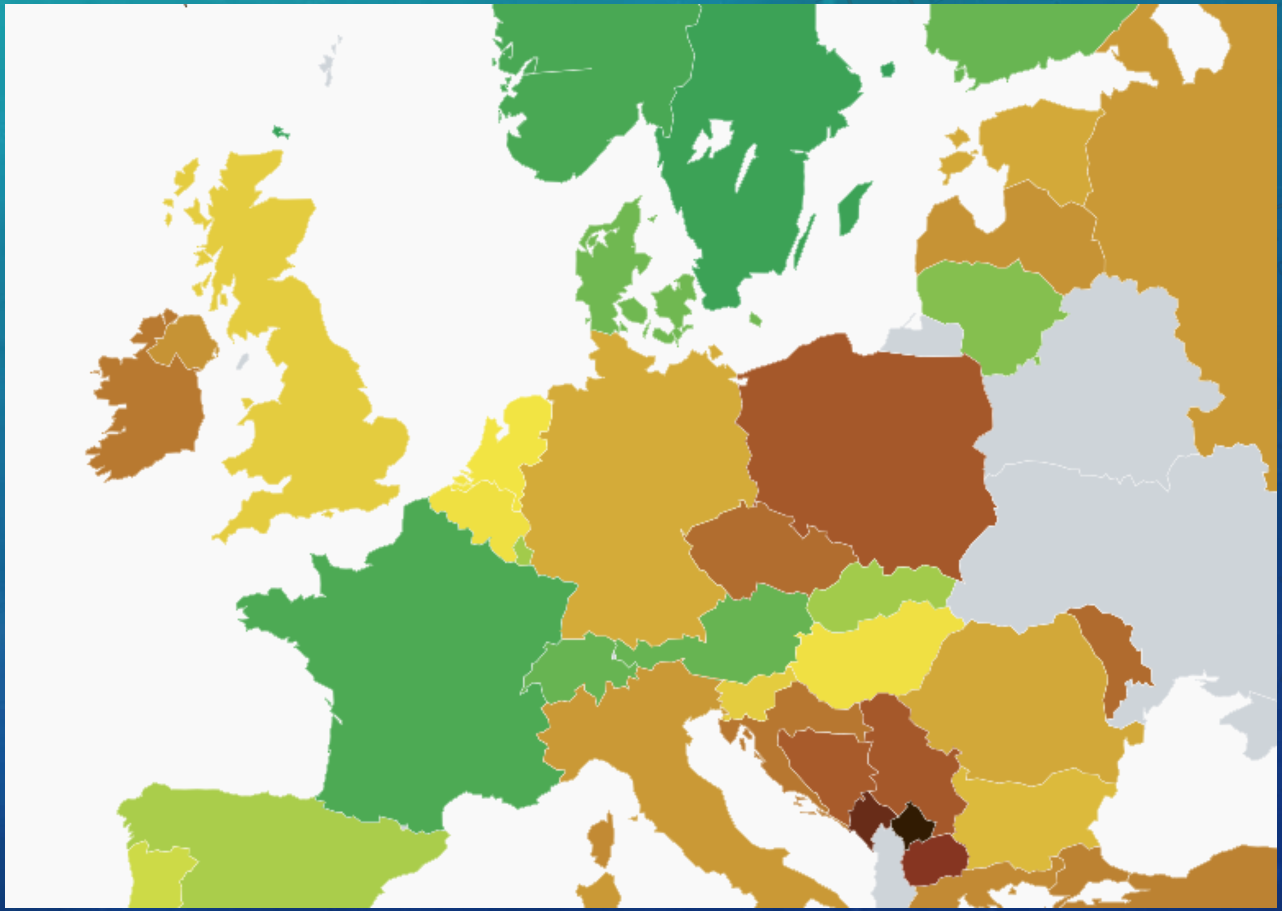
Emissions created during operation

GHG intensity of energy

X

Amount of energy used

- Demand
- Efficiency



Source:
electricitymaps.com
6 September 2024

SOME PRINCIPLES & TACTICS

GSF PRINCIPLES OF GREEN SOFTWARE

1. Emit the least amount of carbon possible.
2. Use the least amount of energy possible.
3. Do more when the electricity is cleaner and do less when the electricity is dirtier.
4. Use the least amount of embodied carbon possible.
5. What you can't measure, you can't improve.
6. Understand the exact mechanism of carbon reduction.

