Estimating the environmental impact of a large Tier 2

D. Whitehouse, D. Bauer, S. Fayer
Imperial College London

Imperial College London

Overview

- The Imperial College Tier 2
- What do we mean by environmental impact?
- Methodology
- Result
- Conclusion

The Imperial College Tier 2

- ~8k worker node cores, hyperthreaded, accommodating ~16k jobs
  - Mainly serving CMS, LHCb and a large number of smaller non-LHC communities
- Memory 2-8 GB/core
- Storage: 85 storage nodes providing ~22 PB
- The Tier 2 is located off campus in a commercial data centre
- Co-located with the Tier 2 are HEP group computing and cloud services
  - These are not included in this survey
- Kit is run until it becomes unreliable rather than operate on a fixed replacement schedule.
  - Typically service life is about 8-10 years

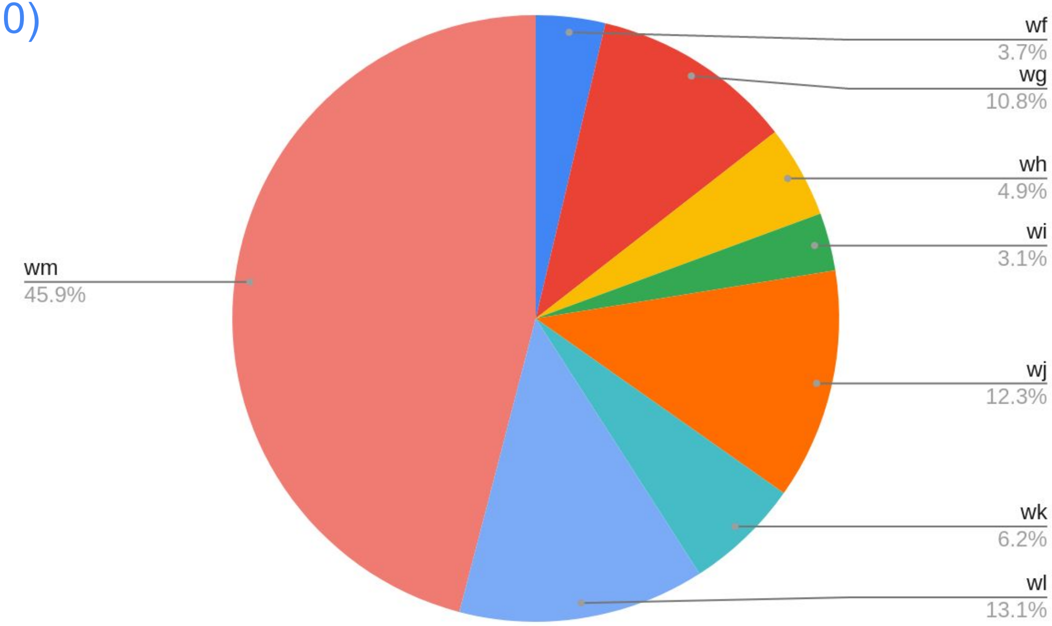
Data centre

- The Tier 2 is hosted at the Virtus (https://virtusdatacentres.com) “London 3” Data Centre in Slough
- 3,000 m^2 net technical space
- 10 MVA incoming supply from National Grid
- 7.2 MW of total IT load
- Design power usage effectiveness (PUE) of ~ 1.3
  - Waste heat reuse in communal parts
- Previously the Tier 2 was hosted on campus in a converted lab space
  - No hot air containment

Imperial College London

Compute

wf: HP ProLiant SL2x170z G6 (2010)
wg: Dell PowerEdge R410 (2011)
wh: Supermicro X9DRT (2014)
wi: Supermicro X10DRT-P (2016)
wj: Dell PowerEdge R430 (2017)
wk: Dell PowerEdge R440 (2019)
wl: Supermicro H11DSU-iN (2020)
wm: Dell PowerEdge R6525 (2020)



WN distribution by core
Total: ~8000 cores

Imperial College London

Storage

sedsk53-60: Dell PowerEdge R510 (2012)

