ASSESSING AND REDUCING THE ENVIRONMENTAL IMPACT OF HPC

Adrian Jackson

a.jackson@epcc.ed.ac.uk

With a lot of help from my friends

THE UNIVERSITY of EDINBURGH

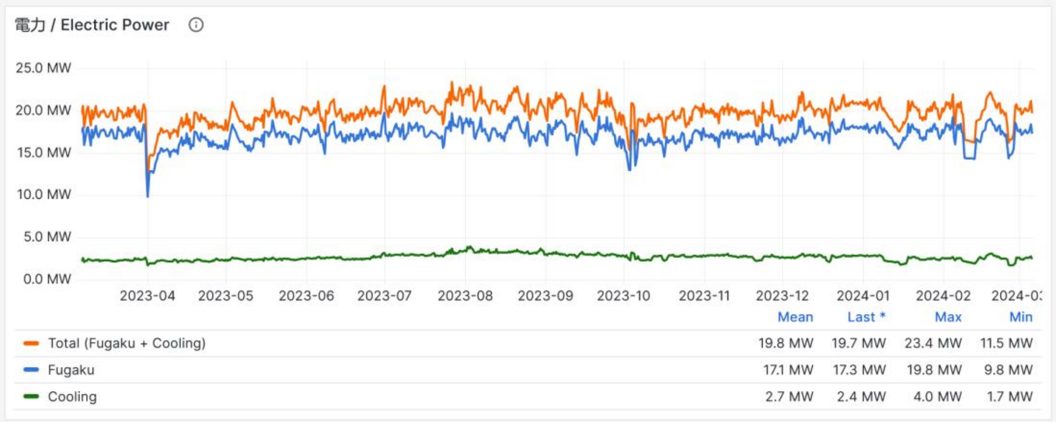
epcc

Landscape

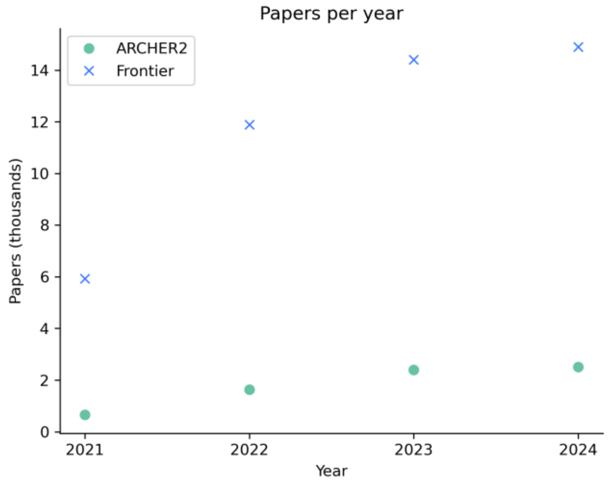
- Research, innovation and partnership continue to underpin the UK's commitment to:
  - achieving a net zero economy by 2050
  - respond to the challenges of climate change
  - live more sustainably.
- UKRI is committed to becoming net zero by 2040
- Digital Research Infrastructure (DRI) part of UKRI spending
  - Implies Net Zero DRI
  - Net means some parts can have climate impacts but that would mean other parts removing climate impacts
- Net zero:
  - Zero overall carbon emissions associated with X
  - Zero overall climate/environmental impact associated with X (general sustainability)

Background

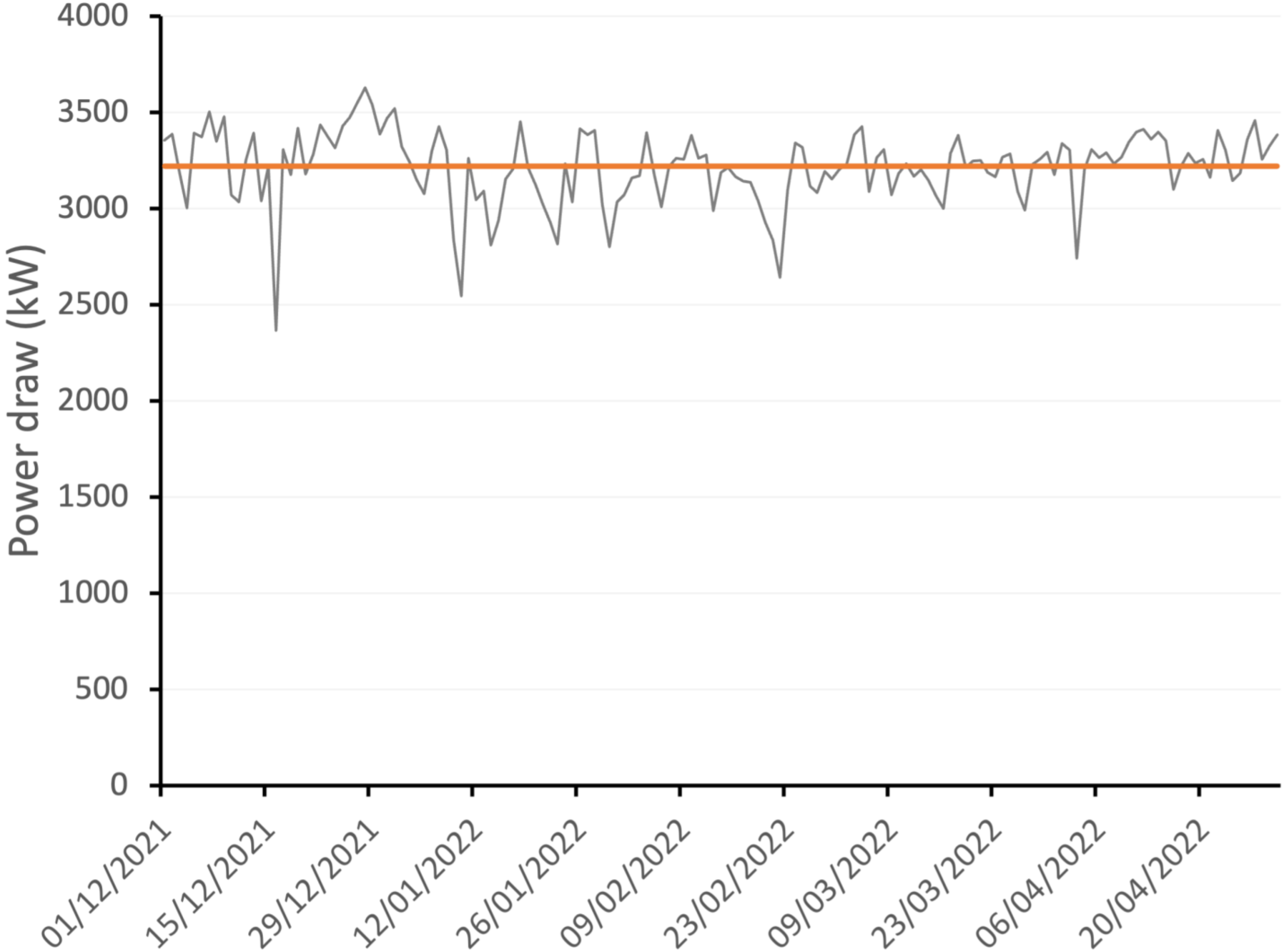
- Large scale HPC using significant amounts of electricity
  - ARCHER2 ~3-4MW (~2500 papers)
  - Frontier ~23MW (+ cooling) (~15000 papers)
  - Aurora ~25MW (+ cooling)