https://learn.greensoftware.foundation

COMPUTE INTENSIVE TACTICS

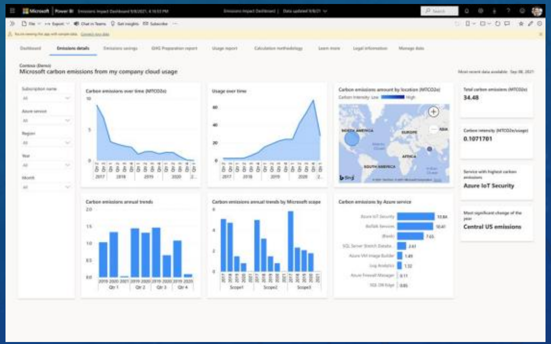
1. Emissions as a Quality Attribute
2. Measurement Culture
3. Unified Policies and Practices
4. Demand Shifting & Shaping
5. Actively Avoiding Waste
6. Account for Lifecycle Emissions

T1: EMISSIONS AS A QUALITY ATTRIBUTE

- We measure performance to focus attention and effort
- Targeting emissions in the same way is a first step to awareness and reduction
- Baseline estimate, then set emissions targets as a software requirement
- Collect averages per site for reference

T2: MEASUREMENT CULTURE

- Difficult and tends to be via estimation
  - But data helps to motivate and focus action
- Create & use reusable models
  - Share expertise and effort across the community
- Add data collection to standard compute programming frameworks
  - Collect resource data & calculate energy usage
  - Estimate energy emissions from GHG intensity
  - Allocate share of embodied emissions (…)