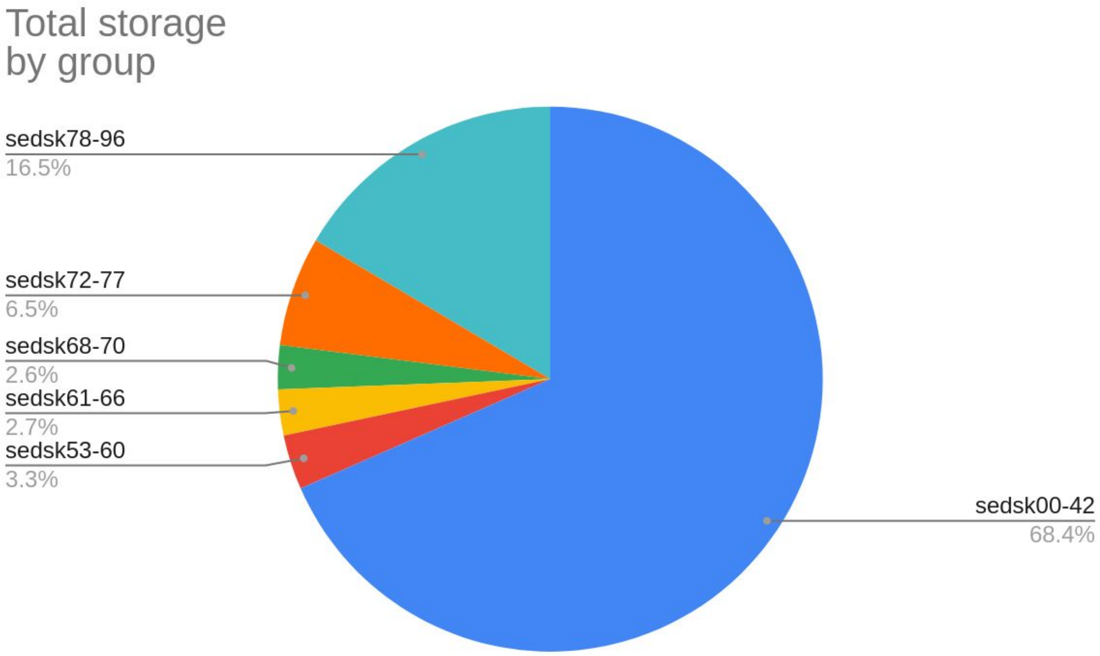
sedsk61-66: Dell PowerEdge R720xd (2014)

sedsk68-70: PowerEdge R730xd (2016)

sedsk72-77: Supermicro X10DRi (2018)

sedsk78-96: Dell PowerEdge R740xd (2019)

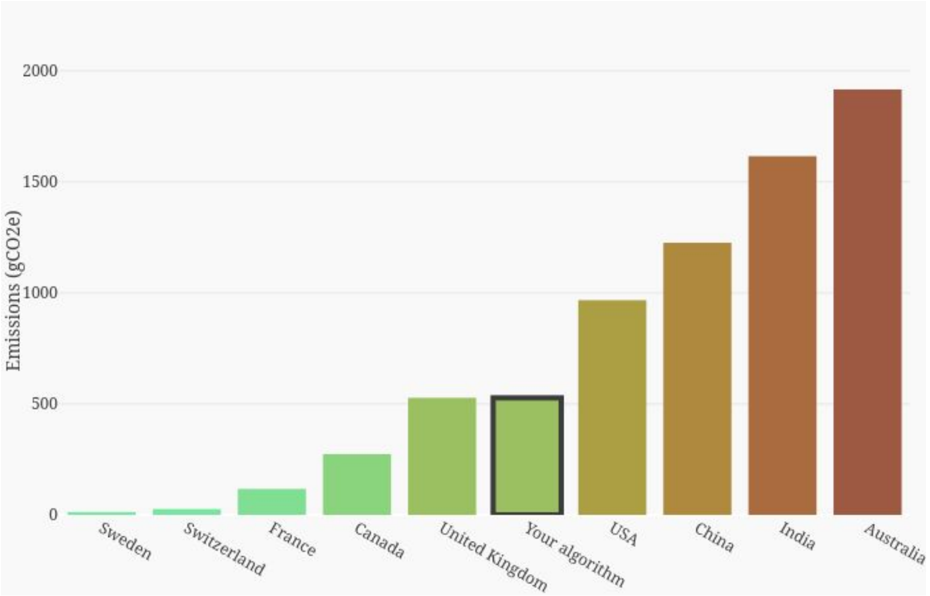
sedsk00-42: Dell PowerEdge R740xd2 (2021/22)



Imperial College London

What do we mean by environmental impact?

