The Frontiers of Society, Science and Technology
ISSN 2616-7433 Vol. 6, Issue 11: 58-62, DOI: 10.25236/FSST.2024.061110

Small Modular Reactors: A Viable Path to Sustainable Data Centers in the Age of Artificial Intelligence

Junchen Lu

BASIS International School Hangzhou, Hangzhou, Zhejiang, 310000, China

Abstract: *Advancements in digital technologies rely on the development of Data Centers (DCs). Yet, the rapidly increasing energy demand for Artificial Intelligence (AI) poses an unprecedented challenge to sustainability. Along with the expansion of the DC industry, this sector's reliance on electricity, substantially from unrenewable sources, exacerbates carbon emissions. This paper discusses the historical trends and current trajectory of energy consumption by DCs in the age of AI and then explores the merits and downsides of adopting Small Modular Reactors (SMRs) as a potential solution to meet the energy demand of future computing facilities. In the end, we assess the future prospects of SMRs to ensure the energy sustainability of AI and its future developments.*

Keywords: *Data Centers; Artificial Intelligence (AI); Energy Consumption; Sustainable Developments; Small Modular Reactors (SMRs)*

1. Introduction

The pace at which Artificial Intelligence (AI) advances is at an unprecedented level. Ever since Large Language Models were introduced, AI has been one of the most prominent sectors in the world. At the same time, centralized computing facilities, namely Data Centers (DCs), have been proliferating, to meet the growing scale of AI models and other applications, including cloud computing. As the energy demand for DCs grows continuously, it is essential to consider its sustainability.

The scale of this increasing energy demand is evident: DCs consumed approximately 460 TWh (Tera-Watt-hours) of energy in 2022, corresponding to about two percent of the world's overall energy use [1]. This is equivalent to the annual energy consumption of 44 million US households (assuming an average US household consumes 10 MWh annually [2]). Following an exponential growth trend, as shown in Figure 1, energy consumption is projected to double by 2030, further putting pressure on sustainability [3].

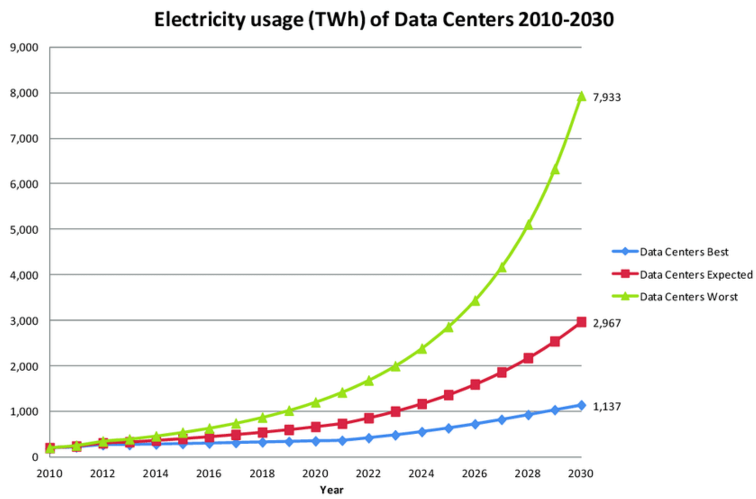

Figure 1: A projection of electricity usage of DCs until 2030 [4]

Why is the consumption growing so dramatically? It should be attributed to the characteristics of AI. For instance, in LLMs, neural networks require multiple times the number of computations, consuming

ISSN 2616-7433 Vol. 6, Issue 11: 58-62, DOI: 10.25236/FSST.2024.061110

several times the energy of traditional search engines to output a response. As an example, a single ChatGPT query requires 2.9 Wh of electricity, compared with 0.3 Wh for a Google search [5], so almost ten times more. If all AI-related applications required a ten-fold increase in energy use, then the information technology industry would require a significantly greater number of DCs, which would drive a highly significant increase in energy and power consumption, as estimated by some sources [5]. Furthermore, training AI models also consumes an increasing amount of energy. GPT-4, for example, required over 50 GWh of energy for its training, 50 times more than GPT-3 [6]. This rapidly increasing demand will have to be met by a corresponding increase in energy generation, well beyond the current output and capacity of energy infrastructure in the United States and, to a similar extent, in the rest of the world.

Various solutions have been introduced to address the issue. Emphasis has been placed on energy sources that are sustainable and with a reduced carbon footprint, so as to not exacerbate current global issues like climate change. Renewable energy sources like wind, solar, and hydroelectric power have been explored as a route to power DCs, owing to their very limited environmental impact. However, it seems improbable, although still a matter of debate, whether these renewable sources alone may be able to meet the current, and especially future, energy demand from DCs.