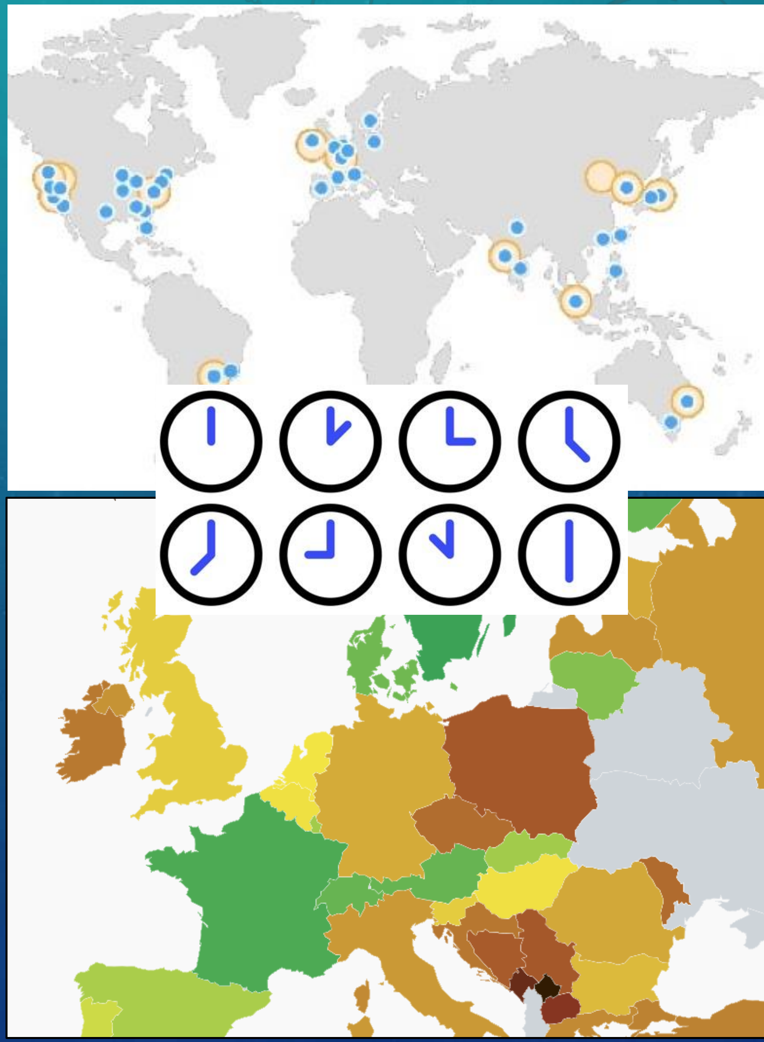
T3: UNIFIED POLICIES & PRACTICES

- Estimating and reducing emissions is much easier with standard environments & policies
  - GHG estimation can be built into runtime
  - Hardware can be allocated efficiently
  - Users can be reminded about emissions
  - Policies can encourage sustainable practices
- Extend existing open source where possible
- Create a “paved road” (not just a set of rules)

T4: DEMAND SHIFTING & SHAPING

- Look for flexibility in when and where workload is executed
  - Run batches at different times
  - Execute workload in different locations (perhaps …)
- Can computational intensity of workload be varied?
  - accuracy, data size, precision

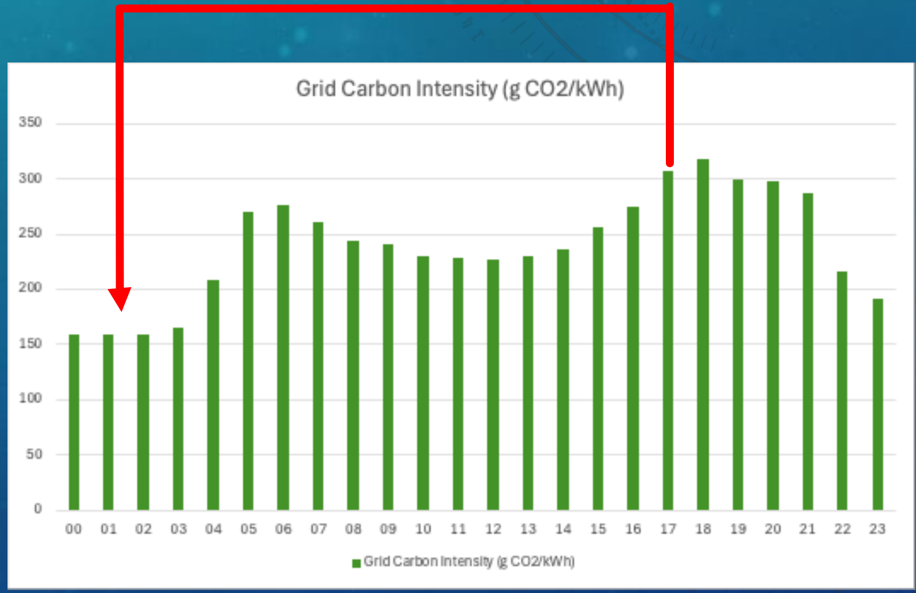
Electricity grids vary in their GHG intensity



https://www.freepik.com/free-vector/clocks-with-different-hand-times-set_263632366.htm

T4: DEMAND SHIFTING & SHAPING

- Move workload to times or places with lower grid carbon intensity
  - Move workload to when local grid is ”greener”
  - Move workload location which is “greener”
  - Simplify workload when high GHG intensity



- Trade offs:
  - complexity
  - hidden emissions (e.g. data movement)
  - simplified workload results may not be useful