- Carbon footprint:
  - **Power consumption**
  - Manufacture (limited information)
- Electricity carbon intensity vary by country
  - We use the UK value
- Data centre uses electricity from renewable resources
  - Good, but this does not give us a free pass



source: http://calculator.green-algorithms.org/

Imperial College London

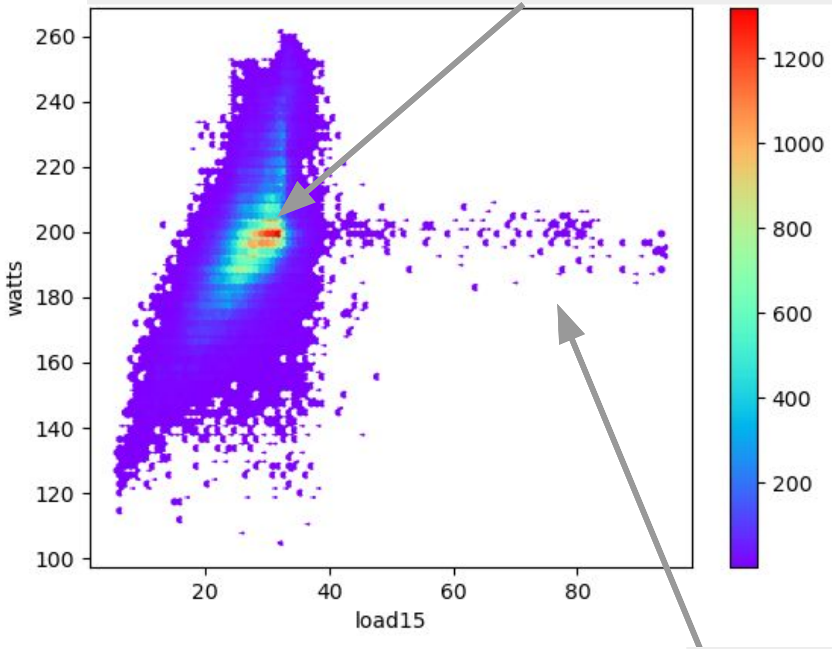
Methodology

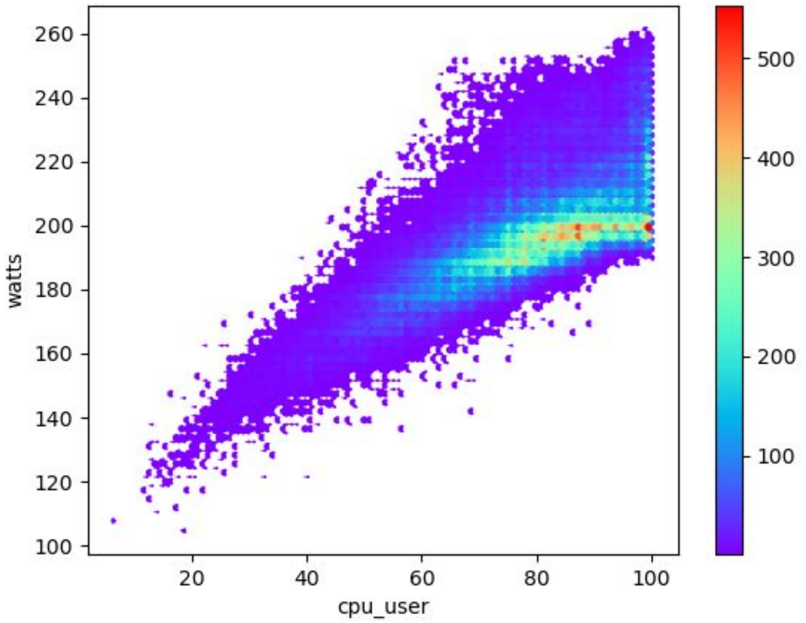
- Use DCMI power interface (via ipmitool) to sample the instantaneous power usage of each machine every 5 minutes.
  - Very old worker nodes (~15% of cluster) do not have power monitoring available: Estimate power consumption from thermal design power (TDP)
  - Older storages nodes with additional chassis for disks do not monitor power to chassis: Estimated values from literature (~15% of cluster)
  - Data set for this talk: February 2023 (28 days)
  - Investigated PDU monitoring, however this is aggregated by phase, which tends to be shared by different types of machines.
  - Where the node type is homogeneous across a single phase, the PDU measurement is used as a cross-check.
- Use vendor provided information on carbon footprint where available.
  - Information is limited and difficult to validate against our actual setup.
- Electricity carbon intensities are country specific: UK 231.12 gCO_2e/kWh (2019)

Imperial College London

load15 and cpu_user vs power usage: Supermicro X9DRT (2014)

Overall load and power usage are fairly constant.