- Two forms of climate impact to consider
  - Active/direct (scope 2)
  - Embodied/embedded (scope 3)
- Some things can be made directly (nearly) Net Zero
  - Electricity generation
- Other things not so much
  - Mining, fuel production, etc...



ARCHER2 context: historical power draw measurements



- Mean power draw from compute node cabinets: 3220 kW
- Measurements taken from the chassis management infrastructure

Power draw by component

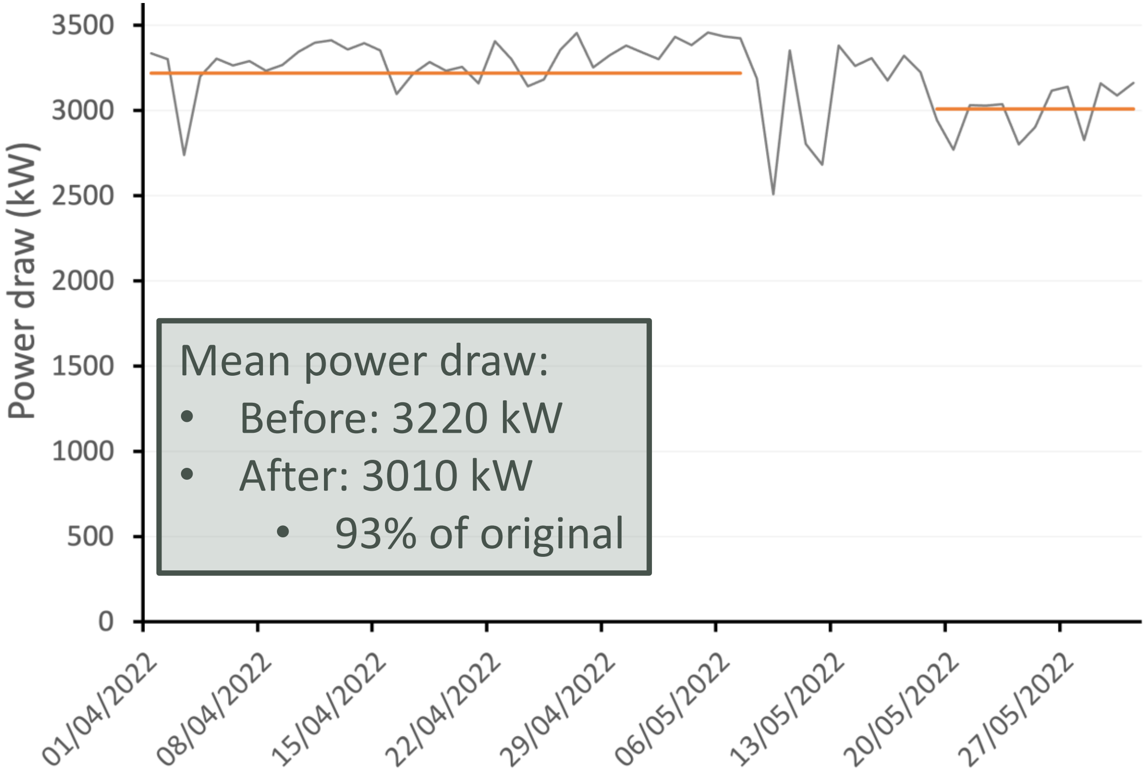
- Estimated loaded power draws for ARCHER2 components:
  - Some values measured by experiments and others provided by HPE engineers

Component	Notes	Idle (each)	Loaded (each)	Approx. %
Compute nodes	5860 nodes	1350 kW (0.23 kW)	3000 kW (0.51 kW)	80%
Slingshot interconnect	768 switches	100-200 kW (0.10-0.25 kW)	540 kW (0.70 kW)	10%
Other Cabinet Overheads	23 cabinets	100-200 kW (4.3-8.7 kW)	210 kW (9.1 kW)	6%
Coolant Distribution Units	6 CDUs	96 kW (16 kW)	96 kW (16 kW)	3%
File systems	5 file systems	40 kW (8 kW)	40 kW (8 kW)	1%
Service nodes	Negligible	-	-	
Total		1800 kW	3900 kW	