T5: ACTIVELY AVOIDING WASTE

- Runtime efficiency => reduced emissions
  - Ensuring high compute utilisation
  - Selecting the right compute env for a workload
  - Avoiding wasted computation
  - Minimise data size and storage duration
- At scale easy to overlook but small % matter
- Automation can help to highlight problems and provide prompt for improvement

```
!!
!! @note Should only be used via class \
!-----------------------------------------------
SUBROUTINE gs_setup(self)
CLASS(psi_gs_eq), INTENT(inout) :: self
INTEGER(4) :: i,io_unit
REAL(8) :: pmin
STACK_PUSH
!---Load GS grid
CALL trimesh_load(self%mesh,TRIM(self%gr
CALL trimesh_local_setup(self%mesh)
!---Load GS field (order)
OPEN(NEWUNIT=io_unit,FILE=TRIM(self%fiel
READ(io_unit,*)self%order
ALLOCATE(self%lagrange)
CALL psi_lag_setup_trimesh(self%lagrange
!---Load GS field (B,P)
ALLOCATE(self%Bvals(3,self%lagrange%ne),
DO i=1,self%lagrange%ne
  READ(io_unit,*)self%Bvals(:,i),self%Pv
END DO
CLOSE(io_unit)
!---
pmin=MINVAL(self%Pvals)
self%pmax=MAXVAL(self%Pvals)
!
self%P_interp%vals=>self%Pvals
```

T6: ACCOUNT FOR LIFECYCLE EMISSIONS

- Operational emissions typically 80% of lifecycle emissions for servers …

  … but embodied emissions also significant

- Difficult for users to estimate
  - Estimate on a site level and allocate to workload (hours spent active in compute and size*duration of storage)
- At DC scale consider embodied emissions vs energy usage when considering upgrades

Bashroush, "*A comprehensive reasoning framework for hardware refresh in data centers.*" *IEEE ToSC, 2018*

OTHER TACTICS TO CONSIDER

Use of public cloud HPC services

Code static analysis tools

Teach automated software testing

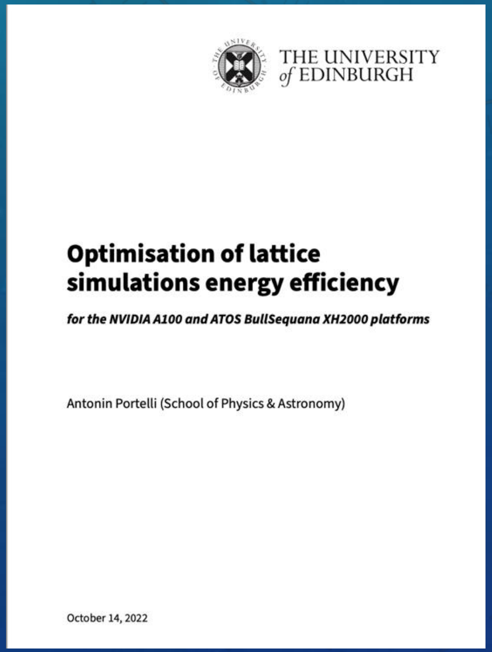
Developer "certification"