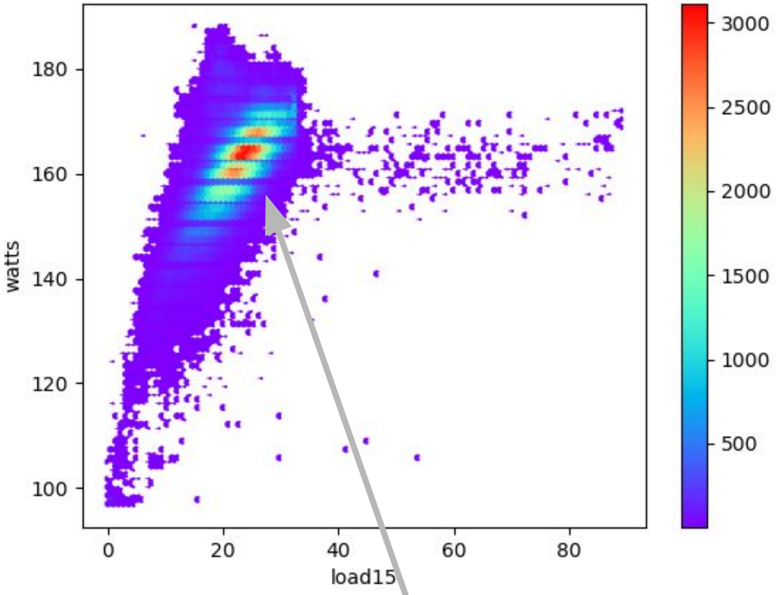


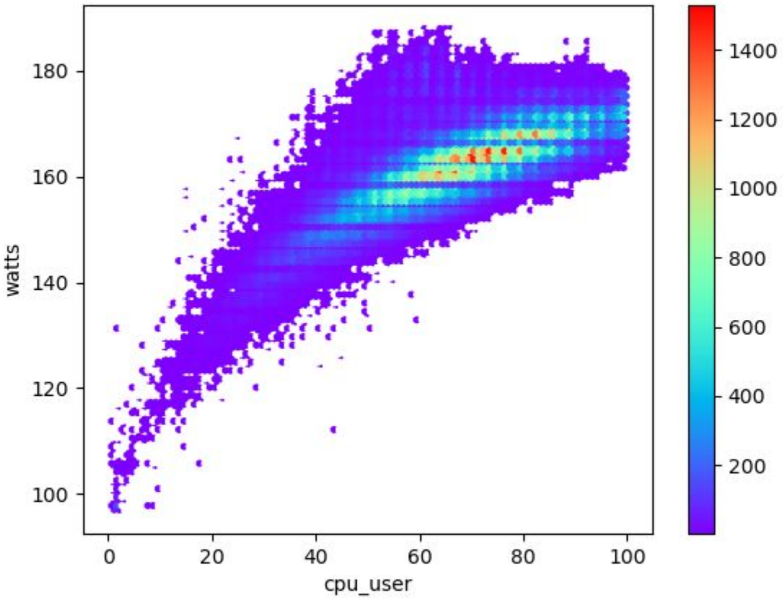
CVMFS issues? Multicore jobs in single core slots?

Imperial College London

load15 and cpu_user vs power usage: Dell PowerEdge R430 (2017)

Little difference in the distribution in newer nodes wrt previous generation.