- Energy use dominated by compute cabinets; storage power not important
- Idle power draw of compute nodes is high

BIOS Setting

- Processor and system can be configured for different runtime modes
- Original ARCHER2 configuration called *Power Determinism*
- *Performance Determinism* keeps processor performance more consistent
    - Performance of multi-node parallel applications is determined by slowest node
    - Any extra power draw for performance above the slowest node is wasted power



Impact on application performance

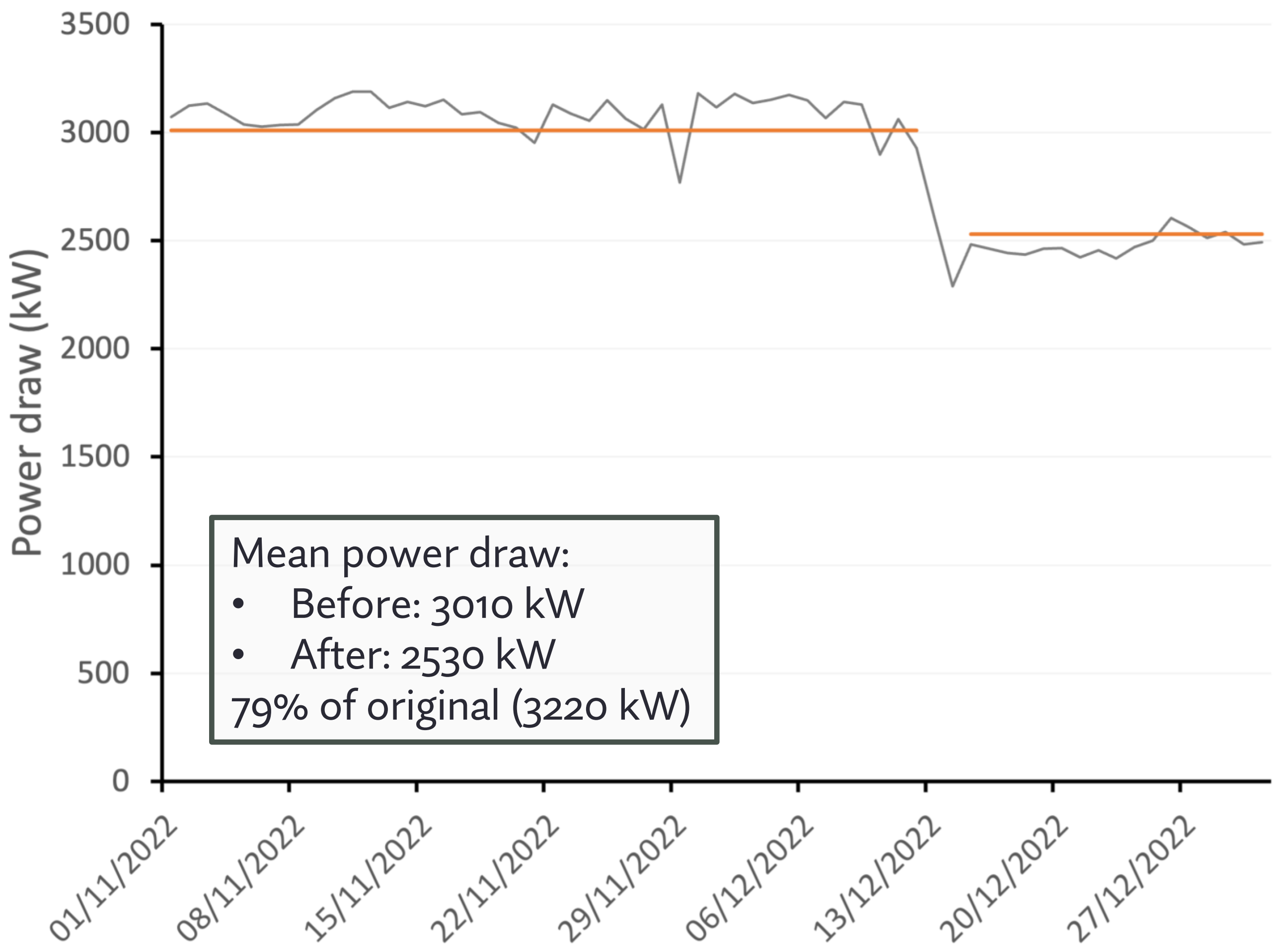
Application benchmark	Number of nodes	Performance ratio PerfMode:PowerMode	Energy[1] ratio PerfMode:PowerMode
CASTEP Al Slab	16	0.99	0.94
OpenSBLI TGV 1024^3	32	1.00	0.90
VASP TiO_2	32	0.99	0.93

[1]Energy measured from on-node energy use counters – only reflects node energy use

- Performance impact is generally low – expected to be lower where more nodes are used
- Energy savings measured using cabinet power in line with energy savings measured on compute nodes
  - Suggests that overheads on top of compute node power do not affect conclusions

CPU Frequency – impact on power draw

- As well as configuring the processor/node overall, can modify processor behaviour on the fly
- ARCHER2 configuration:
  - 2.25GHz processor frequency
  - turbo boost enabled
- New configuration
  - 2.00 GHz (no turbo boost)
- Can be done on a per application/per job basis



CPU Frequency – impact on performance

Application benchmark	Performance ratio	Energy usage ratio	Energy to performance ratio
VASP CdTe	0.95	0.88	1.08
GROMACS 1400k atoms	0.83	0.92	0.9
CP2K H2O 2048	0.91	0.93	0.98
LAMMPS Ethanol	0.74	0.92	0.8
CASTEP Al Slab	0.93	0.88	1.05
ONETEP hBN-BP-hBN	0.92	0.82	1.12
Nektar++ TGV 128 DoF	0.80	0.80	1