Awards or recognition for projects achieving sustainability standards

EXAMPLES FROM PRACTICE

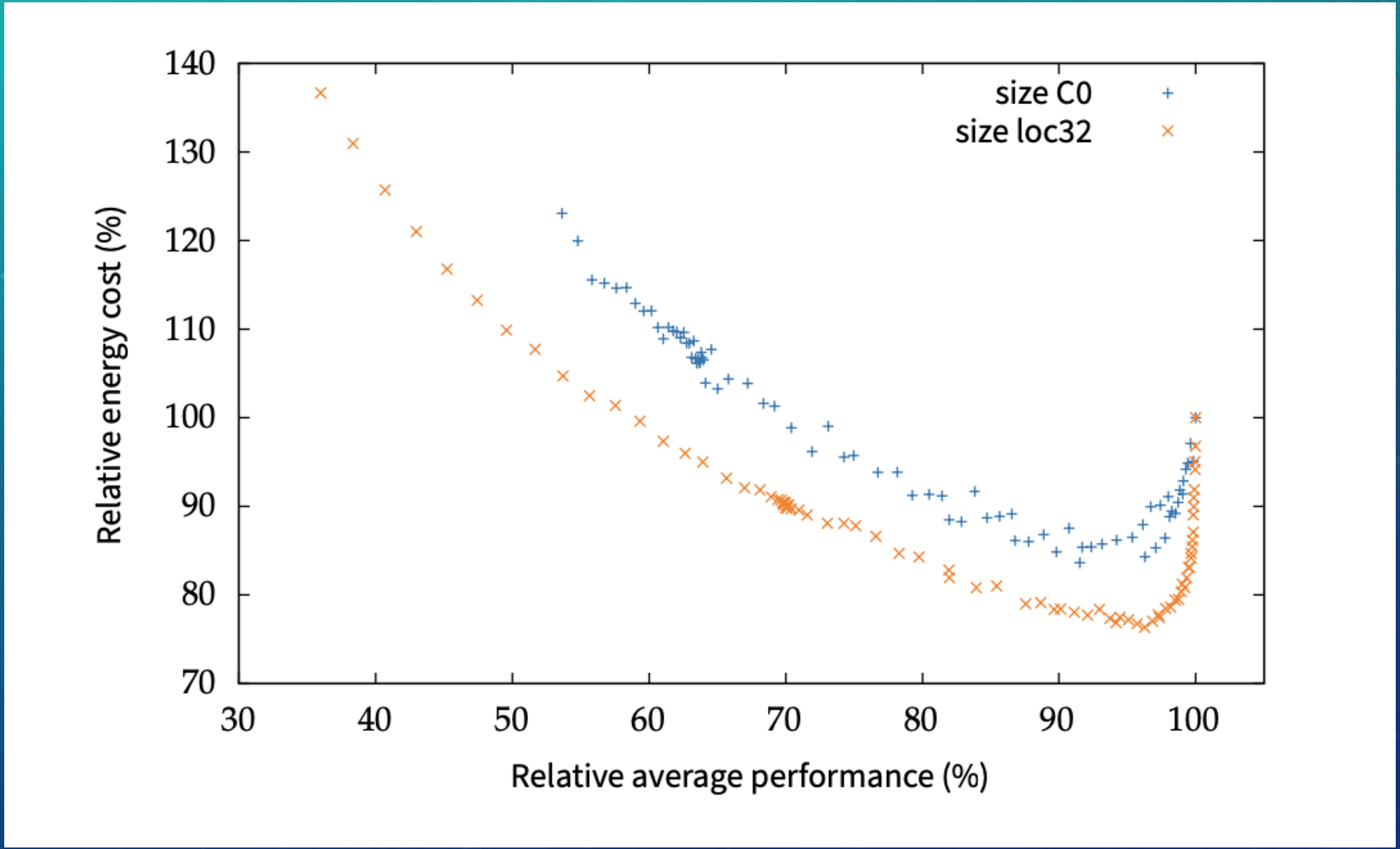
CASE STUDY – STFC DiRAC DOWNCLOCKING

- STFC DiRAC "Tursa" Supercomputer in Edinburgh, UK
  - 448 A100-40 GPU, 224 AMD CPUs, 112 TB of memory
- Lattix QCD simulation
  - "DWF" benchmark from the "Grid" library
- Tested energy and performance impact of reducing GPU clock speed
- Reducing clock speed from 1.4GHz to 1.0GHz results in 10% performance reduction but 16-24% energy saving