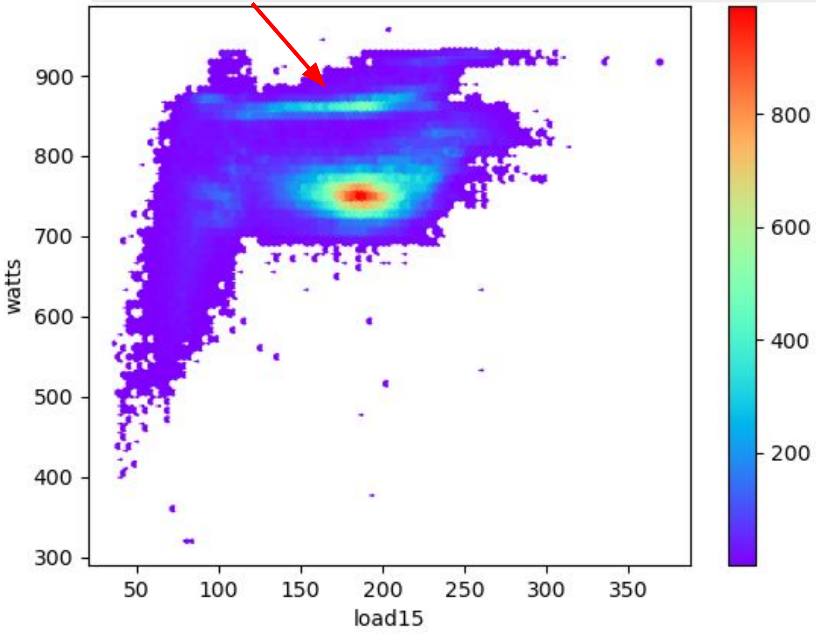


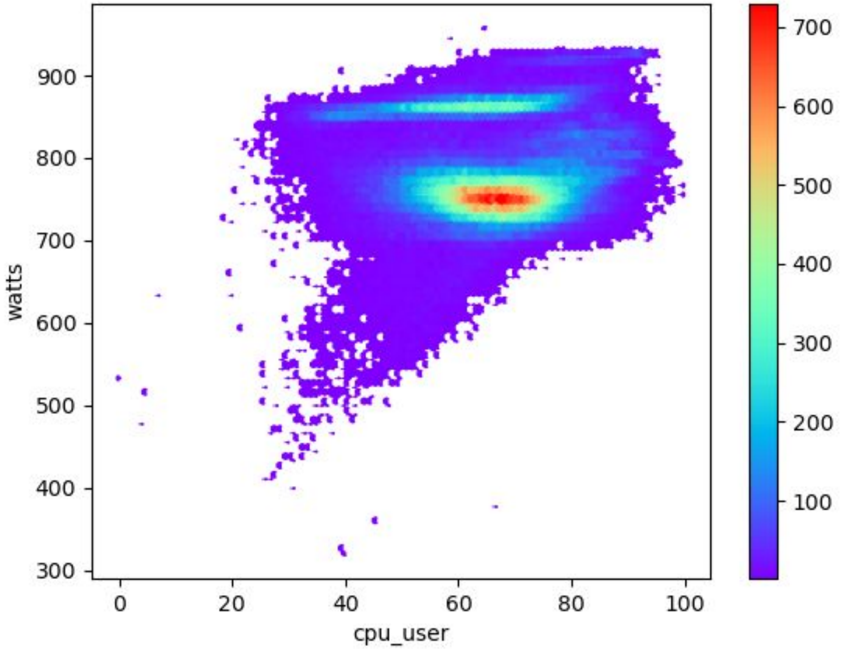
lines are an artefact of the power measurement resolution (~15 W)

Imperial College London

load15 and cpu_user vs power usage: Dell PowerEdge R6525 (2020)

Possibly an architecture feature artifact (Turbo Core)?





Imperial College London

Worker nodes

Power usage/core suggest that our policy of running nodes until they die might keep our capacity up, but might not be the best for the environment.

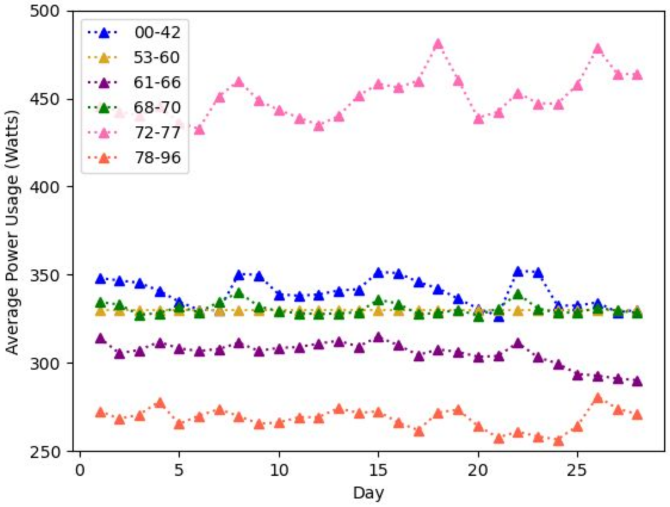
wf:	HP ProLiant SL2x170z G6	(2010)	17.0 W/core
wg:	Dell PowerEdge R410	(2011)	14.2 W/core
wh:	Supermicro X9DRT	(2014)	12.3 W/core
wi:	Supermicro X10DRT-P	(2016)	10.2 W/core
wj:	Dell PowerEdge R430	(2017)	10.1 W/core
wk:	Dell PowerEdge R440	(2019)	10.9 W/core
wl:	Supermicro H11DSU-iN	(2020)	5.4 W/core
wm:	Dell PowerEdge R6525	(2020)	6.1 W/core