- All applications are more energy efficient at 2.0 GHz
- Looking at cost-efficiency would suggest:
  - Frequency set to 2.25 GHz: GROMACS and LAMMPS, Nektar++ [due to increased residency costs]
  - Frequency set to 2.0 GHz: VASP, CASTEP, ONETEP, CP2K
- Default frequency: 2.0 GHz with strong advice to users to test impact on their software

CPU Frequency – impact on performance

- Reserved a full cabinet (256 nodes) and filled with copies of benchmarks
- Initially focused on applications which would be running at 2.0 GHz

Experiment	Cabinet energy use (kWh)[1]	Node energy use (kWh)[2]	Overheads (kWh)	% Overheads	Cabinet ratio to 2.25 GHz	Node ratio to 2.25 GHz
8-node VASP, 256 nodes, 2.25 GHz	43.9	35.3	8.6	19.6%		
8-node VASP, 256 nodes, 2.00 GHz	38.5	30.4	8.1	21.0%	0.88	0.86

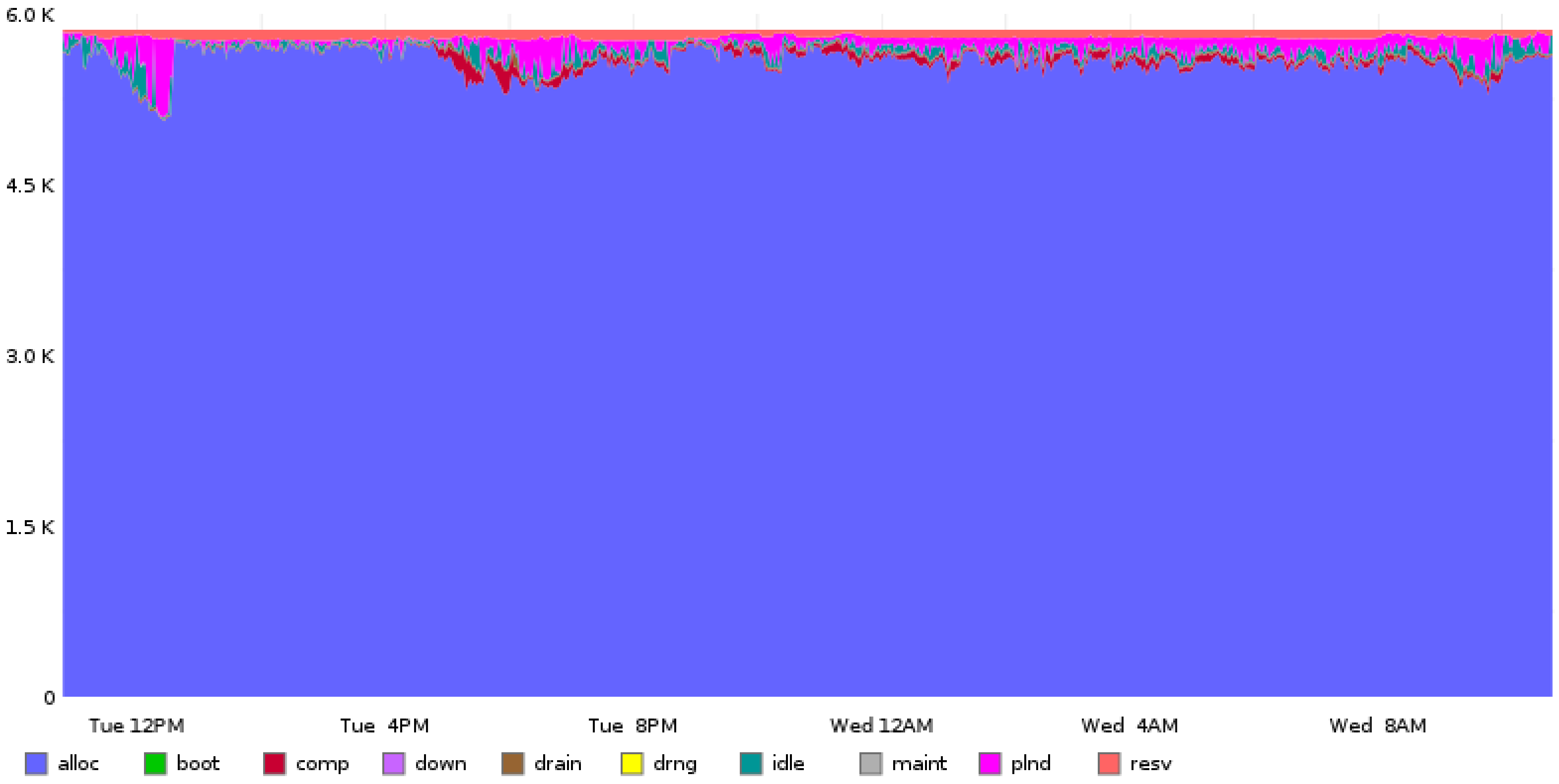
Experiment	Cabinet energy use (kWh)[1]	Node energy use (kWh)[2]	Overheads (kWh)	% Overheads	Cabinet ratio to 2.25 GHz	Node ratio to 2.25 GHz
4-node ONETEP, 256 nodes, 2.25 GHz	128.2	108.3	19.8	15.5%		
4-node ONETEP, 256 nodes, 2.00 GHz	107.8	88.5	19.3	17.9%	0.84	0.82

[1] Calculated from instantaneous cabinet power draw measurements during benchmark runtime
[2] Sum of energies from all calculations in set that filled 256 nodes

- Energy savings measured at the node level clearly propagate to full cabinet energy use
  - Cabinet energy use includes interconnect switches and power overheads

