ATL-SOFT-PROC-2025-001
06 January 2025

Carbon, Power, and Sustainability in ATLAS Computing

Emma Sian Kuwertz[1], *Zachary* Marshall[2,*], *Daniel* Schien[1], *Paul* Shabajee[1], and *Rodney* Walker[3]

[1]University of Bristol
[2]Lawrence Berkeley National Laboratory
[3]Ludwig Maximilians Universität

Abstract. The ATLAS Collaboration operates a large, distributed computing infrastructure: almost 1M cores of computing and over 1 EB of data are distributed over about 100 computing sites worldwide. These resources contribute significantly to the total carbon footprint of the experiment, and they are expected to grow by a large factor as a part of the experimental upgrades for the HL-LHC at the end of the decade. This contribution describes various efforts to understand, monitor, and reduce the carbon footprint of the distributed computing of the experiment. This includes efforts to construct a full life-cycle assessment model for the carbon impact of ATLAS distributed computing, all with the goal of making recommendations for sites to reduce their carbon footprint for the HL-LHC.

1 Introduction

With global temperatures rising and the undeniable impact of climate change affecting populations around the world, the environmental sustainability of large-scale science must be carefully examined. Advances, scientific and otherwise, always come with a cost, but it is the responsibility of the community to minimize the cost to society at large. With this in mind, the ATLAS collaboration has turned some of its attention to improving efficiency, reducing waste, and recommending a variety of actions to help reduce the environmental impact of the experiment. The ATLAS experiment [1, 2] is one of two general-purpose experiments operating at the Large Hadron Collider (LHC) at CERN. ATLAS operates a significant worldwide distributed computing system [3] that is expected to grow by a large factor in the coming decade [4].

CERN and the HEP community have been giving increasing attention to issues around environmental sustainability [5, 6], including the sustainability of computing. The focus of most studies to date, particularly within the HEP community, has been on power consumption [7]. This is an important aspect, but particularly with the reduction in carbon-intensity of the power grid, the carbon footprint of component production (“embedded” or “embodied” carbon) is of growing importance [8, 9]. Recommendations have to take into account the entire life-cycle of hardware, not just performance under a benchmark, in order to ensure that the realized environmental impact of a change will be positive.

---

[*]e-mail: ZLMarshall@lbl.gov

There are at least two fronts on which the problem of sustainability in computing can be approached. The first is for the experiment to take action by modifying policies and practices, optimize software, and improve awareness to reduce the usage and waste of computing resources. The second is largely for the sites and hardware operators, involving sustainability-minded procurement and improvements to data center efficiency via either configuration or physical site upgrades, for example. These two fronts are not independent; the most power-efficient hardware configuration depends on the software workloads being run. The greatest reduction in environmental impact can only be achieved by ensuring that all parties are working hand-in-hand.

Section 2 introduces the scale and variation in the computing resources used by the ATLAS experiment. Improvements to user behavior and the reduction of waste is discussed in Section 3, and changes to computing policy that can improve environmental impact are discussed in Section 4. Opportunities for sites and hardware improvements are discussed briefly in Section 5, and conclusions and plans are provided in Section 6.

2 Growing resources

The ATLAS Collaboration currently operates about 700,000 cores of computing, with peaks up to over one million cores thanks primarily to the contributions of High-Performance Computing (HPC) systems, as well as other "beyond-pledge",[1] opportunistic, and volunteer computing resources. These computing resources are paired with over one exabyte of storage worldwide, including about 450 PB of active disk storage and 600 PB of archival tape storage. The resources are spread over more than 100 sites world, including WLCG sites, HPC systems, commercial cloud computing, the ATLAS trigger computing system, and volunteer computing (ATLAS@Home).

In the coming years, the experiment will undergo a significant upgrade, with the new "high-luminosity" LHC (HL-LHC) running period beginning in 2030. As a result of the accelerator and detector upgrades for the HL-LHC, the computing resources required by the experiment are expected to increase to 3–5 times those of today. The experiment is expected to run until 2041, and by the end of operation the required computing resources are expected to be 10 times those of today. Limiting this growth of resources to only what is necessary to deliver the physics program of the experiment is an ongoing challenge. The growth also presents an opportunity: most computing sites will be buying significant new resources in the coming years, and if good recommendations can be made towards sustainable purchasing, the environmental impact of the upgrade and future operation of the experiment can be significantly reduced.