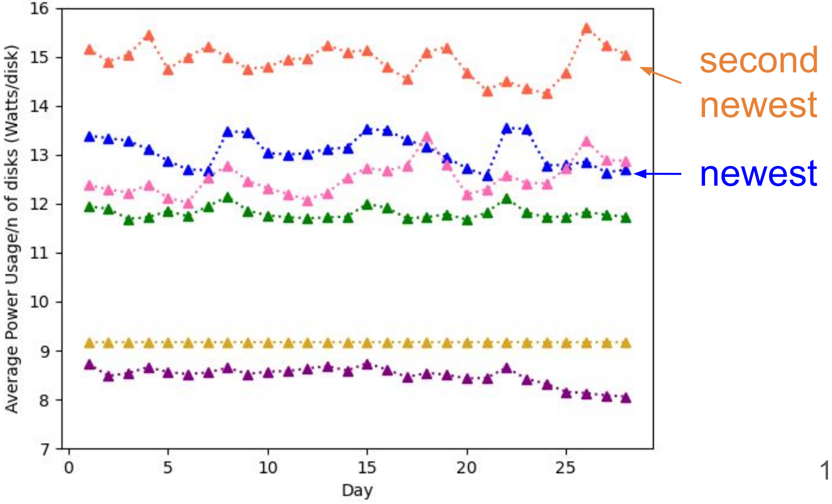
Storage

Group	Manufac-turer	Year of Purchase	All disks/ active disks	Disk Size (TB)	Storage per server (TB)
00-42	Dell	2020/21	26/22	16	352
53-60*	Dell	2012	36/30	3	90
61-66**	Dell	2014	36/30	4	120
68-70	Dell	2016	28/24	8	192
72-77	Supermicro	2018	36/30	8	240
78-96	Dell	2019	18/16	12	192

*estimate
**measured + estimate for extra chassis

Number of disks alone is not the whole picture.