Usage and renewable energy

- Once you have installed a system, use it as much as possible
  - Make good use of the embodied energy
- Electricity impacts can be reduced
  - ARCHER2 electricity is renewable
  - Purchased with REGO certificates (100% renewable generation)
- Leaves embodied costs (scope 3)



HPC climate impacts

- Three (maybe four) components to consider for climate impact
  - Active direct: Electricity to run the systems
  - Active indirect: Electricity to run the infrastructure
  - Embodied direct: The impacts to manufacture and install the systems
  - Embodied indirect: The impacts to manufacture and install the infrastructure/facilities

(Active direct and Active indirect: PUE)

- Net Zero implies taking all of these to zero
  - Active could be "easy"
  - Embodied then comes to dominate

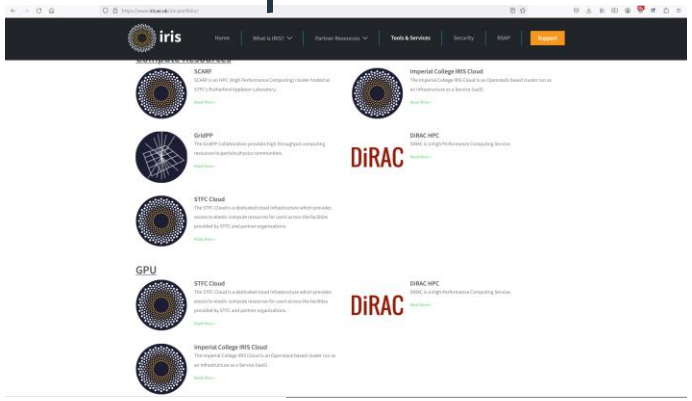
Power **U**sage **E**ffectiveness

$$\frac{\text{Total Power Used}}{\text{Power Used for Compute}}$$

PUE: 1 – Efficiency of 100%
PUE: 2 – Efficiency of 50%

Embodied Climate Impacts

- Aim to estimate embodied carbon for a computing infrastructure
  - Net Zero Scoping Projects: IRISCAST
- Funded by the UKRI Net Zero Digital Research Infrastructures Scoping Project
  - https://net-zero-dri.ceda.ac.uk/
- Capture/calculate the climate cost of IRIS during a set of snapshot periods
  - 24-hour run of the full infrastructure (subset of IRIS)
  - Capture as much details as possible
- Estimate/ballpark the "impact" of the whole system



Carbon model

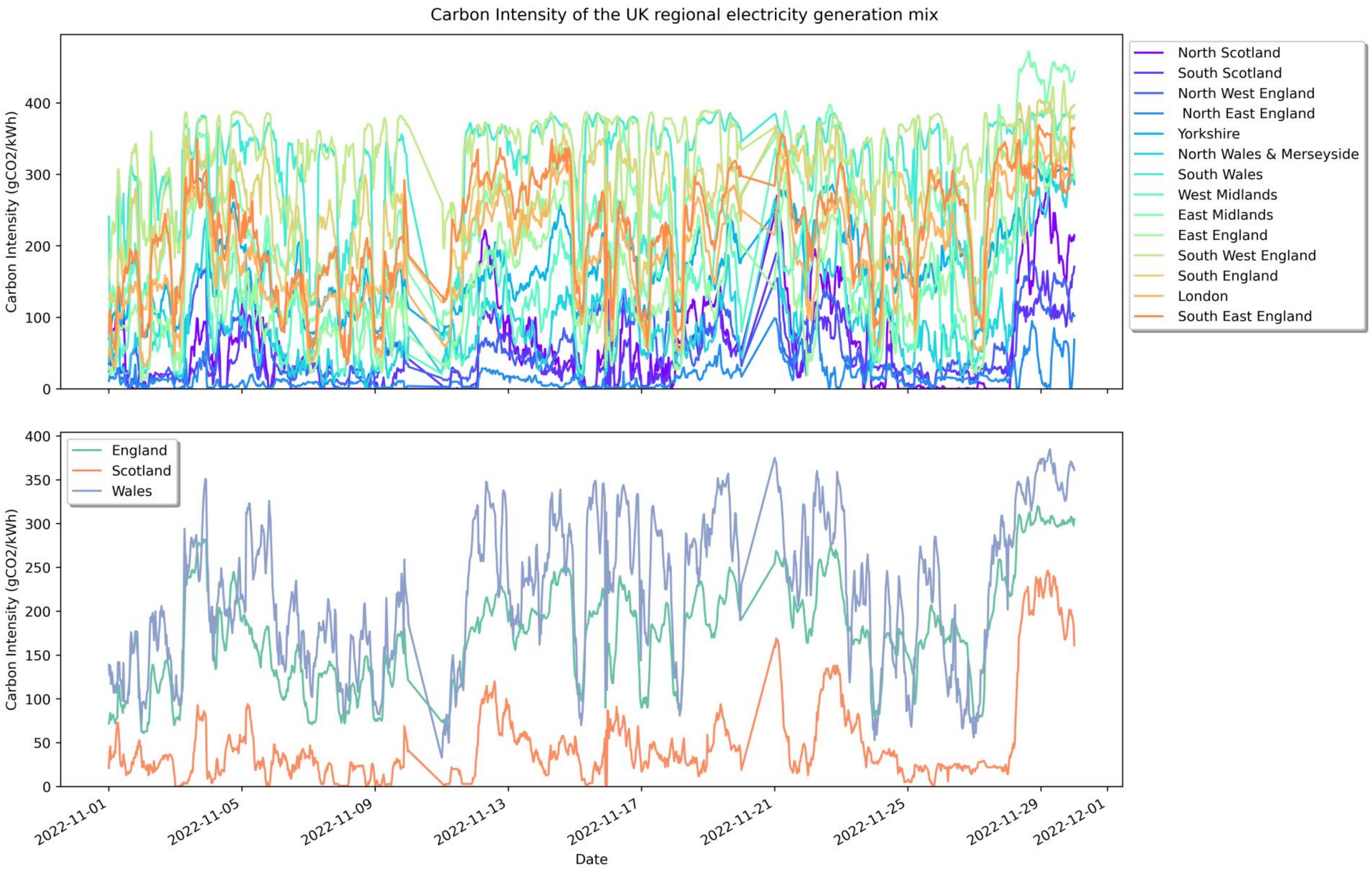
- Carbon model utilises energy values and a energy carbon intensity factor to create an overall carbon usage estimate