The current computing resources are diverse in terms of hardware, configuration, and management approaches: in some cases, the sites are completely dedicated to the ATLAS experiment and the collaboration has almost full control over their operation; in others, the collaboration is only one user of a system designed for other clients or use-cases. This diversity is only expected to increase in the coming years as the experiment strives to take advantage of more opportunistic resources and the computing hardware market continues to diversify. The diversity often results in limitations in terms of what actions can be taken towards sustainability. The most effective actions, therefore, are policy-driven actions that can immediately and positively affect all sites worldwide.

---

[1] In this context, "beyond-pledge" resources can be thought of as resources in a computing center that were not promised to the experiment but were delivered. These can be substantial in case a computing center serves multiple organizations and one is unable to utilize all their requested CPU, for example, or if the data center has old, less reliable hardware that they wish to use but not rely on.

3 Improving user actions and reducing waste

One of the most important and straightforward actions for the experiment to take is to raise awareness of environmental sustainability issues among users, production managers, and site administrators. To that end, all user jobs and production tasks now report an effective carbon footprint, with documentation of the footprint and real-world comparisons for reference [10]. The footprint that is reported is averaged over the world-wide site data for several reasons:

- Individual site data regarding power usage efficiency[2], renewable power sources (e.g. the availability of solar panels at a site), and specific actions taken by the site to reduce their carbon footprint (e.g. changing cooling set-points, exhaust heat re-use, recycling) are not sufficiently robust to be certain that a site's reported footprint is accurate.
- ATLAS computing resources are always busy; if a user re-locates their analysis job, a production job will take its place. Therefore there is no impact on the worldwide carbon footprint of the experiment if a user relocates their job to run in Norway instead of the USA, for example.
- Users aggressively re-locating their jobs to single sites could create significant congestion, and the resulting movement of data could negatively impact the total carbon footprint of the experiment.
- Using a global average ensures a direct link between software performance improvements and a reduced carbon footprint. A faster job run at the same time will report a smaller footprint. This would not be the case if a site-specific footprint were used and the two jobs ran at different sites.

Nevertheless, site-specific information is tracked and recorded. The job monitoring database includes information about the carbon intensity of the local power grid at each site during the job's run time.

Reducing waste is the most important first action to take towards environmental sustainability. There are many different sources of waste, and significant effort is required to systematically improve each one. Many examples of waste are down to human behaviors: users retrying jobs that will never succeed, not testing software sufficiently before running at large scale, being too conservative about data requirements and therefore creating unnecessarily large secondary datasets, or being too conservative about deleting unnecessary data or delaying the creation of data until it is required. In this last category, for example, some analysis formats are created for far more datasets than strictly required, simply because shortening the list of samples takes time, and there is some risk that more production could be needed in the future and induce delays. Educating users about the environmental impact of their choices is one way in which this waste can be reduced — and, of course, by reducing the waste, resources are freed for running critical and necessary production. In order to ensure that newcomers consider environmental sustainability, the introductory analysis software tutorial also includes a short section explaining the carbon footprint calculation for jobs, and pointing out that software performance and efficiency are the responsibility of every user. Some users and software developers in the experiment have renewed efforts towards software optimization and improving job efficiency because of a recognition of the environmental impact of their software.

---

[2] Power usage efficiency, or PUE, is defined as the ratio of power consumed by the data center to power consumed by the IT elements in the data center. A number close to one normally means almost all power in the data center is going to IT elements (e.g. very little power is going to heating or lights; all the power is going to CPU, disk, network, and other computing resources). When waste heat is considered, this number can be below one.

ATLAS also collected responses to a survey in 2023 that asked sites about actions towards sustainability. Most sites expressed a desire to improve their environmental impact, and several indicated that they were already taking some actions, mostly including improvements to the site efficiency and hardware configuration. One outstanding issue was that sites expressed significant uncertainty around the correct action to take with regard to retiring old hardware. Some run hardware as long as possible in order to amortize the embodied carbon cost of the hardware over a longer period; some renew hardware as quickly as possible in order to improve their computing power per Watt. Data-driven recommendations for old hardware in particular is one of the key recommendations that should be made prior to the HL-LHC period.

4 Computing policy changes

ATLAS has many computing policies that have direct environmental impact; however, until recently these were mostly examined from a financial standpoint. In some cases, these align very well: software or computing system optimization that results in reduced computing resource requirements provides a financial and environmental benefit. In some cases, the environmental impact is unclear and needs to be examined with some care before putting policies into place.

