Estimating the environmental impact of a large Tier 2

Daniela Bauer[1,*], *Simon* Fayer[1], and *Dan* Whitehouse[1]

[1]Blackett Laboratory, Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom

Abstract.
Recent years have seen an increasing interest in the environmental impact, especially the carbon footprint, generated by the often large scale computing facilities used by the communities represented at CHEP. As this is a fairly new requirement, this information is not always readily available, especially at universities and similar institutions which do not necessarily see large scale computing provision as their core competency. We present the results of a survey of a large WLCG Tier 2 with respect to power usage and carbon footprint leveraging all sources of information currently available to us. We show that it is possible to estimate the environmental impact with respect to power usage without having to invest in dedicated monitoring equipment. Manufacturers however do not yet provide sufficient information to allow for a detailed analysis of the carbon footprint of equipment manufacture, but even with the available information it is clear that this cannot be ignored.

1 Introduction

The Imperial College High Energy Physics (HEP) group hosts a large Tier 2 WLCG site consisting of approximately 8000 compute cores and 22 PB of storage. It serves the CMS, LHCb and ATLAS experiments, plus a number of smaller non-LHC communities carrying out research in areas as diverse as neutrino physics, astronomy and bio-informatics. Unlike other university based Tier 2 sites in the UK which host their Tier 2s on-site, the Imperial College Tier 2 is located in a commercial data centre[1] approximately 30 km to the west of the College campus. Until 2019 the Tier 2 was hosted on campus in a converted laboratory space, without hot air containment; in our experience the challenge of keeping equipment at adequate temperatures in these conditions outweighs the convenience of access to e.g. carry out repairs, hence the overall availability and reliability of the cluster has increased since moving into commercial space. The Tier 2 is co-located with the HEP group's computing, a cloud service for non-WLCG communities and a number of service nodes; these are excluded from the survey.

Our survey focuses on the two main contributing factors to the carbon footprint: power consumption and manufacture. Currently vendors do not routinely supply estimates of the manufacturing carbon footprint for specific configurations, generally only producing an example per node type, if any at all. As the result of this survey will show, the information provided is not yet sufficient to be used as a basis for purchasing decisions; however this is

*e-mail: lcg-site-admin@imperial.ac.uk

slowly changing. As is common among Tier 2s in the UK, we run most of our hardware until it becomes unreliable rather than operate a fixed replacement schedule. In practice this means that we retire most hardware around the 8-10 year mark.

In principle the carbon footprint of the power used by the Tier 2 depends on the source of the electricity and the amount of heat that can be recycled. While the rise of so-called green energy is certainly encouraging, we decided to use the *average* carbon intensity in the UK[2] as our conversion factor, as this represents the most realistic estimate of the environmental impact resulting from power consumption. It should be noted that the carbon footprint due to electricity usage is highly country specific. Figure 1 shows an example on how the carbon footprint changes depending on where the power was generated.

The data centre that we use is operated by Virtus[1] who operate a chain of facilities in the UK. They provide 3000 square meters of technical space and have a 10 MVA incoming supply from the National Grid with an IT load of 7.2 MW total. The advertised power usage effectiveness (PUE) is 1.3, i.e. for every unit of power consumed another 30% is spent on providing cooling, lighting and other ancillary functions. Some of the waste heat is reused in communal areas of the facility.

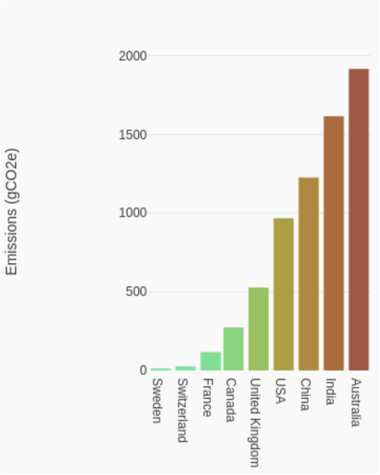


Figure 1: Country dependent carbon footprint. Figure by the Green Algorithm project[3].

