Data Centers: The Engineering Marvels Powering Our Modern Economy

Ariel Gonzalez

Florida International University

CGS 3095 Section RVC – Fall 2026

Abstract

This research summarizes current and projected estimations of energy consumption and carbon emissions resulting from global data center expansion. It uses a simple heuristic to estimate what the current energy consumption and carbon emissions would have been if equivalent computational workloads ran in on-premises data centers. It also summarizes some of the most important business and engineering efforts the world's leading cloud service provider has done as part of its commitment to sustainable data center operations. It also presents simple illustrative scenarios showing how software architects can help design applications for lower power consumption. It briefly analyzes the contribution of data centers from a utilitarian, deontological, and bioethical and human rights perspective. Lastly, it provides a brief global analysis.

Data Centers: The Engineering Marvels Powering Our Modern Economy

1. INTRODUCTION

Data centers, the engineering marvels powering our modern economy, are the evil villains and arch-nemeses of ecology, and many of them alone can surpass the energy consumption of a small nation. This is how the media has portrayed data centers for years. It has also been the way academics and sustainability advocates have presented data centers in lectures at some of the most prestigious universities—lectures I have attended and can describe in first person. Their energy consumption is undeniably gargantuan. The International Energy Agency (IEA) reports that data centers consumed around 415 terawatt-hours (TWh), or about 1.5% of global electricity consumption, in 2024. This demand soared by 17% in 2025 and, more alarmingly, is expected to reach around 945 TWh by 2030, representing just under 3% of total global electricity consumption [1, 2]. Rapid industry expansion has also strained water resources. Mostly used for cooling, a 1-megawatt (MW) data center using traditional cooling methods would use about 6.75 million gallons of water per year [3].

Have the 13,030 data centers across the globe [4] been built without any concern for the environment? Have cloud service providers made any effort to optimize their infrastructure and preserve natural resources? Have they attempted to develop more energy-efficient cloud technologies to reduce energy consumption? Are there any software or architectural patterns that architects can apply to build more energy-efficient applications? While all of these questions are critically important, the latter carries a much deeper personal overtone. Having spent nearly twenty years architecting web applications, the author had never considered architecting web applications for lower energy consumption. To answer these questions, the author turned to the leading and most widely respected cloud service provider, Amazon Web Services (AWS).

2. REDUCING ENVIRONMENTAL IMPACT

The greatest concern about the massive increase in data center energy consumption is carbon emissions. While nuclear and renewable energy sources produce relatively few direct or indirect emissions, coal remains the largest source of electricity, accounting for 34 percent of total global generation [5]. With estimated emissions of 2.31 pounds per kilowatt-hour (lb/kWh) [6], the remaining 321.3 TWh of energy that would need to be generated from coal in 2030 to meet global data center energy demand would already equate to approximately 336.7 million metric tons of CO_2e ($MMTCO_2e$) annually. Continuing to increase data center capacity globally without investing in renewable energy projects would further exacerbate the problem.

In 2016, Amazon enabled the Eastern Shore, the first grid-scale solar project in Virginia. A year later, the company enabled the Virginia – Southampton solar project. In both cases, Amazon signed a long-term Power Purchase Agreement. The projects generate a combined total of 381,000 megawatt-hours (MWh) of electricity per year [7, 8]. Since then, the company has invested in over 700 carbon-free energy projects in 28 countries, with capacity to generate over 40 gigawatts (GW) of electricity. In 2025, BloombergNEF again recognized Amazon as one of the world's leading corporate purchasers of carbon-free energy, and noted that it has the largest carbon-free energy portfolio of any corporation globally [9]. Furthermore, Amazon co-founded and was the first signatory of The Climate Pledge, a commitment to reach net-zero carbon emissions across its operations by 2040 [10], which makes the company not just the leading cloud service provider, but also a champion of sustainability.