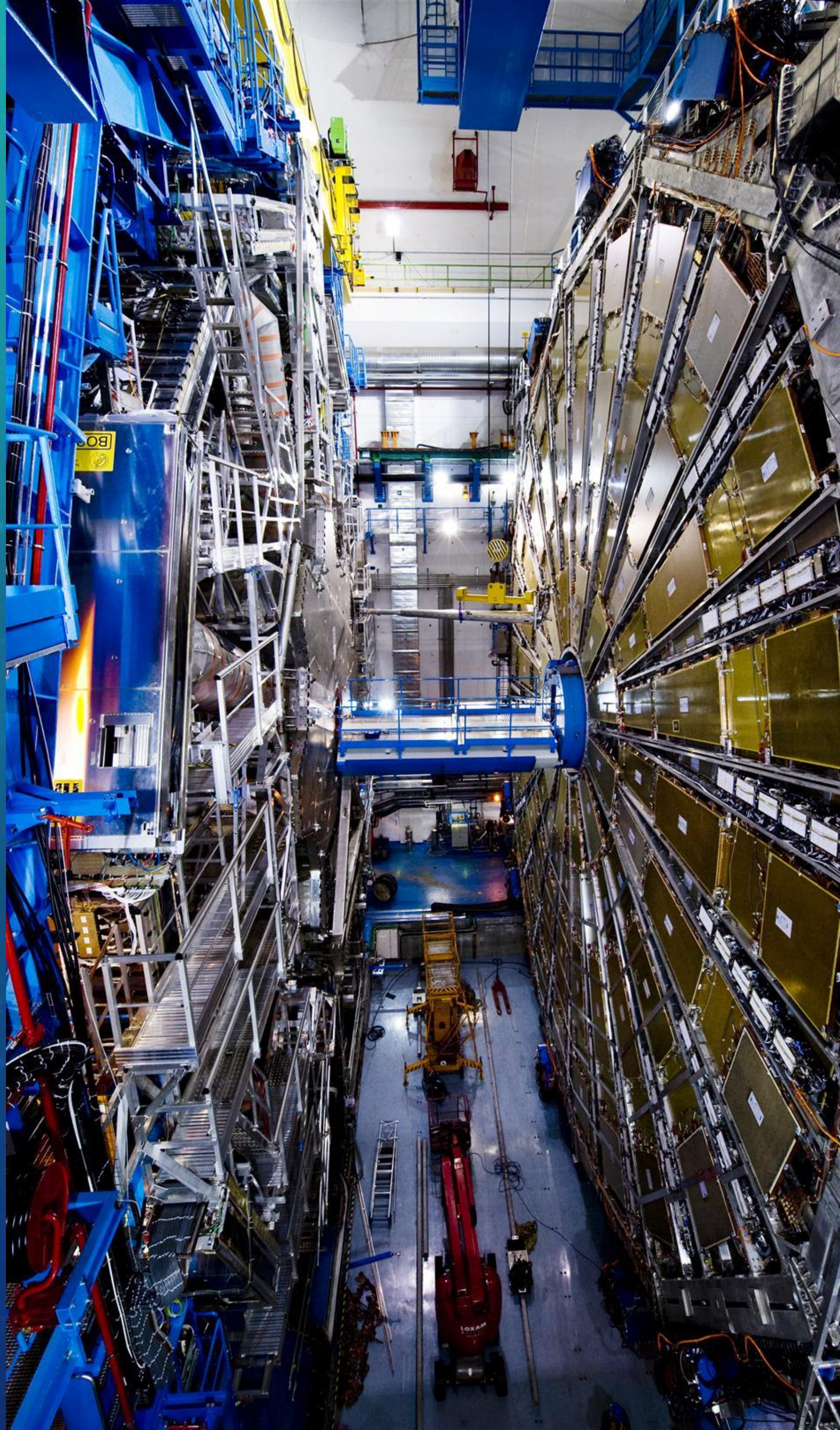
Antonin Portelli, *Optimisation of lattice simulations energy efficiency*, DiRAC TR, 2022, https://zenodo.org/records/7057319

CASE STUDY – STFC DiRAC DOWNCLOCKING



CASE STUDY – ATLAS

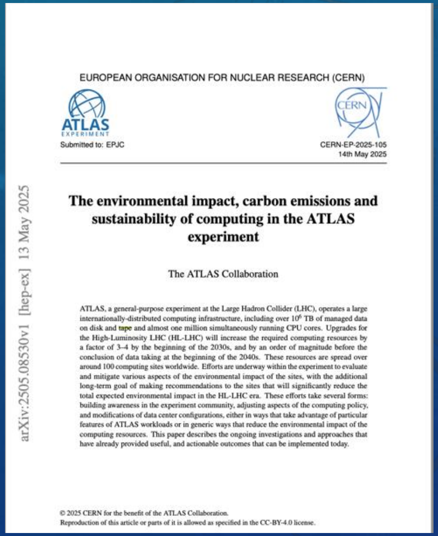
- Large experiment at CERN LHC
- Huge distributed infrastructure
  - 700k cores, 10^6 TB of NAS storage, 100 sites
- Expecting significant growth by 2030+
  - 1.5-3m cores, 3-4 x 10^6 TB storage
- Strategic desire to manage and minimise GHG emissions