2 Tier 2 Hardware Overview

Our compute cluster is described in table 1. Approximately 60% of the cores were purchased within the last 3 years, on the other hand 15% of cores are older than 10 years. The corresponding storage is shown in table 2. Due to the increasing size of storage disks, the storage capacity is dominated by the newer (2019 and later) storage nodes, which make up 85% of our storage capacity and 75% of our storage nodes. We have not taken switches, head-nodes for compute and storage and hardware we use for monitoring into account.

3 Power monitoring

We use the DCMI[4] interface via ipmitool[5] to monitor the power usage of our nodes. We sample the instantaneous power usage every 5 minutes and store it in a PostgreSQL

Node group	Year of manufacture	Vendor	Model	Total number of cores in group
wf	2010	HP	ProLiant SL2x170z G6	288
wg	2011	Dell	PowerEdge R410	840
wh	2014	Supermicro	X9DRT	384
wi	2016	Supermicro	X10DRT-P	240
wj	2017	Dell	PowerEdge R430	960
wk	2019	Dell	PowerEdge R440	480
wl	2020	Supermicro	H11DSU-iN	1024
wm	2020	Dell	PowerEdge R6525	3584

Table 1: Tier 2 compute nodes as of May 2023.

Node Group	Vendor	Year of Manufacture	Total / Active Disks	Disk Size in TB	Storage / Server in 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

Table 2: Tier 2 storage nodes as of May 2023.

database together with the current load and CPU usage as reported by the Linux kernel. For some very old worker nodes (~15% of the cluster) no power monitoring was available, instead we estimated their power usage from the thermal design power (TDP) of their CPUs. Similarly we found that some of the older storage nodes with an additional disk-array chassis did not monitor the power supplied to the chassis, however we were able to estimate their power consumption from literature[6]. Some very old storage nodes had no power monitoring available at all, here we used the power consumption as measured from the most similar model available.

We investigated using the usage provided by our power distribution units (PDUs), however these report power values aggregated by phase. The majority of our PDUs have multiple classes of machine on a single phase, preventing any sensible use of the values. Where the node type was homogeneous across a phase we used the data from the PDUs as a cross check to validate the output from the DCMI interface. Both sets of data were found to be consistent.

The dataset analysed in these proceedings was taken in February 2023 and comprises 28 days worth of readings. All code relating to this project is stored in an internal repository and can be made available on request.

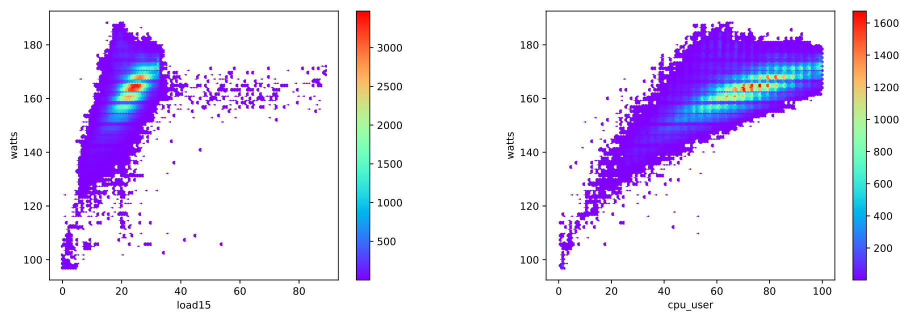
4 Power usage

4.1 Worker nodes

As the job pressure on the Tier 2 cluster is fairly high and we run a mix of jobs on each node, we expected the load and power usage to be fairly constant and this was indeed the case, as can be seen in figures 2 and 3. We suspect that the long tails seen in the load values in figure 2 are either due node configuration issues (e.g. where a software area on a node is not correctly mounted resulting in multiple fast job failures), or multi-core jobs that were assigned to a single core slot; however these events are rare. The distributions look similar for all generations of nodes, with the exception of the Dell PowerEdge R6525, shown in figure 3 which shows a distinct high peak power usage artefact. We attribute this to the "Turbo Core" feature of the processor architecture.