$$C_{ax}^{p} = E_{x}^{p} \times CM_{e}^{p}$$

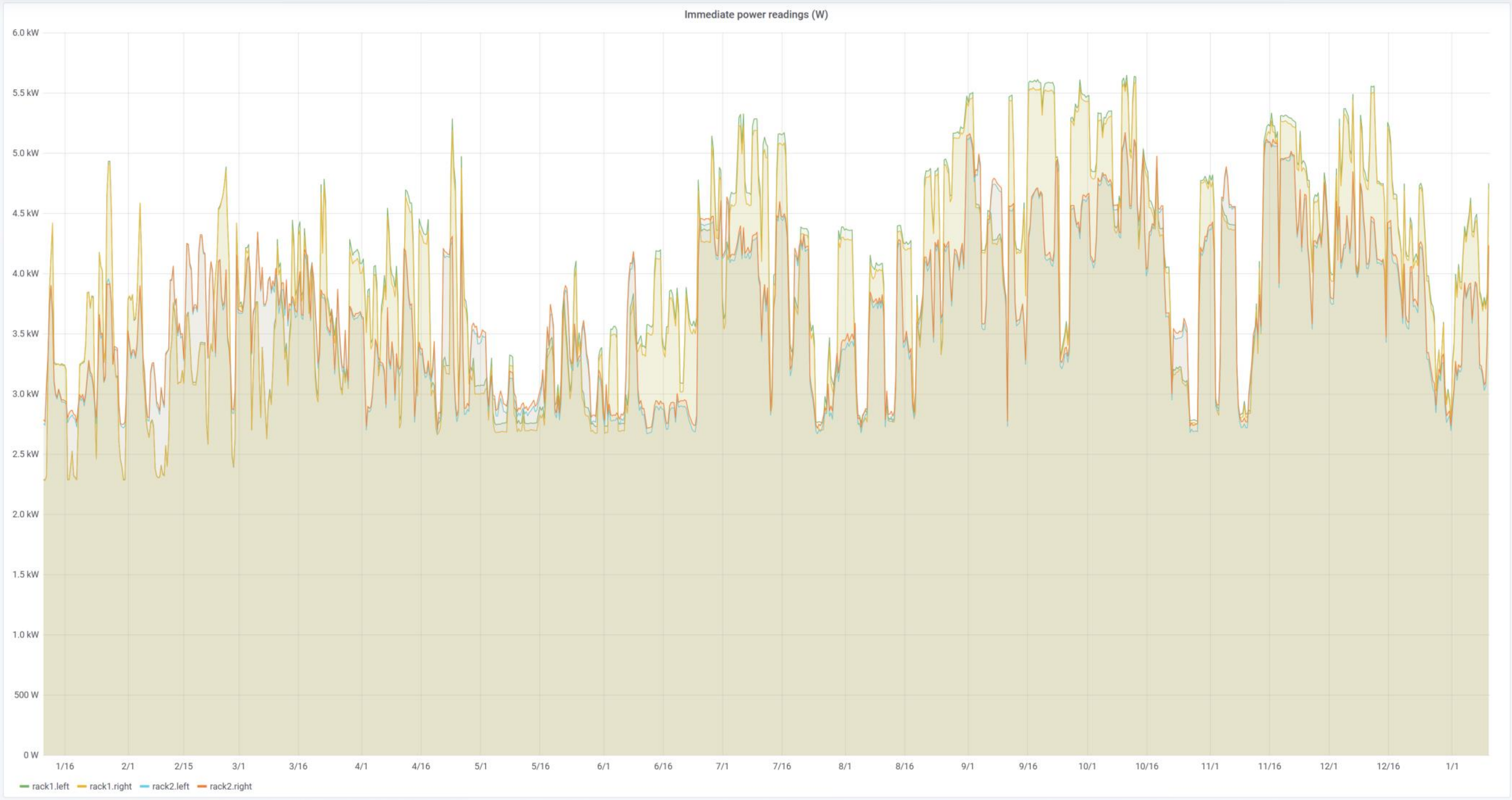
- Carbon intensity factor can be derived from various forms
  - Embodied carbon has multiple components
    - Electrical carbon intensity when manufacturing
    - Raw material extraction costs
    - Transport
    - Decommissioning and recycling

Carbon intensity

- Lower-level details are also available
  - Could account for local hosting conditions
  - Not currently considering this level of detail