CASE STUDY – ATLAS

4 elements of their sustainability initiative[1]:

- Create **awareness** of GHG emissions
  - Scientists, developers, administrators
- **Policies and standard procedures** to drive emissions reduction
- **ATLAS specific** administration practices
- **General data centre** GHG reduction practices



EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: EPJC

CERN-EP-2025-105
14th May 2025

arXiv:2505.08530v1 [hep-ex] 13 May 2025

The environmental impact, carbon emissions and sustainability of computing in the ATLAS experiment

The ATLAS Collaboration

ATLAS, a general-purpose experiment at the Large Hadron Collider (LHC), operates a large internationally-distributed computing infrastructure, including over 10^6 TB of managed data on disk and tape and almost one million simultaneously running CPU cores. Upgrades for the High-Luminosity LHC (HL-LHC) will increase the required computing resources by a factor of 3–4 by the beginning of the 2030s, and by an order of magnitude before the conclusion of data taking at the beginning of the 2040s. These resources are spread over around 100 computing sites worldwide. Efforts are underway within the experiment to evaluate and mitigate various aspects of the environmental impact of the sites, with the additional long-term goal of making recommendations to the sites that will significantly reduce the total expected environmental impact in the HL-LHC era. These efforts take several forms: building awareness in the experiment community, adjusting aspects of the computing policy, and modifications of data center configurations, either in ways that take advantage of particular features of ATLAS workloads or in generic ways that reduce the environmental impact of the computing resources. This paper describes the ongoing investigations and approaches that have already provided useful, and actionable outcomes that can be implemented today.

1. https://arxiv.org/abs/2505.08530 (2025)

CASE STUDY – ATLAS

Examples of actions from the initiative:

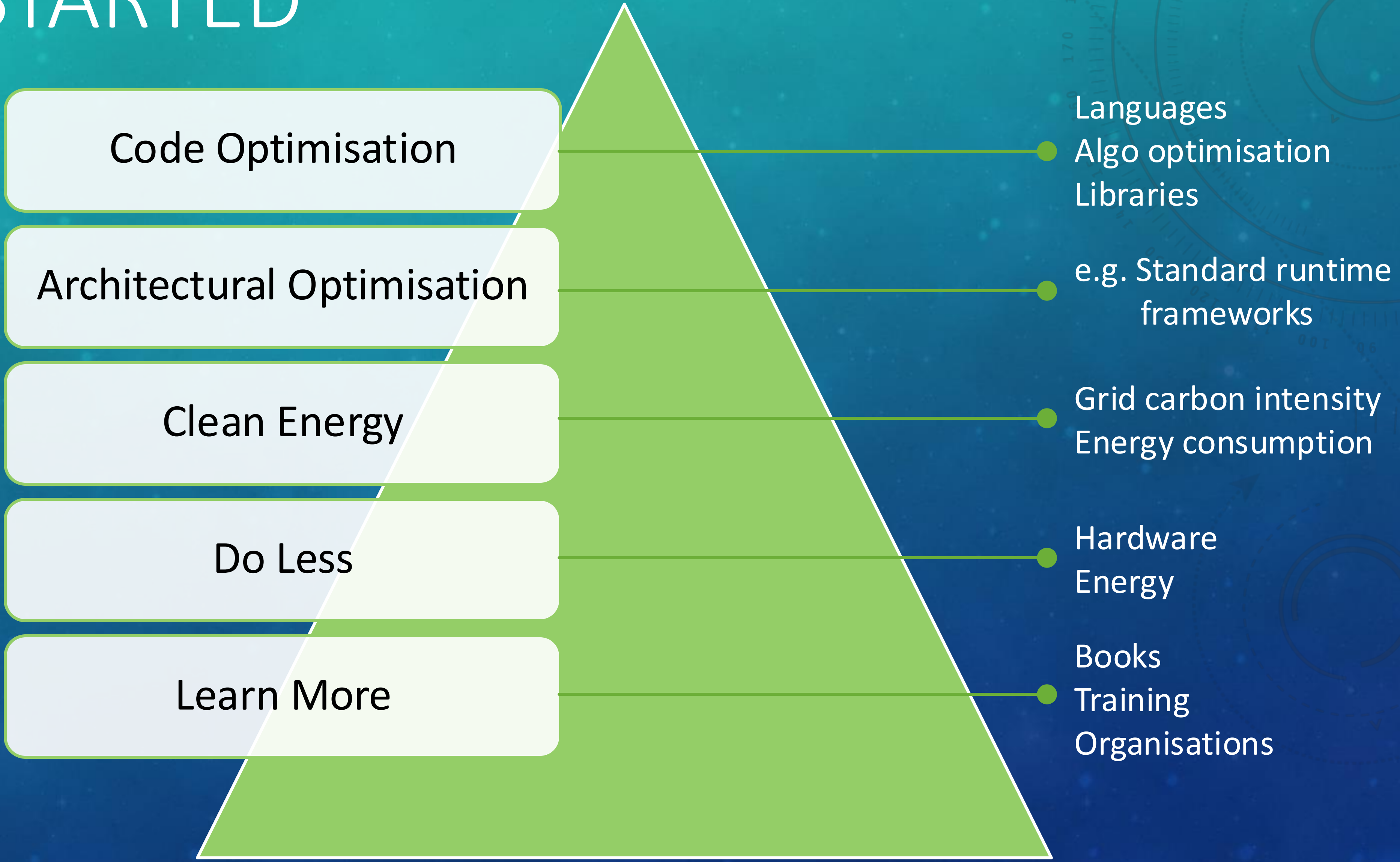
- Include GHG impact of computing in **end-user training**
- Provide GHG **emissions estimates** in job output report
- Encourage the use of **tape vs disk** - lower overall emissions[1]
- Research on when and how to use **compression**[2]
- Automated testing of incoming tasks before releasing entire workload – **avoid waste in case of errors**

1. https://arxiv.org/abs/2404.06335 (2024)

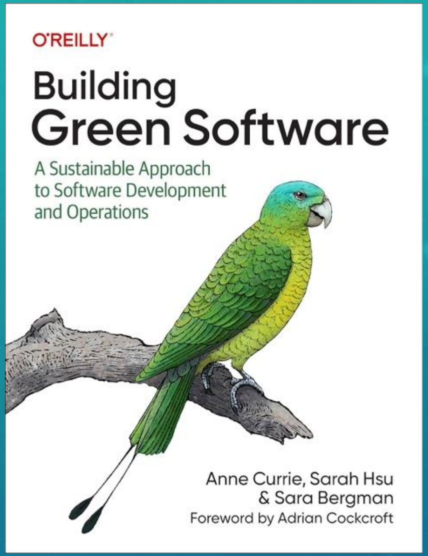
2. https://doi.org/10.1051/epjconf/202429503027 (2024)

BEGINNING YOUR JOURNEY

GETTING STARTED



RESOURCES



Green Software for Practitioners
LFC131
THE LINUX FOUNDATION | Education
OPEN SOURCE BEST PRACTICE

https://tinyurl.com/lsf-gsw-course

Green Algorithms
Towards environmentally sustainable computational science

www.green-algorithms.org

GreenDiSC

https://www.software.ac.uk/GreenDiSC

GreenDIGIT

https://greendigit-project.eu/

Green Software Foundation

https://greensoftware.foundation

ClimateAction.tech

https://climateaction.tech

SDIA

https://sdialliance.org

RESOURCES

ELECTRICITY MAPS

WattTime

climatiq

Software Carbon Intensity Specification (GSF)

Tech Carbon Standard (Scott Logic)

Thank you … questions?

Eoin Woods
www.eoinwoods.info
eoin.woods@artechra.com
@eoinwoods.bsky.social
threads.net/@eoinwoodz