3. COMPUTER ENGINEERING TO REDUCE ENVIRONMENTAL IMPACT

However, while the dates above might erroneously give the impression that the company only began taking action a decade ago, plenty of empirical evidence suggests otherwise.

3.1 Engineering More Efficient Central Processing Units

Amazon has shown visionary leadership and a true commitment to engineering excellence in optimizing its infrastructure. In 2015, when growth indicators for cloud service demand were not as strong as they are today, the company assumed the business and financial risk of developing its own processor. The initial AWS Graviton processor featured a modest 16 cores and launched in 2018 alongside the A1, the first generation of instances powered by an ARM processor. In retrospect, perhaps only AWS and a few individual industry experts had truly understood the significance of what the company had done.

The AWS Graviton processor was engineered to consume significantly less memory and deliver significantly more processing power. The AWS Graviton5, released in 2025 to power the 9th Generation of AWS instances, features 192 cores, consumes less energy per core than any predecessor, and delivers more processing power. As the company recently reported, AWS Graviton-based Amazon EC2 instances use up to 60% less energy than comparable Elastic Compute Cloud (EC2) instances for the same performance [11]. AWS was the first cloud vendor to develop its own processor, a true business and engineering triumph that, as of 2024, saves an estimated 83,000 metric tons of CO_2e each year [12].

3.2 Engineering More Efficient Virtualization Platforms

Although the introduction of the AWS Graviton processor was significant, the company had already made substantial changes to increase performance efficiency. Several years earlier, in 2013, Amazon began designing custom hardware to improve networking. In 2017, the company introduced the AWS Nitro System to eliminate the overhead of the software-based Xen hypervisor. Nearly a decade later, the much more advanced AWS Nitro System encompasses

multiple cards, a security chip, and a hypervisor that runs with near-zero overhead [13]. Except for a few legacy instances still supported, all other instances today run on the Nitro System [14].

3.3 Core Principles

The media, shrouded in over-sensationalism, does not report these technological advancements, and apparently scholars also are not keeping up with the research. Perhaps they have also embraced the same over-sensationalism, reporting extravagant energy-consumption figures in presentations instead of crediting AWS's incredible engineering achievements and its unshakable commitment to sustainability.

It is imperative to pause to reflect on the questions that brought the research to this junction. Yes, responsible cloud computing is possible. AWS has proven that it is. Perhaps it is the result of business and engineering teams with superhuman abilities, or perhaps it is nothing at all. AWS has consistently shown great customer obsession, ownership, curiosity, high standards, the ability to think big, and the ability to build trust—among themselves as an organization and with their customers. After all, according to John Rossman, "unlike most organizations, Amazon's leadership principles are not simply suggested guidelines for new hires or empty verbiage from a mission statement buried in the employee manual" [15]. Perhaps it is the genuine embrace of those core leadership principles that makes everything else possible.

4. SOFTWARE ENGINEERING TO REDUCE ENVIRONMENTAL IMPACT

However, one question remains unanswered. Are there any software or architectural patterns that architects can apply to build more energy-efficient applications? Concrete evidence in the AWS literature indicates that it is possible, but doing so most effectively requires a balance between software patterns and solution architecture.

4.1 Load Balancing

Intuitively, we can argue that one of the most crucial ways to improve resource utilization and reduce energy consumption when hosting web applications in the cloud is to use load balancing. Traditional, high-volume load balancing in on-premises data centers was generally done using hardware appliances (e.g., Cisco GSS, Cisco CSS). However, the high-performance switching appliances could not dynamically adjust the cluster size based on incoming application traffic, leaving clusters heavily overprovisioned during off-peak hours. Cloud-based load balancing offers much greater flexibility. A web application hosted in AWS can use an application load balancer, target group, and auto scaling group to create a highly effective load-balancing strategy. Auto scaling groups can provide the missing functionality mentioned above by dynamically adjusting the cluster size based on incoming application traffic. Furthermore, by combining session affinity and software patterns to implement a second-level cache within each cluster instance, software architects can achieve even greater resource utilization and lower energy consumption. The cache eliminates the need to constantly read reference datasets over the network, drastically improving application performance and eliminating unnecessary operations across different stacks [16].