One positive example is the tape carousel system that ATLAS uses to actively and frequently recall data from tape and thereby reduce the need for spinning disk. The tape carousel has been in place for several years and has been effective in reducing the need for disk. Tape requires less power to operate per TB of storage than disk, thanks to the large volume of tapes that are idle in the library compared to the number of drives. Improving the tape carousel further, therefore, is an effective way to reduce the environmental impact of the experiment.

One policy the experiment is currently examining is the competition between data preservation and data reproduction. Old data must be regularly deleted to make space for new data and new analyses. In some cases, an analysis might be close to publication and might request that their data be preserved “just in case” it is required during collaboration or journal review. The solutions are requiring the analysis team to update to use newer data (often labor intensive, and therefore strongly disfavored); preserving the data on disk; preserving the data on tape (because these datasets are planned for deletion, and tape re-packing is a complex undertaking, this is also disfavored); and re-producing the data on demand in the case that the analysis team requires it. Between the two favored options — disk storage and reproduction — there are essentially two considerations: how quickly can the data be reproduced if the demand is made, and what is the relative environmental impact of storage compared to reproduction. The reproduction has already been shown to be fast enough for users. Rough calculations suggest that if the probability of recall is significantly below 10% then reproduction is likely to be the more environmentally friendly approach, but these calculations must be improved in order to ensure that the correct policy path is taken.

Some automatic approaches to waste reduction have also been taken. For example, the experiment uses HammerCloud for automatically identifying sites with problems and taking them offline, avoiding large numbers of failing jobs. In the case that the CPU at the site is simply left idle, the power savings is about 50%; in the case that the CPU can be re-purposed for other jobs that are not failing (e.g. some local batch system jobs, or jobs from another VO that are not affected by the same problems), 100% of the wasted power can be saved. The production system also uses a system of “scouting jobs”, where 10 jobs from a large task are run first as a test and must mostly succeed before the rest of the jobs are released. In that way, if the task is badly configured, only a small number of jobs will fail and a relatively small amount of CPU will be wasted. Similarly, within the distributed production system there are automatic triggers for retrying jobs that fail. These triggers are designed to retry jobs that have a transient failure, as well as to relocate jobs that appear to have a site-specific failure, in order to reduce the load on production managers. Of course, if the policies for retrying jobs are incorrect, a job may be retried many times before being identified as truly lost, resulting in significant wasted CPU. These policies need to be regularly re-examined to ensure that they are up to date.

5 Hardware opportunities

There are many trends in hardware and the worldwide power grid that present opportunities for more environmentally sustainable computing. Among these are the rise in renewable energy and waste heat re-use technology. The increasing volume of renewable energy in some countries has already resulted in periods where power is free and emission-free when considering the cost of operational production [11].[3] At the same time, however, there are expected to be periods when the power demand is greater than the renewable capacity. Decreasing the computing load during these periods helps reduce the environmental impact of the experiment.

One way to reduce load is to reduce the frequency of the CPUs run at the data centers. In fact, for typical ATLAS workloads that are often memory-bandwidth bound, reducing the frequency of the CPU results in greater throughput per Watt as well as a reduced operational carbon footprint. A second and more drastic approach is to provide check-pointing, whereby software can be stopped in the middle of a run, the state written to disk, and restarted at some later time. Check-pointing allows CPUs to be completely switched off, which can quite significantly reduce a site's power load. It has an additional use-case in areas where brownouts are frequent: by check-pointing all jobs immediately prior to a brownout (either when it is scheduled or by making use of the UPC of the data center), no work is lost when the data center has to be briefly shut down.

There are many other site considerations that can offer significant improvements to environmental sustainability. For large data centers, ensuring that the building has an extremely low PUE and that waste heat is used effectively provides significant benefits [12]. In many cases, constructing a new data center pays off in carbon footprint terms in under 10 years. Many of these issues are being actively explored (see, e.g., Ref. [13]), and most of them are not specific to the ATLAS experiment or LHC computing. Moreover, in many cases ATLAS does not have control over the configuration of a site. However, one action that can be taken is to ensure that when site measures are tested, HEP software is included in the suite of benchmarks that are run. For example, in the case of frequency scaling described above, including some memory-bound HEP software for benchmarking and not only CPU-bound applications like LINPACK can help ensure that sites draw the right conclusions for the real applications that will be run there.