When normalising power usage to the number of cores we see a factor of three in reduction of power consumption over the course of a decade. Details are given in table 3 and table 4 summarises the power consumption of all worker nodes in the Tier 2.

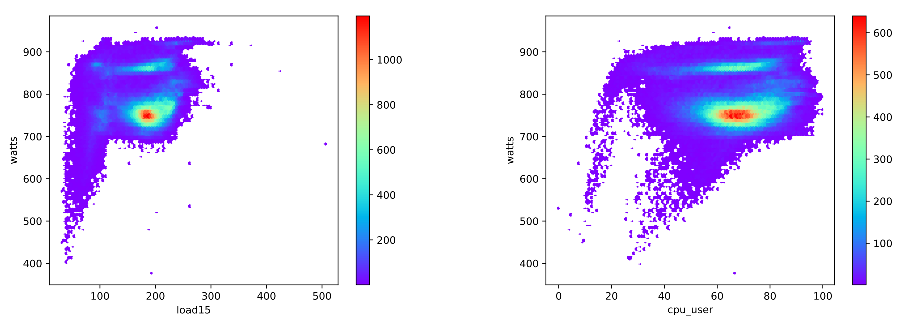
WLCG workloads are accounted for in HEPSPEC06-hours[7]. The Tier 2 site delivered 1,498,292,348 HEPSPEC06-hours between 1st March 2022 and 28th February 2023. This corresponds to a headline figure of 0.162 gCO_2e per HEPSPEC06-hour.



(a) Power usage in Watts vs Load15

(b) Power usage in Watts vs user CPU usage (%)

Figure 2: Dell PowerEdge R430, 2017 (cf. *wj*-group in table 1).



(a) Power usage in Watts vs Load15

(b) Power usage in Watts vs user CPU usage (%)

Figure 3: Dell PowerEdge R6525, 2020 (cf. *wm*-group in table 1).

Node group	Year of purchase	Power measurement in Watts	Power per core in Watts
wf	2010	136	17.0
wg	2011	171	14.2
wh	2014	197	12.3
wi	2016	204	10.2
wj	2017	162	10.1
wk	2019	351	10.9
wl	2019	693	5.4
wm	2020	784	6.1

Table 3: Power usage per core in different generations of worker nodes. Worker nodes are running predominantly WLCG-type workloads.

Power in Watts	kWh/year	$kgCO_2e$/year PUE=1	$kgCO_2e$/year PUE=1.3
66523	582741	187060	243178

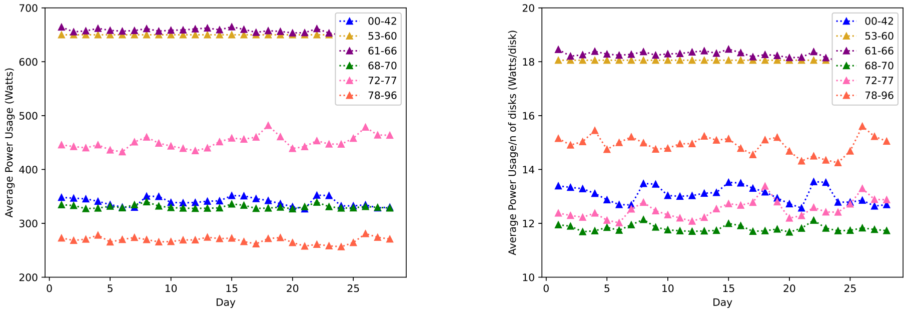
Table 4: Combined power usage of Tier 2 worker nodes. PUE=1 is given for reference, as data centre efficiencies are improving: For example Google currently reports an average PUE of 1.10 across their data centres[8].