4.2 Properly Sizing Application Nodes

Properly sizing application instances is crucial to avoid overprovisioning resources. AWS offers more than 1,160 instance types across 6 major instance families (e.g., General Purpose, Compute Optimized, Memory Optimized), so that customers can choose the right instance for their specific workloads. Solution architects, ultimately responsible for selecting instances based on customer requirements, can holistically consider instance specifications, patterns provided by software architects, and architectural design to reduce energy consumption [17].

4.3 Job Scheduling and Batch Processing

Historically, scheduling jobs in on-premises data centers required an application or scheduling server (e.g., Quartz). While many offered strong functionality and scalability, these servers had to run continuously, even when no jobs were running, unnecessarily consuming energy. AWS EventBridge is a serverless event bus with great scheduling capabilities, including support for CRON syntax. The EventBridge scheduler can trigger a Lambda function, which can then use the Elastic Compute Cloud (EC2) API to start an instance or group of instances to run a particular job. The instances can perform the job(s) and automatically shut down after completion, optimizing resource utilization and, consequently, energy consumption [16].

5. ETHICAL ANALYSIS

Based on this research, we can analyze data centers through three philosophical frameworks: Utilitarianism, Deontology, and Bioethics and Human Rights.

Data centers have enabled sustainable transformations across a broad spectrum of industries by enabling dematerialization, powering clean energy innovation, accelerating circular supply chains, and fostering sustainable practices [18]. In addition, they have enabled the development of large platform ecosystems and software-intensive systems [19, 21] while using power more efficiently and producing lower carbon emissions than on-premises data centers [20]. Consequently, from a utilitarian perspective, data centers used by any industry result in a net ethical good.

From a deontological perspective, when they are used to solve complex issues like global supply chain optimization, agricultural research, or extreme weather forecasting, they become a moral imperative, as we have an absolute duty to build the infrastructure needed to solve such difficulties.

However, from a bioethics and human rights perspective, their importance is even greater. For example, AWS has enabled some of the largest pharmaceutical platform ecosystems in existence today, including Pfizer, Merck, Moderna, AstraZeneca, and Novo Nordisk. Medical advancements and the resulting benefits to humanity from advanced genetic treatments, accelerated drug discovery, pipelines that speed clinical trials, precision medicine patient matching, and automated pharmaceutical manufacturing are a humanitarian necessity.

Their contributions to humanity far surpass the limited resources they utilize.

6. GLOBAL ANALYSIS

Given the extensive empirical evidence in the literature identified through this research, we conclude that data centers have significantly improved the utilization of computational resources and, consequently, helped preserve natural resources. Running workloads on data centers can be up to 4.1 times more energy efficient and can reduce carbon footprint by up to 99 percent [20]. This research suggests that progressively consolidating computational resources in data centers has enabled scientific discoveries that would not have been possible with isolated on-premises data centers. However, even if it were possible—disregarding the duplicated infrastructure waste of running small on-premises data centers worldwide—it would cost the world another 1,286.5 TWh of energy annually and require coping with a staggering additional 1,347,992,610.18 MMTCO2e.

7. CONCLUSION

Data centers, the engineering marvels powering our modern economy, are not the evil villains the media and many credulous scholars believe they are. The extensive research supporting this paper demonstrates unequivocally that data centers are the only embodiment of the digital revolution that has contributed to the rapid development of so many other industries by democratizing advanced technologies, allowing businesses to gain a global presence in minutes, making serverless architectures ubiquitous, enabling faster experimentation, and, more importantly, considering mechanical sympathy [21, 22, 23].

Thus, let it be known to the industry and to the world: responsible cloud computing is possible. Amazon has proven that it is.

