



2024

US DATA CENTER ENERGY USAGE REPORT SUMMARY

LAWRENCE BERKELEY
NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY



+1 463-266 4496



INFO@VEDENI.ENERGY



VEDENI.ENERGY.COM

UNDERSTANDING THE GROWTH OF DATA CENTER ENERGY USE

The 2024 United States Data Center Energy Usage Report explores U.S. data centers' rapidly growing energy consumption and the technological, operational, and environmental dynamics that drive it. The increasing digitization of the economy has fueled unprecedented demand for data centers, and while efficiency improvements initially held energy use steady, a sharp shift occurred with the rise of new hardware technologies. These include GPU-accelerated servers designed for artificial intelligence (AI) applications and the computational demands of cryptocurrency mining. The report outlines historical trends, current impacts, and future scenarios, providing a roadmap for understanding how this critical sector evolves.

The report contextualizes the data center industry's rise alongside expanding cloud computing and digital services. Efficiency improvements in the 2010s helped maintain relatively flat energy consumption despite increasing demand, primarily through innovations in cooling systems, improved server utilization, and reduced server idle power. However, the growing adoption of AI and the computational intensity of cryptocurrency mining have introduced new challenges, causing energy use to climb significantly.

METHODOLOGY

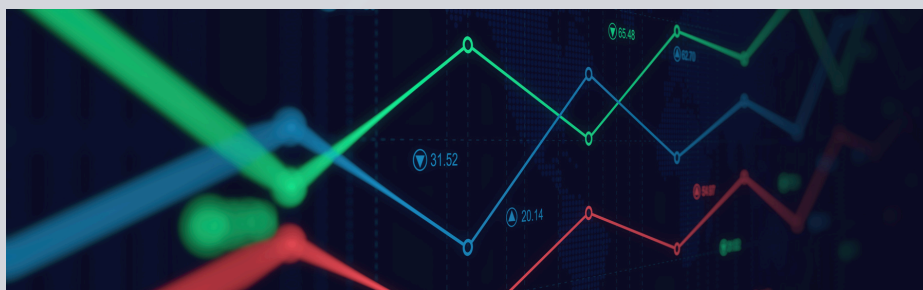
ROOTED IN DETAIL

The report employs a bottom-up modeling approach to provide accurate estimates and projections, starting with granular data at the equipment level and scaling up to industry-wide trends. This methodology, informed by historical shipment data and operational practices, offers a robust foundation for assessing past energy use and predicting future scenarios. The model integrates critical updates, including detailed characterizations of AI hardware, cooling systems, and regional variations in carbon intensity. These refinements make the methodology well-suited for capturing the complex interplay between technological advancements and energy demands. The model's key innovations lie in its ability to account for emerging technologies, particularly the rapid deployment of AI hardware. It also evaluates variations in cooling efficiency, considering both physical systems and climatic conditions at data center locations. Including carbon and water intensity metrics highlights the broader environmental impacts of data center operations, offering a comprehensive view of their resource use.



HISTORICAL TRENDS AND SHIFTING PARADIGMS

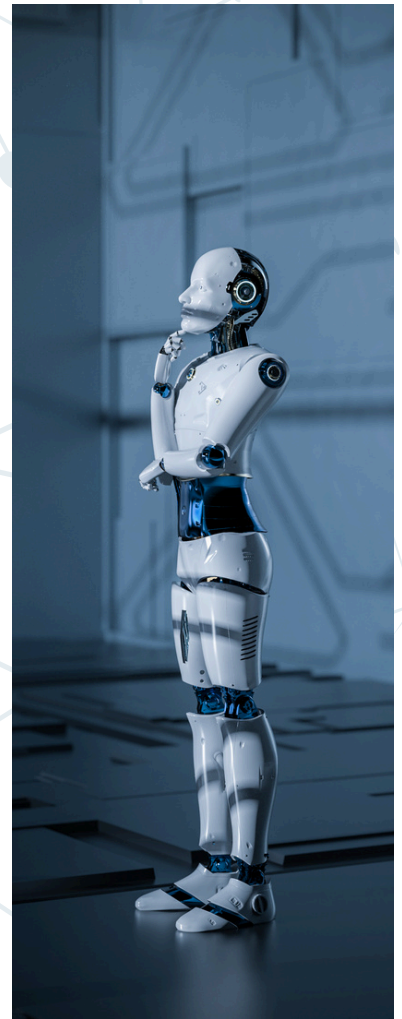
From 2014 to 2023, U.S. data centers experienced a significant transformation. Energy use increased from 60 terawatt-hours (TWh) to 176 TWh, representing 4.4 percent of total U.S. electricity consumption in 2023. This shift can largely be attributed to the rise of GPU-accelerated servers, which have become critical to AI workloads. By 2018, the growing adoption of these specialized servers had reversed the trend of minimal energy growth observed in previous years. The report demonstrates that these changes depart from the era of flat energy consumption driven by efficiency gains.