4.2 Storage nodes

Similarly to our compute cluster, the Tier 2 storage hosts a mix of data for the different communities we support. We do not have storage nodes that are dedicated to a single experiment or data type. As with the compute we see that power consumption of individual node groups is fairly constant, as illustrated in figure 4(a). We were not able to monitor the power usage of our oldest nodes directly, and have used values found in literature. Hence for the oldest group (53-60 (2012)), we use a constant value of 650 W per server, while for our second oldest group (61-66 (2014)), we were able to measure the power usage of the main storage server and add a constant value of 400 W for the connected storage chassis.

Over time the storage capacity of individual nodes has increased, mainly through an increase in disk size (see table 2 for details). Our initial assumption that power consumption would scale primarily with the number of disks contained in a server was not borne out by our data, see 4(b). However as figure 5 shows, the power needed to provide 1 TB of accessible storage does indeed drop over time, with newer nodes being more efficient. Assuming our estimates for the oldest nodes are correct, these account for 25% of our storage carbon footprint but only for 5% of storage total (see figure 6). This is clearly not ideal, even if the over all contribution of these old nodes to the Tier 2 footprint is only around 5%. The differences for nodes manufactured in 2016 and later are much less pronounced.

Averaged over all nodes (see table 5) this allow us to estimate a headline figure of 3.7 $kgCO_2e$ per year per TB.



(a) Average power usage per node group

(b) Daily average per disk

Figure 4: Daily averages by node group. Note that the values for the oldest nodes (groups 53-60 (2012) and 61-66 (2014)) are estimates.

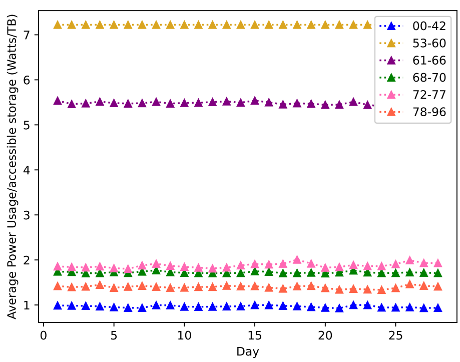
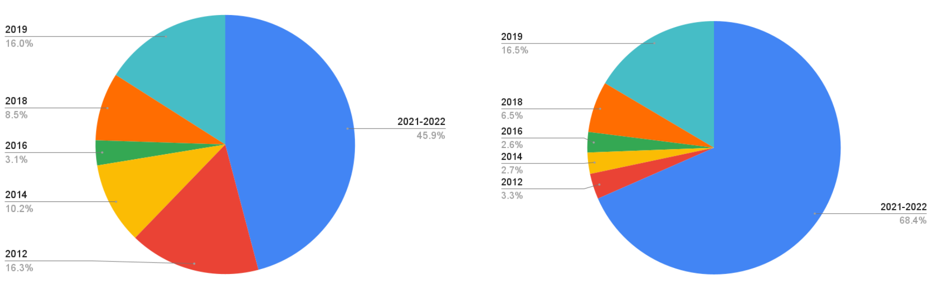


Figure 5: Daily average disks per TB.



(a) Carbon Footprint

(b) Accessible Storage Space

Figure 6: (a) Carbon Footprint from power usage of Tier 2 storage nodes grouped by year of manufacture. Note that the usage for the two oldest groups of nodes is based on estimates, (b) Accessible storage space grouped by year of manufacture.

Power in Watts	kWh/year	$kgCO_2e$/year PUE=1	$kgCO_2e$/year PUE=1.3
31876	279102	64473	83814

Table 5: Summary power usage for storage. Again, PUE=1 is given for reference, as data centre efficiencies are improving.

5 Manufacturing Carbon Footprint

As concern for the environment increases, manufacturers are providing increasingly detailed information about the manufacturing carbon footprint. However the information currently publicly available usually refers to sample servers, rather than to the particular specification of nodes as purchased, see e.g. [9]. But even with this limited information it is clear that manufacturing accounts for a non-negligible fraction of the carbon footprint of our Tier 2. Here we assume 1600 $kgCO_2e$ per node as an estimate of the manufacturing carbon footprint. Using an 8 year life-cycle, i.e. 200 $kgCO_2e$/node/year, this results in a manufacturing carbon related footprint for the whole Tier 2 site of 67,400 $kgCO_2e$ per year.