Active energy usage

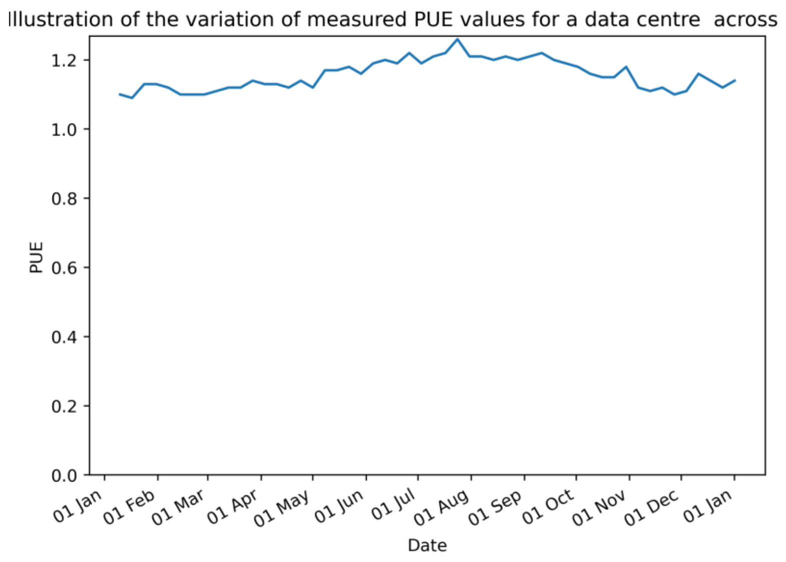


Collected active energy

System	Cumulative energy used (kWh)
QMUL	1299.7
CAM	261.5
DUR	8699.9
STFC CLOUD	3903.2
STFC SCARF	4271.3
IMP	943.9
Total	**19379.5**

Metric	Carbon Intensity Value (gCO_2/kWh)	Active Energy Carbon (C_{ax}^{p}) ($kgCO_2$)
Low	50	969
Medium	175	3391.4
High	300	5813.8

PUE and Active Carbon Estimate



- Facility level active energy
  - Not collected for the IRISCAST snapshot process
  - Modelling using a range of PUE factors
    - Low: 1.1
    - Medium: 1.3
    - High: 1.6

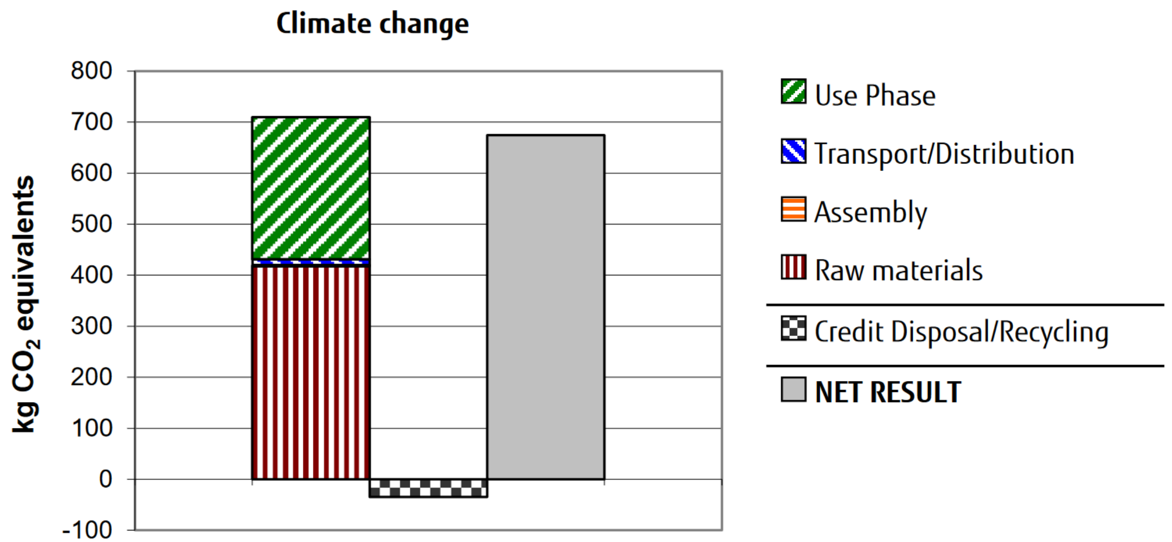
Metric	Low			Medium			High		
Active Energy Carbon (kgCO2)	969			3391.4			5813.8		
PUE Estimate	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Active Energy Carbon including Facilities (kgCO2)	1066	1260	1550	3731	4409	5426	6395	7558	9302

Lowest low: **1066 kgCO2** Highest high: **9302 kgCO2**

Embodied carbon estimates

- Estimated two values based on single node (*made up values*):
  - 400 $kgCO_2$ 1100 $kgCO_2$
- Amount of embodied carbon to be attributed to the snapshot varies depending on lifetime assumptions
  - How long you keep the hardware affects how much embodied carbon gets apportioned to a given period

	Embodied carbon estimate ($kgCO_2$)		Embodied carbon estimate ($kgCO_2$)	
	400	1100	400	1100
Server lifespan estimate	Embodied emission ($kgCO_2$ per 24 hours)		Snapshot embodied emissions ($kgCO_2$)	
3	0.36	1.00	876	2409
4	0.27	0.75	657	1806
5	0.22	0.61	526	1445
6	0.18	0.50	438	1204
7	0.16	0.43	375	1032



https://www.fujitsu.com/global/documents/about/environment/Life%20cycle%20analyses%20of%20Fujitsu%20Desktop%20ESPRIMO%20P9010%20June%202021.pdf

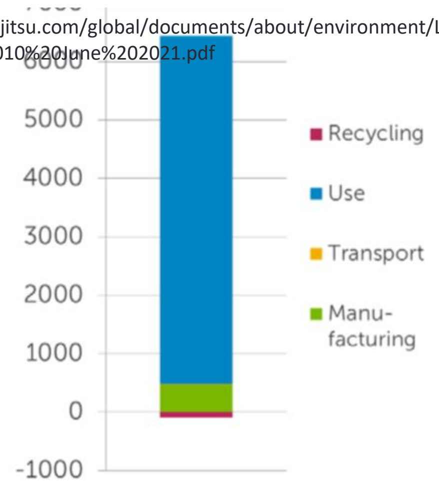
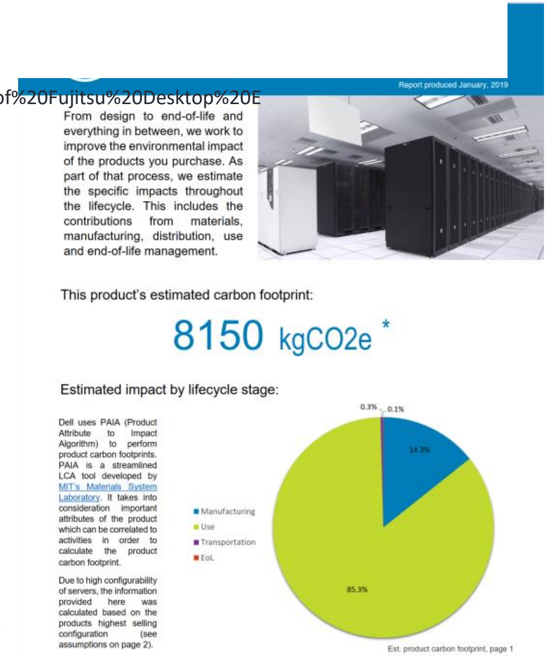