At the same time, there is a sustainable and scalable technology that has been available for many decades, which is nuclear energy. However, it shall be noted that nuclear energy has its drawbacks, first and foremost the time it takes to build a power plant, as well as the issue of disposal of nuclear waste. The vital question then becomes whether DCs can be powered by a new technology of small-scale reactors that could be built to be dedicated to DCs. One such technology, namely Small Modular Reactors (SMRs) has been recently discussed as a possible solution. Although still an early-stage technology, it is potentially a viable option, as it combines the benefits of nuclear power – such as high capacity, high energy density, and no carbon emission – with improved safety and flexibility, and, most importantly, reliable and continuous generation of power at the level required to operate single DCs.

2. Data centers – challenges and opportunities

Given the increasing importance of DCs for information technology applications, it is crucial for one to understand how DCs are structured, their energy consumption, and their operation costs, in order to be able to assess the current and projected demand for new technologies like AI, and their future sustainability.

DCs have complex compositions with diverse energy demands, though the primary sources of energy consumption, illustrated by Figure 2, can be traced back to two main areas [7]: IT equipment (e.g. server and storage, networking, etc.); and infrastructure facilities (e.g. cooling, power systems, etc.).

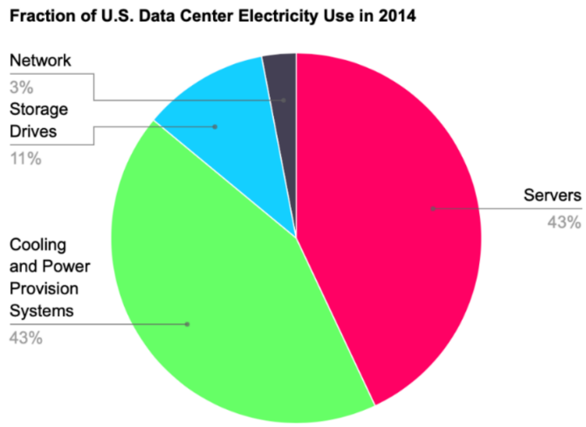

Figure 2: The pie chart above [8] shows a typical distribution of DC's power consumption sources, yet this may vary depending on the types of individual DCs.

ISSN 2616-7433 Vol. 6, Issue 11: 58-62, DOI: 10.25236/FSST.2024.061110

Servers are responsible for receiving, processing, and sending data to support network services, and are one of the major areas where energy is consumed in DCs. Central processing units (CPUs), and, increasingly, graphic processing units (GPUs) perform computations. Because AI is highly complex and very power-hungry, it relies on a large number of processing units to perform training, testing, and inference, resulting in a significant energy demand. The storage of data, performed mainly by memory, hard drives, and solid-state drives, also consumes significant amounts of power (~11% of the total energy budget of a DC), since these storage devices require power to maintain data integrity and perform read/write operations.

Networks, namely the interconnection between devices, usually consume a small but non-negligible amount of energy (~3% of the total energy budget of a DC) when data transmission takes place at large bandwidth, because of the great amounts of data that are transferred per second. Its routers, maintaining the transfer of data packets both within the DCs and toward external networks, also consume energy when handling large traffic. Altogether, the data processing hardware (computing, storage, transmission) utilizes more than half (~57%) of the power needed to operate a DC. A large-scale DC (~100 MW) can be estimated to consume around 900 GWh in a year, and the portion consumed by data processing hardware alone is about 500 GWh. Given an average annual electricity consumption by a single household of around 10 MWh/year, the power consumption for IT equipment at a single DC is equivalent to the power used by ~50000 homes.

While computing hardware represents a significant fraction of a DC energy budget, the consumption of non-computational hardware components is almost as prominent.

The cooling systems, which perhaps have always dominated the consumption share of DCs, are vital due to the strict requirements for temperature control at DCs. Highly dense processing units result in enormous heat dissipation, therefore cooling is essential to maintain the operating temperature within a safe range, less than 27°C typically. Higher temperatures may result in the reduction of computing efficiency, or damage to the microprocessors. Using air as a heat-conducting medium, the cooling system maintains its circular flows, removing heat from servers and other electronic equipment. Currently, cooling equipment consumes about 35% of the DC power budget. However, there have recently been increasing investments in liquid-based cooling for DCs. As some liquid cooling technologies have been proven to reduce energy consumption [9], it could be a future trend.

The power system is responsible for ensuring a continuous supply of power to every device in a DC. Electricity is distributed by Power Distribution Units, which contain transformers and voltage regulators. Power conversion, namely converting alternating to direct currents, is not highly efficient and thus results in non-negligible energy losses. Power systems account for ~10% of the energy consumption of a DC.