6 Conclusions

Combining the contributions from electricity usage for compute, storage and the carbon cost due to manufacture, we arrive at a total of 394,392 $kgCO_2e$ per year for the Tier 2, with compute being the largest contribution at around 60%, see figure 7.

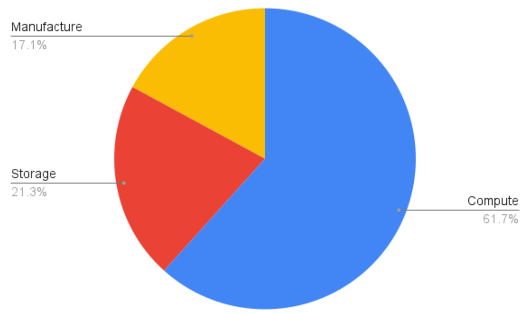


Figure 7: Overview of the different contributions to the yearly carbon footprint of the Imperial College Tier 2.

When analysing a carbon footprint it is necessary not to look at the numbers in isolation. Clearly there has been great progress over the last decade in efficiency in delivering both compute and storage. However even with the limited information currently available it is obvious that manufacture accounts for a sizeable portion of the carbon footprint. This indicates that in order to minimise the carbon footprint, one should disregard the manufacturer suggested lifetime of 4 years[9] and that our strategy to run the servers for at least 8 years is actually preferable. The issue is different when trying to optimise for power consumption alone where hardware efficiency improves on a yearly basis.

The PUE of any given data centre is generally constant. A specific PUE target is normally selected when the facility is constructed and any significant change to this would require a large amount of investment and potential downtime. Where possible, picking a data centre with a lower PUE could represent a significant amount of carbon savings over the lifetime of the contract; although it's likely that this would only be one of many considerations when deciding where to house equipment.

It is also necessary to put the overall numbers into context. Taking the popular 'metric' of comparing a carbon footprint to the carbon footprint of aircraft travel, the authors took their own journey to CHEP2023 as an example: At the authors home institution (Imperial College), business travel is assigned a carbon equivalent on booking by the travel agent[10]. For the journey LHR to JFK this amounts to 818 kgCO_2e per person. Hence the total yearly carbon footprint of the Tier 2 is the equivalent of 482 computing professionals crossing the Atlantic, or less than a single return journey of a fully occupied plane on this route. While this does not absolve us from the need to take the environmental impact of our computing work into account, there may be other aspects of our field where more significant improvements in carbon footprint can be made.

References

[1] Virtus Data Centres, https://virtusdatacentres.com, [accessed 2023-06-08]

[2] Department for Business, Energy & Industrial Strategy, *2022 Government Greenhouse Gas Conversion Factors for Company Reporting*, https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2022

[3] Lannelongue, L., Grealey, J., Inouye, M., *Green Algorithms: Quantifying the Carbon Footprint of Computation* Adv. Sci. 2021, 2100707. https://doi.org/10.1002/advs.202100707, https://www.green-algorithms.org, [accessed 2023-06-08]

[4] Data Center Manageability Interface, https://www.intel.com/content/dam/www/public/us/en/documents/technical-specifications/dcmi-v1-5-rev-spec.pdf, [accessed 2023-06-08]

[5] ipmitool, https://github.com/ipmitool/ipmitool, [accessed 2023-08-25]

[6] https://www.dell.com/community/FluidFS/MD1200-Power-Requirements-Questions/td-p/3754493, [accessed 2023-06-08]

[7] HEPiX, https://w3.hepix.org/benchmarking/HS06.html, [accessed 2023-08-23]

[8] *Google Data Centre PUE performance* https://www.google.co.uk/about/datacenters/efficiency/, [accessed 2023-08-23]

[9] https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-r6525.pdf, [accessed 2023-08-23]

[10] Egencia: Bringing sustainability into the functionality of booking corporate travel, https://www.egencia.com/en/sustainability-in-booking-tool, [accessed 2023-08-25]