Figure 2: Total product carbon footprint [kg CO_2eq] of the Dell PowerEdge R710 in the US.

Manufacturing, which includes raw material extraction, subassembly manufacturing, transportation of the subassemblies and final assembly in Austin, TX has the next biggest impact, but at 7% (471 kg CO_2eq, green) it represents only a fraction of the total GHG emissions.

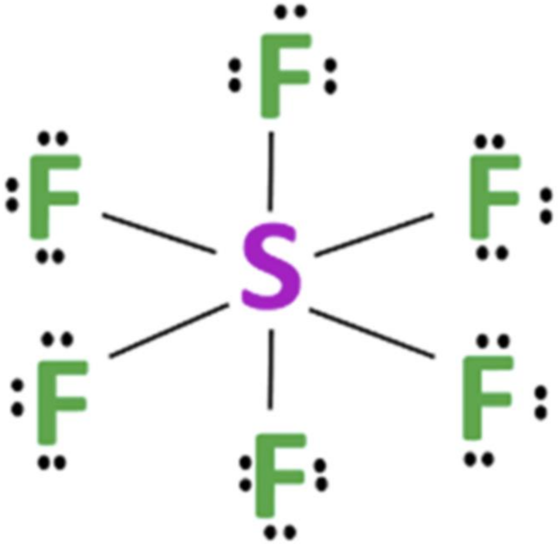


https://i.dell.com/sites/csdocuments/CorpComm_Docs/en/carbon-footprint-poweredge-R430.pdf

https://i.dell.com/sites/content/corporate/corp-comm/en/Documents/dell-server-carbon-footprint-whitepaper.pdf

Other stuff

- Buildings
  - Data centre building have some contribution
- Infrastructure embodied carbon
  - No inclusion of cooling and power hardware in the embodied estimates, needs to be included
- Transformer, coolant, and fire suppressant gases?
  - 1kg of SF_6 is approximately equivalent to 23,500kg of CO_2
  - Potential for hidden impacts like these



Total estimated impacts (24 hour period)

		Total carbon footprint estimate ($kgCO_2$) (Percentage active carbon)		
Server embodied carbon estimate ($kgCO_2$)	**Server lifespan estimate**	**Low – Low Active Carbon Assumptions**	**Medium – Medium Active Carbon Assumptions**	**High – High Active Carbon Assumptions**
400	3	1950 (55%)	5293 (83%)	10186 (91%)
	5	1600 (67%)	4943 (89%)	9836 (95%)
	7	**1449** (74%)	4792 (92%)	9685 (**96%**)
1100	3	3483 (**31%**)	6826 (65%)	**11719** (79%)
	5	2519 (42%)	5862 (75%)	10755 (86%)
	7	2106 (51%)	5449 (81%)	10342 (90%)

Overall impact

- 24-hour period
  - Embodied carbon estimate:
    - 375 - 2409 $kgCO_2$ → 140 – 9000 $kgCO_2$
  - Active carbon range:
    - 1066 – 9302 $kgCO_2$ → 4000 – 34000 $kgCO_2$
- Active looks bigger, 2-10x bigger
  - Can address in the near term
  - Active will reduce as the energy mix gets cleaner
- Embodied has very wide error margins
  - Other things look much lower impact (i.e. buildings)
- Comparator:
  - https://www.carbonindependent.org/22.html
  - Typical flight CO_2 emissions: 92 $kgCO_2$ per passenger per hour
  - Imaginary 24-hour flight for one person 2208 $kgCO_2$

"AWS to buy Cumulus data centre"

Currently 48MW, max available 960MW, likely 480MW

Carbon model

- Things the carbon model misses?
  - Minimise movement
    - Keep data where it is
    - Key for maximising energy efficiency
  - Minimise waiting on movement
    - Get data to where it needs to be before it is required
    - Key for maximising user efficiency
  - Minimise the maximum hardware required
  - Ensure “optimal” hardware used for a particular problem/use case
  - The “value” of what is being run

Net Zero really?

- Pitfalls we need to avoid
  - Focussing on the negative, the positive really should be something we should be pushing
  - Focussing on the metrics, because easily measured metrics don't really tell you anything
  - Ignoring upgrade benefits
  - Ignoring rebound effects
  - Ignoring displacement effects
- Properly Net Zero just for DRI assumes some carbon equivalent reduction
  - Could be heat recovery schemes or some carbon capture etc…

Mitigating climate impacts from high performance computing

- Optimising operation of HPC facilities can help at a global level
  - Reduce overall power/energy required
  - More work to do on this (i.e. per job active management)
- Ensuring applications are optimised helps overall science per system/watt/£/$
- Ensuring people only run what is necessary also helps
  - Culture and behaviour research interesting on HPC usage
- Improving the efficiency of manufacturing electronics also a big part of this
  - Maintaining and repurposing hardware can help