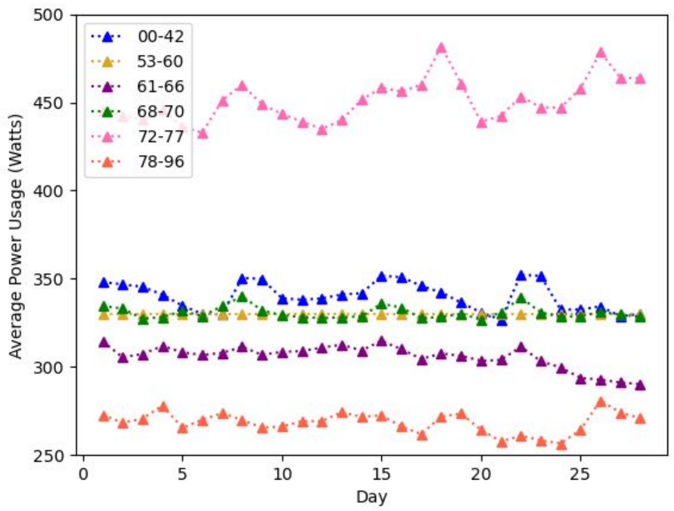


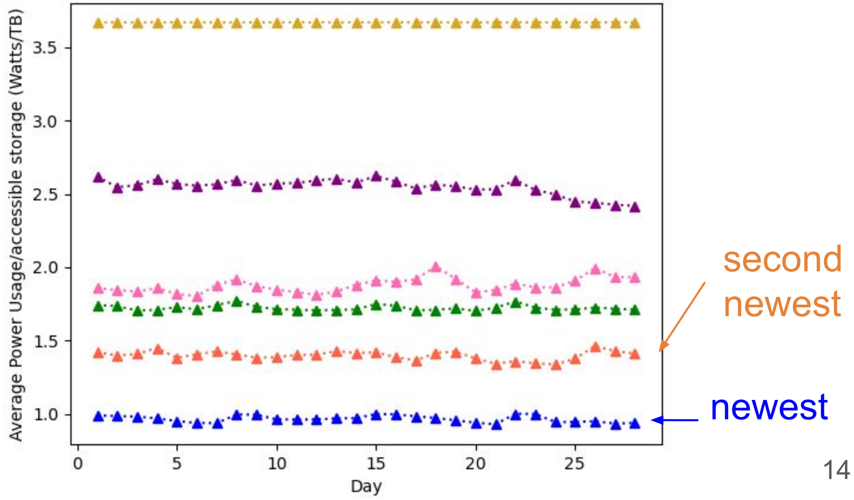
Imperial College London

Storage

Group	Manufac-turer	Year of Purchase	All disks/ active disks	Disk Size (TB)	Storage per server (TB)
00-42	Dell	2020/21	26/22	16	352
53-60*	Dell	2012	36/30	3	90
61-66**	Dell	2014	36/30	4	120
68-70	Dell	2016	28/24	8	192
72-77	Supermicro	2018	36/30	8	240
78-96	Dell	2019	18/16	12	192

*estimate
**measured + estimate for extra chassis





However, newer storage servers use less energy per usable TB. Phew!

Imperial College London

Results: Power consumption: Worker nodes

Node group	Power measure-ment in Watts	Power per node group in Watts	Power per core in Watts	Energy per year (kWh) per node	Energy per year per node group (kWh)	kg CO2e/year per node	PUE = 1 total kg CO_2e/year	actual PUE = 1.3 total kg CO_2e/year
wf (2010)	136	4896	17.0	1191	42889	382	13,767	17,898
wg (2011)	171	11970	14.2	1498	104857	481	33,659	43,757
wh (2014)	197	4728	12.3	1726	41417	554	13,295	17,283
wi (2016)	204	2448	10.2	1787	21444	574	6,884	8,949
wj (2017)	162	9720	10.1	1419	85147	456	27,332	35,532
wk (2019)	351	5265	10.9	3075	46121	987	14,805	19,246
wl (2020)	693	5544	5.4	6071	48565	1,949	15,590	20,266
wm (2020)	784	21952	6.1	6868	192300	2,205	61,728	80,247
SUM		66,523 W			582,741 kWh		187,060 kg CO_2e	243,178 kg CO_2e

Imperial College London

Headline figure:
HEPSPEC06 hours delivered (12 months) 1,498,292,348 -> 0.162 g/HEPSPEC06/hour

Results: Power consumption: Storage nodes

Group	Average power usage (W)	power per node group (W)	energy per node group/year (kWh)	energy per TB/year (kWh)	kg CO2e/ year/TB	PUE = 1 total kg CO2e/year	actual PUE = 1.3 total kg CO2e/year
00-42 **(2020/21)**	340	14,620	128,071	8.5	2.0	29,584	38,460
53-60 **(2012)**	330	2,640	23,126	32.1	7.4	5,342	6,945
61-66 **(2014)**	306	1,530	13,403	22.3	5.2	3,096	4,025
68-70 **(2016)**	331	993	8,699	15.1	3.5	2,009	2,612
72-77 **(2018)**	451	2,706	23,705	16.5	3.8	5,476	7,118
78-96 **(2019)**	268	5,092	44,606	12.2	2.8	10,304	13,395
SUM		27,581 W	241,610 kWh			55,812 kg CO_2e/year	72,555 kg CO_2e/year

Imperial College London

Headline figure: 3.3 kg CO_2/year/TB

Manufacturing Carbon Footprint

- New servers ([Dell-R6525](#) in this example) have started being sold with catchy "lifetime" carbon footprint.
- Manufacture carbon footprint retrieved by squinting.
- No (public) information available for older nodes.
- Carbon footprint due to usage depends heavily on local factors, but we can measure it - unlike for manufacture.

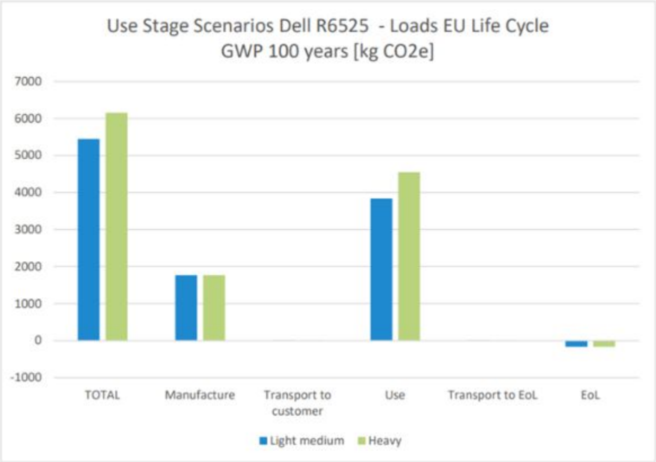
5440 kgCO2e +/-*

Estimated impact by lifecycle stage is outlined in the below graph:

*The product carbon footprint data generated in this report was created using the GaBi 10 Software system for life cycle engineering, developed by Sphera Solutions Inc.