6 Summary and planned efforts

There is significant ongoing work in the ATLAS collaboration towards understanding and improving the environmental impact of its computing. These efforts include storage footprint and configuration optimization, frequency scaling and software check-pointing, understanding the cost of networking components, the optimization of cooling systems, site power consumption, the adoption of new, more efficient computing platforms, and software optimization and efficiency improvements. One issue that these studies have made clear is that

---

[3] For example, hydroelectric power does not create emissions, but the creation of the power plant may have.

limited models can result in counter-productive recommendations. For example, if one ignores the embodied carbon of hardware, one should replace hardware continuously — which would clearly be harmful.

There are many important open questions that should be understood before the significant resource increases that will precede the HL-LHC operation period. One key question is the trade-off between, or relative cost of, various resources (CPU, disk, and tape), where choices in the computing model can exchange one resource for another. Another important question is towards the adoption of new hardware, like GPU accelerators. For a single application that is able to efficiently run on a GPU, it is clear that there is an operational benefit and reduction in operational carbon compared to the same application running on a CPU. However, for ATLAS, only some fraction of workloads will be able to make use of GPU acceleration. Therefore, the improvement when the GPUs are in use must be balanced against their embodied cost, as well as the operational costs when they are idle.

Extrapolations into the future can affect recommendations significantly. The decarbonization of the worldwide power grid will increase the importance of embodied carbon compared to operational carbon. Efforts towards more modular, reusable hardware components could reduce waste in some parts of a site. At the same time, trends towards larger packages that are internally more efficient may result in more difficulty in reusing or recycling components.

The significant increase in computing resource requirements that the ATLAS experiment will face at the start of the HL-LHC era offers an opportunity for the re-optimization of the environmental impact of computing. Recommendations to the sites of best practices, as well as a careful re-examination of computing policies with environmental considerations in mind, can result in a significant reduction in waste and the overall carbon footprint. These studies also provide an important window into the optimization of computing systems for future large-scale experiments in the years to come.

References

[1] ATLAS Collaboration, JINST **3**, S08003 (2008)

[2] ATLAS Collaboration, JINST **19**, P05063 (2024), `2305.16623`

[3] ATLAS Collaboration, *Software and computing for Run 3 of the ATLAS experiment at the LHC*, CERN-EP-2024-100 (2024), `2404.06335`

[4] ATLAS Collaboration, *ATLAS Software and Computing HL-LHC Roadmap*, CERN-LHCC-2022-005 (2022), `https://cds.cern.ch/record/2802918`

[5] CERN Occupational Health and Safety and Environmental Protection unit and the CERN Education, Communications and Outreach group, *CERN ENVIRONMENT REPORT 2019-2020* (2021), `https://hse.cern/environment-report-2019-2020`

[6] CERN Occupational Health and Safety and Environmental Protection unit and the CERN Education, Communications and Outreach group, *CERN ENVIRONMENT REPORT 2021-2022* (2023), `https://hse.web.cern.ch/environment-report-2021-2022/energy`

[7] D. Britton, S. Campana, B. Panzer-Stradel, EPJ Web Conf. **295**, 04001 (2024)

[8] U. Gupta et al. (2020), `2011.02839`

[9] Boavizta Working Group, R. Lorenzini, *Digital and Environment: How to evaluate server manufacturing footprint, beyond greenhouse gas emissions?* (2021), last accessed 18 August 2024, `https://www.boavizta.org/en/blog/empreinte-de-la-fabrication-d-un-serveur`

[10] PanDA Collaboration, *Estimation of PanDA job carbon footprint*, last accessed 18 August 2024, `https://panda-wms.readthedocs.io/en/latest/advanced/carbon_footprint.html`

[11] Federal Ministry for Economic Affairs and Climate Action, *Renewable Energy*, last accessed 18 August 2024, `https://www.bmwk.de/Redaktion/DE/Dossier/erneuerbare-energien.html`

[12] V. Ljungdahl, M. Jradi, C. Veje, Applied Thermal Engineering **201**, 117671 (2022)

[13] EE HPC WG and Center of Expertise for Energy Efficiency in Data Centers, *Center of expertise for energy efficiency in data centers: Data center master list of energy efficiency actions*, last accessed 18 August 2024, `https://datacenters.lbl.gov/resources/data-center-master-list-energy`