REFERENCES

[1] International Energy Agency. (2026, April 16). *Data centre electricity use surged in 2025, even with tightening bottlenecks driving a scramble for solutions*. https://www.iea.org/news/data-centre-electricity-use-surged-in-2025-even-with-tightening-bottlenecks-driving-a-scramble-for-solutions

[2] International Energy Agency. (2025). *Energy demand from AI*. In Energy and AI. https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

[3] Orr, R., & Klesner, K. (2016, June 17). *Ignore data center water consumption at your own peril*. Uptime Institute Journal. https://journal.uptimeinstitute.com/dont-ignore-water-consumption

[4] Programs.com. (2026, October 30). *Measuring the data center boom: Facts and statistics (2026). https://programs.com/resources/data-center-statistics*

[5] International Energy Agency. (2026, April 20). *Global Energy Review 2026: Electricity supply*. https://www.iea.org/reports/global-energy-review-2026/electricity-supply

[6] U.S. Energy Information Administration. (2024, December 11). *How much carbon dioxide is produced per kilowatthour of U.S. electricity generation?* https://www.eia.gov/tools/faqs/faq.php?id=74&t=11

[7] Community Energy. (2026). *Amazon Solar Farm – Eastern Shore*. https://www.communityenergyinc.com/projects/amazon-solar-farm-us-east

[8] Community Energy. (2026). *Amazon Solar Farm Virginia – Southampton*. https://www.communityenergyinc.com/projects/amazon-solar-farm-virginia-southampton

[9] Amazon.com. (2026). *Amazon sustainability*. https://sustainability.aboutamazon.com/

[10] Amazon Web Services. (2026). *AWS sustainability*. https://aws.amazon.com/sustainability/

[11] Amazon Web Services. (2026). *Sustainability with AWS Graviton*. Retrieved November 3, 2026, from https://aws.amazon.com/ec2/graviton/graviton-sustainability/

[12] Amazon Web Services. (2025, July). *2024 Amazon sustainability report: AWS summary*. Amazon. https://sustainability.aboutamazon.com/2024-amazon-sustainability-report-aws-summary.pdf

[13] Amazon Web Services. (2026). *AWS Nitro System*. https://aws.amazon.com/ec2/nitro/

[14] Amazon Web Services. (2026). *Instance Type Explorer*. https://aws.amazon.com/ec2/instance-explorer/

[15] Rossman, J. (2024). *The Amazon way: Amazon's 14 leadership principles (3rd ed.)*. Clyde Hill Publishing.

[16] Shrivastava, S., Srivastav, N., Artasanchez, A., & Sayed, I. (2023). *AWS for Solutions Architects: The definitive guide to AWS Solutions Architecture for migrating to, building, scaling, and succeeding in the cloud (2nd ed.)*. Packt Publishing.

[17] Sard, P., & Wadia, Y. (2024). *AWS Certified Solutions Architect – Professional Exam Guide (SAP-C02): Gain the practical skills, knowledge, and confidence to ace the AWS (SAP-C02) exam on your first attempt*. Packt Publishing.

[18] Patel, P. G., Dua, I. K., & David, S. (2025). *Sustainable cloud development: Optimize cloud workloads for environmental impact in the GenAI era*. Packt Publishing.

[19] Tiwana, A. (2013). *Platform ecosystems: Aligning architecture, governance, and strategy*. Morgan Kaufmann.

[20] Amazon. (2026). *AWS Cloud. Amazon Sustainability*. https://sustainability.aboutamazon.com/products-services/aws-cloud#increasing-efficiency

[21] INCOSE. (2023). *INCOSE systems engineering handbook: A guide for system life cycle processes and activities (5th ed.)*. Wiley.

[22] Amazon Web Services. (2026). *AWS Well-Architected*. https://aws.amazon.com/architecture/well-architected/

[23] Amazon Web Services. (2026). *AWS Well-Architected Tool*. https://aws.amazon.com/well-architected-tool/