The transformation was not confined to AI hardware alone. Storage technologies shifted from traditional hard disk drives to flash storage, while networking equipment evolved to support higher-speed operations. These developments have compounded the sector's energy needs, particularly as storage density increases and network efficiency improvements struggle to offset growing demands. The rapid deployment of AI hardware is a pivotal energy consumption driver. The report reveals how AI servers operate at higher power and utilization rates than conventional servers, emphasizing their energy intensity. Adopting GPUs and other accelerators for tasks like AI training and inference has introduced new variables in predicting energy trends, as these workloads operate differently from traditional enterprise computing tasks.

PROJECTIONS AND SCENARIOS FOR FUTURE GROWTH

The report forecasts substantial increases in energy consumption, projecting that U.S. data centers could use between 325 and 580 TWh annually by 2028. This would account for 6.7 to 12 percent of total U.S. electricity consumption. These projections are presented as a range, reflecting uncertainties around factors such as the rate of GPU deployment, operational practices, and the efficiency of future cooling technologies.



The introduction of AI-specific servers is expected to continue reshaping the energy landscape. By 2028, AI servers could constitute a significant portion of the total installed server base, driving up operational and cooling demands. While liquid cooling systems and other innovations may mitigate some of these increases, the report underscores the need for systemic changes in data center design and operation to manage this growth sustainably.

The projections also emphasize the importance of regional considerations. Data centers are concentrated in specific geographic areas, and the energy sources powering these regions influence the carbon intensity of their operations. As such, the environmental impacts of this growth will vary depending on the energy mix of local grids.

THE ROLE OF CRYPTO CURRENCY MINING

The report dedicates a section to analyzing the impact of cryptocurrency mining, particularly Bitcoin. While this sector is relatively niche compared to broader data center operations, its energy intensity is disproportionately high. The computational demands of blockchain validation contribute to significant electricity consumption, and future trends will depend on market dynamics, including cryptocurrency values and mining efficiency. The report's findings suggest that aggressive growth scenarios in cryptocurrency mining could exacerbate challenges for energy infrastructure, highlighting the need for targeted policy interventions.



OPPORTUNITIES

FOR INNOVATION & COLLABORATION

Despite the challenges posed by rising energy demands, the report identifies several pathways for addressing these issues. Technological innovation remains a cornerstone of these efforts. Developing advanced cooling systems, such as liquid and evaporative cooling, offers opportunities to reduce the energy intensity of data center operations. Similarly, server design and operational efficiency improvements could offset some of the projected growth in demand.

Collaboration between data center operators and utilities is another critical area for progress. Enhanced transparency in energy use and benchmarking initiatives can support better planning and investment in infrastructure. The report calls for integrating data center needs within broader energy infrastructure planning, ensuring the sector's growth aligns with the transition to renewable energy sources.

Future research priorities include exploring the unique energy demands of AI workloads, developing more precise forecasting models, and evaluating the long-term implications of emerging technologies. These efforts will require coordination among industry stakeholders, policymakers, and researchers.



A BROADER CONTEXT FOR ENERGY CHALLENGES

The analysis in this report aligns with broader trends observed in the energy sector. Reports from organizations such as the International Energy Agency (IEA) and consulting firms like McKinsey highlight similar concerns about the impact of AI and cloud computing on global energy demand. However, the report distinguishes itself by offering a detailed, equipment-level perspective that bridges the gap between macroeconomic projections and operational realities.

Data centers are unique in their rapid evolution and dependence on cutting-edge technologies compared to other energy-intensive sectors. This dynamism presents both challenges and opportunities. While energy demand is expected to rise sharply, the scalability of data center operations and their integration with renewable energy sources offer pathways for sustainable growth.