Although it is necessary to look at the energy budget of a DC when assessing its sustainability, looking at the financial costs is equally important to ensure the adoption of a viable sustainable energy supply. While the construction of DCs is expensive, the cost of energy consumption tends to exceed construction costs after a few years of operation [10]. Optimizing the latter is relatively more beneficial to the economy.

DCs receive energy from different sources, most commonly electrical grids and renewables. However, a shortcoming of electrical grids in the sense of sustainability is that, as shown in Figure 3, less than only ~20% of the electricity is generated from renewable/clean sources [11].

Besides grid electricity, there is an increasing number of DCs migrating to renewable energy [12]. Yet, as DCs increase in number and size, renewable energy sources like solar and wind power become rather impractical, due to their significant requirements in land and resources. Taking a mid-scale DC (35 MW) as a representative example, if all of its power were to come from renewable sources like solar, this would require 175000 square meters, or ~43 acres of land (assuming a typical solar power generation density of 200 W/m^2 [13]) to accommodate a properly sized solar farm. Wind power generation is even less viable because of the greater requirements for the land area since wind farms have a much lesser power density, which is 2-3 W/m2 [14]. Furthermore, the intermittency of solar energy makes it even more impractical. A 35 MW DC consumes roughly 800 MWh daily. Assuming 4 full sun-hours per day, and a 200 W/m^2 solar energy density production, one would need approximately a square kilometer, or roughly 250 acres. On top of that, solar-powered DCs have to store energy in batteries for night-time operation, when solar power is unavailable. A battery system with a capacity of 800 MWh/day would occupy a volume of roughly 40000 cubic feet (or about 13 average single-family homes) and would require 3200 tons of lithium (assuming that the DC uses lithium-ion batteries with a gravimetric energy density of 250 Wh/kg and a volumetric energy density of 700 Wh/L [15]). These additional energy storage requirements make the sole reliance on sun energy even more challenging.

ISSN 2616-7433 Vol. 6, Issue 11: 58-62, DOI: 10.25236/FSST.2024.061110

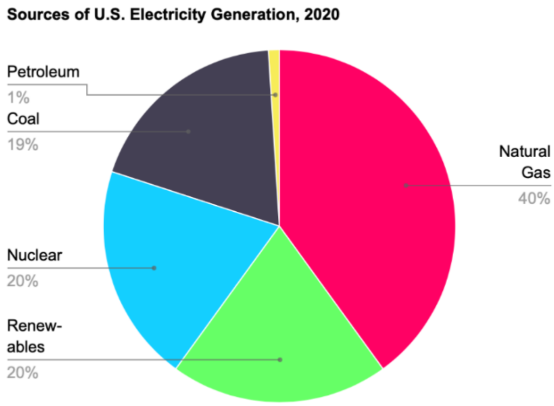

Figure 3: The pie chart above [11] shows the distribution of U.S. electricity generation sources in 2020.

3. Small modular reactors – a possible solution

Having reviewed the challenges for renewable energy supply to DCs, we can ask ourselves the following question: Are there other clean energy sources that can provide a practical solution for DCs? The natural candidate is nuclear power [16]. Though not renewable, nuclear power is carbon-free. Its high energy density implies that nuclear power makes a more efficient use of land resources. Unlike wind and solar power, nuclear power reactors operate day and night, providing an uninterrupted supply of energy for continuous DC operations.

However, a typical utility-scale nuclear power plant, producing around 1 gigawatt of power, is oversized for the needs of a single DC. Thus, small modular reactors (SMRs), which can generate 300 megawatts of power or less, will be more suitable for DCs. SMRs are a smaller and more flexible type of nuclear reactor. Their smaller size allows for greater flexibility in their deployment. They are also modular, meaning that they can be constructed and installed at two different locations.

The size of SMRs allows them to be built in the proximity of DCs, enabling a more stable and consistent supply of energy. This reduces significantly the reliance on backup energy sources due to power interruptions. The modular design of SMRs also makes them scalable, allowing DCs to increase their power capacity incrementally as demand increases. This flexibility enables modules to be transported in case of repairs. Moreover, the design of SMRs provides more safety for energy-intensive operations like DCs. Its passive safety systems significantly increase the emergency response time in case of accidents [17].

On the topic of sustainability, it should be noted that the nuclear waste of SMRs must be handled properly in order not to cause harm to the environment. Yet with more and more sophisticated disposal technologies, this appears to no longer be a major issue [18].