To view the full Life Cycle Analysis report click here.

Documentation for all GaBi datasets can be found online (Sphera Solutions Inc., 2020).



Imperial College London

Manufacturing Carbon Footprint - Fineprint

- That’s not quite our server.
- If you are a vendor reading this slide, can our kit please come with manufacturing carbon footprint of the actual server we have just bought?

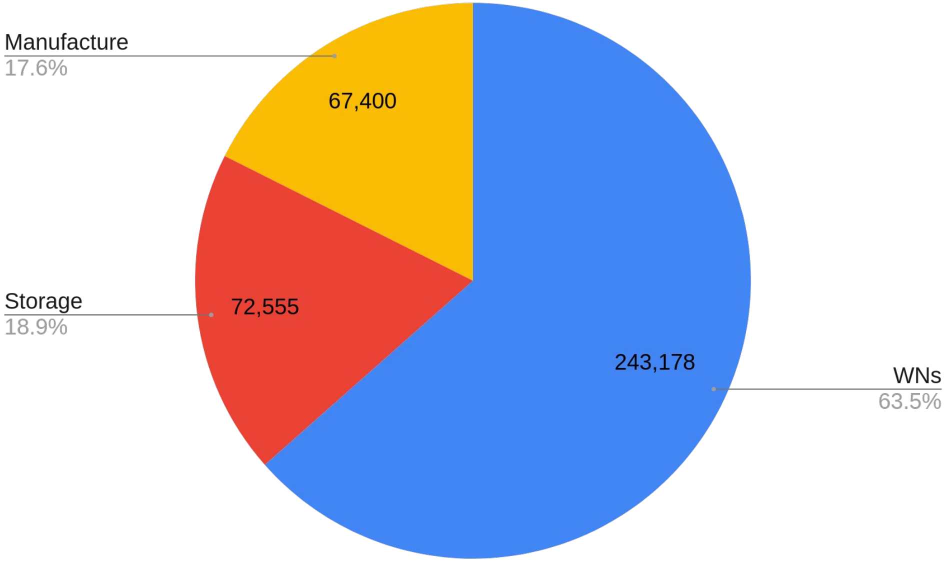
Table 1: Assumptions

	Assumptions			
Component	R6515	R7515	R6525	R7525
Lifetime of product	4 Years			
Use location	EU & US			
Mainboard	12 layers, OSP fin-ishing		14 Layers	
Memory	8x Micron 16GB		16x Micron 16GB	
Hard Drives	2x 4TB SATA			
Processor	1x AMD EPYC 7452		2x AMD EPYC 7452	
Chassis	1U Rack Chassis with up to 4 x 3,5" Hard Drives for 1CPU configuration incl. rails and bezel	2U Rack Chassis with up to 8x3,5" Drives for 1 CPU Configuration incl. rails and bezel	1U Rack Chassis with up to 4x3,5" Drives for 2 CPU Configuration incl. rails and bezel	2U Rack Chassis with up to 8x3,5" Drives for 2CPU Con-figuration incl. rails and bezel

- Total number of machines: 84 storage servers and 253 worker nodes.
- 1600 kg CO_2e per node assuming 8 year lifetime: 200 kg per node per year: Total for site: 67,400 kg CO_2e.

What follows is a back of the envelope calculation….

Summary: kg CO_2e per year per Tier 2



PUE = 1.3: **383,000 kg** = 445 sysadmins crossing the Atlantic in a plane*

*One sysadmin going from LHR to JFK: 860 kg CO_2 e

Conclusion

- Full carbon equivalent impact not easy to estimate accurately.
- Highly dependent on manufacturer for information that we can't measure, e.g. embodied carbon cost.
- In-built server power monitoring good, but could be improved:
  - Careful PDU bank to Server mapping would allow better cross-checking of results.
  - Servers could contain more power sensors for specific components.
- Things are improving:
  - Greener energy supplies.
  - Heat reclamation becoming more commonplace.
  - Increasing communication with vendors around embodied carbon.

Backup: References

Department for Business, Energy and Industrial Strategy: *2022 Government Greenhouse Gas Conversion Factors for Company Reporting*

https://accounting.egi.eu/egi/

Imperial College London