CONCLUSION

PREPARING FOR A HIGH-ENERGY FUTURE

The 2024 United States Data Center Energy Usage Report underscores the urgency of addressing the rising energy demands of data centers. The findings illustrate the importance of technological innovation, strategic planning, and collaboration in navigating this complex landscape. By focusing on efficiency and sustainability, stakeholders can ensure that data centers continue to support economic growth while minimizing their environmental impact. This report provides a foundation for understanding and addressing the challenges ahead, positioning the sector as a leader in transitioning to a digital and decarbonized future.



VEDENI ENERGY ANALYSIS

The 2024 United States Data Center Energy Usage Report, produced by Lawrence Berkeley National Laboratory (LBNL), provides a comprehensive analysis of the escalating energy demands within the U.S. data center industry. The report indicates that data center energy consumption has tripled over the past decade and is projected to double or triple again by 2028, potentially accounting for up to 12% of total U.S. electricity consumption.

GLOBAL CONTEXT & COMPARISONS

This trend is not confined to the United States. Globally, data centers are experiencing similar surges in energy consumption due to the integration of artificial intelligence (AI) and other high-performance computing tasks. The International Data Corporation (IDC) forecasts that global electricity consumption by data centers will more than double, reaching 857 terawatt-hours (TWh) by 2028. AI-specific workloads are anticipated to grow even faster, consuming 146.2 TWh by 2027.

In Europe, the surge in AI-driven data processing has prompted concerns about energy sustainability. Reports indicate that the rapid expansion of data centers, essential for AI operations, is exerting unprecedented pressure on energy resources, potentially complicating climate policies to reduce carbon emissions.

TECHNOLOGICAL INNOVATIONS AND ENERGY EFFICIENCY

The industry is exploring various technological innovations to mitigate the environmental impact of increased energy consumption. Liquid cooling systems have emerged as a promising solution, with companies like Super Micro Computer introducing liquid-cooled AI data center systems that can reduce power usage by up to 40%. Developing neuromorphic hardware, which mimics the human brain's neural architecture, also offers the potential for more energy-efficient AI processing. Research suggests that integrating neuromorphic systems into data centers could significantly enhance energy efficiency, though widespread adoption faces challenges with hardware and software integration.

ENERGY INFRASTRUCTURE AND POLICY IMPLICATIONS

The escalating energy demands of AI data centers are influencing energy infrastructure and policy decisions. In the United States, tech companies are exploring partnerships with nuclear power providers to secure stable energy supplies for data centers. For instance, Microsoft has entered into an agreement to reopen the Three Mile Island nuclear plant to power its AI centers, highlighting a trend toward seeking reliable and sustainable energy sources.

In Malaysia, the government has indicated that global tech companies investing in data centers should prepare for higher costs for water and energy, reflecting concerns over resource sustainability. The country is encouraging data centers to utilize renewable energy, aligning with its goal of achieving 70% renewable energy capacity by 2050.

FUTURE OUTLOOK AND RECOMMENDATIONS

The convergence of AI growth and data center energy consumption presents challenges and opportunities. While AI workloads are inherently energy-intensive, AI also offers the potential for enhancing energy efficiency within data centers. Implementing AI-driven efficiency tools can help manage and optimize energy usage, contributing to more sustainable operations.

A multifaceted approach is necessary to address the dual challenges of supporting AI advancements and ensuring energy sustainability. This includes continued investment in energy-efficient technologies, strategic partnerships with renewable energy providers, and the development of policies that promote sustainable practices within the data center industry. By adopting these strategies, the industry can strive to balance the demands of technological innovation with the imperative of environmental stewardship.



About Vedeni energy



Vedeni Energy offers specialized services designed to help businesses navigate the complexities of the modern energy landscape. Our offerings are tailored to meet the unique needs of utilities, independent power producers, regulatory bodies, and other stakeholders, ensuring success through strategic insights, expert guidance, and innovative solutions.



VedeniSpark+, a service provided by Vedeni Energy, is designed to help startups and established companies secure the capital funding necessary for growth and success. Our team of seasoned advisors works closely with clients to develop tailored funding strategies that align with their business goals and financial requirements.



Vedeni Energy's VedeniIQ+ service provides actionable wholesale electric power market intelligence that enables clients to make informed decisions confidently. Our expert analysis and reporting distill complex energy market information into clear, concise insights, helping organizations elevate their market strategies, influence policy, and identify new opportunities.



+1 463-266 4496



VEDENI.ENERGY.COM



INFO@VEDENI.ENERGY



WHITESTOWN, IN, 46075, USA

TO LEARN MORE, VISIT US AT

WWW.VEDENI.ENERGY