Cost is another factor of consideration in order to assess the viability of using SMRs for DCs. Median estimates for the construction cost of a 45-megawatt SMR range from $4000 to $16300 per kilowatt [19]. Assuming a $10000/kW cost per unit of power, the overall construction cost can be estimated at around 350 million dollars for a 35-megawatt DC. However, the operation cost of SMRs is a mere $100 per megawatt hour [20], which is only a very small fraction of the operating costs of an average DC. Also, with energy supply being a significant component (~60-70%) of the total operational costs of DCs [21], the adoption of more efficient power sources can have a substantial impact on the operating budget of a DC.

SMR projects are also receiving increasing support, accelerating their deployment [22]. Given the great opportunities presented by SMRs, it is all the more essential to weigh their risks in relation to their benefits. Although SMRs are generally less expensive than traditional nuclear power plants, the initial capital investments are still significant. Furthermore, as a new technology, it must undergo

ISSN 2616-7433 Vol. 6, Issue 11: 58-62, DOI: 10.25236/FSST.2024.061110

comprehensive regulatory review in order to ensure that it meets safety and environmental standards. Given the early stage of this new technology, more investigations are required before SMRs can be fully deployed as power sources for DCs.

4. Conclusion

So what are we to conclude? AI's large and rapidly growing requirements for energy are not only a concern to DCs' economy but also put pressure on their sustainable development. As integrations of numerous computing devices, DCs have multiple sources of energy consumption, including servers and cooling systems as the major share. Both are projected to increase along with the expansion of AI models in both number and complexity. Therefore, providing clean energy to data centers is vital to ensure their sustainability. SMRs emerge as a promising solution, offering not only zero carbon emissions but also a more reliable source of electricity. However, SMR technology is still in its nascent stage and represents just one possibility among many. Further exploration and consideration are necessary to determine the optimal path forward.

References

[1] "Electricity 2024 - Analysis and forecast to 2026", International Energy Agency, 2024.
[2] "Electricity use in homes," U.S. Energy Information Administration, May. 09, 2019.
[3] M. Gooding, "US data center power consumption to double by 2030," Data Center Dynamics, Jan. 15, 2024.
[4] L. M. Hilty and B. Aebischer, "On global electricity usage of communication technology: trends to 2030," ResearchGate, Apr. 2015.
[5] Goldman Sachs, "AI poised to drive 160% increase in power demand," Goldman Sachs Intelligence, Jun. 28, 2023.
[6] A. Cohen, "AI is pushing the world towards an energy crisis," Forbes, May. 23, 2024.
[7] EnergyConsult, "Energy efficiency policy options for AUS/NZ data centres," Energy Rating, Apr. 2014.
[8] E. Masanet and N. Lei, "How much energy do data centers really use?" Aspen Global Change Institute, Mar. 2020.
[9] R. Brown, "Energy performance improvement in data centers: A review of metrics, benchmarks, and drivers," Energy and Buildings, vol. 152, pp. 19-31, Oct. 2017.
[10] J. L. Lucas Simões Gomes, G. B. Brant, A. O. Fernandes, and P. R. Freire Cunha, "Modeling service execution on data centers for energy efficiency and quality of service monitoring," ResearchGate, May. 2014.
[11] U.S. Environmental Protection Agency, "U.S. electricity grid & markets," EPA Green Power Markets, 2020.
[12] IS Partners, "How data centers are using renewable energy," IS Partners LLC, Jul. 19, 2023.
[13] Blair, "On Renewables and compromises, Intermission: Energy Density and Power Density," A Chemist in Langley, 2015.
[14] D. Roy, P. Vaughan, and M. Yelland, "Physical Footprint comparison," Greens for nuclear energy, 2020.
[15] John, "The Energy Density of a Lithium-ion Battery," U Fine Battery, Jan. 27, 2024.
[16] Office of Nuclear Energy, "3 Reasons Why Nuclear is Clean and Sustainable," Energy.gov, Mar. 31, 2021.
[17] Directorate-General for Energy, "Small Modular Reactors explained," European Commission, 2024.
[18] World Nuclear Association, "Storage and Disposal of Radioactive Waste - World Nuclear Association," World Nuclear Association, Apr. 30, 2024.
[19] A. Abdulla, I. L. Azevedo, and M. G. Morgan, "Expert assessments of the cost of light water small modular reactors," Proceedings of the National Academy of Sciences, vol. 110, no. 24, pp. 9686–9691, Jun. 2013.
[20] D. Schlissel, "Eye-popping new cost estimates released for NuScale small modular reactor," Institute for Energy Economics and Financial Analysis, Jan. 11, 2023.
[21] T. R. Comerford, "Power Requirements, Energy Costs, and Incentives for Data Centers," BLS & Co., Nov. 1, 2015.
[22] P. Lookadoo, J. Bolinger, L. Dickinson, "DOE's New $900M Funding for Small Modular Reactor Project Teams", Haynes Boone, July. 17, 2024.