

Powering AI: The Energy Demands of Data Centers

Overview

Since the release of generative artificial intelligence (AI) models like ChatGPT, Gemini, and Perplexity, American businesses, consumers, and the federal government have rapidly integrated these tools into their daily operations. This widespread adoption has increased attention to the substantial computing power required to train and run AI systems. Unlike traditional computing tasks, training advanced AI algorithms requires immense processing power and vast data, making it significantly more energy-intensive. As AI models grow more complex, the data centers that support them consume more electricity, placing mounting pressure on the strained U.S. electricity grid. Expanding the grid's capacity to meet this new demand will take years—if not decades—prompting industry leaders and policymakers to seek alternative solutions. Without proactive measures to address this growing demand, regions with high concentrations of data centers—such as Northern Virginia, California, and Texas—could face rising utility costs and, potentially, rolling blackouts. On-site mini-grids and alternative energy sources can alleviate data centers' burden on the grid, promoting their energy security and reliability. This Basic will evaluate the challenges facing the U.S. electricity grid as demand from AI data centers grows and explore the potential of alternative energy sources to alleviate grid pressure.

Energy Demand

Data centers across the U.S. currently account for 4% of the nation's electricity use, comparable to the total annual consumption of New York State. A 2024 McKinsey & Company report projects U.S. data center electricity consumption will triple from 25 gigawatts (GW) in 2024 to more than 80 GW by 2030, making up 12% of total U.S. demand. One GW of electricity is enough to power 750,000 U.S. homes, and the average natural gas power plant produces 500 megawatts (MW), with 1,000 MW equalling a GW. This demand growth is driven by increasing digitalization, cloud migration, and the development of AI. Meeting a quarter of this demand by the decade's end would require the U.S. to build 50 to 60 GW of additional power infrastructure.

Generative AI algorithm queries require 10 to 100 times more electricity than standard internet searches due to their use of **graphic processing units** (GPUs). GPUs process graphics and perform complex calculations, allowing generative AI to learn from existing data and create new content. Historically, data centers relied on central processing units to facilitate internet queries, running at roughly 200 watts per chip. AI developers shifted from CPUs to GPUs because the latter can process graphics more efficiently in parallel; however, their increased computing power requires more energy. The rising processing needs of AI Center Forward Basics May 2025

Center Forward Basics

Center Forward brings together members of Congress, not-for profits, academic experts, trade associations, corporations and unions to find common ground. Our mission: to give centrist allies the information they need to craft common sense solutions, and provide those allies the support they need to turn those ideas into results.

In order to meet our challenges we need to put aside the partisan bickering that has gridlocked Washington and come together to find common sense solutions.

For more information, please visit <u>www.center-forward.org</u>

Key Terms:

- **Graphic Processing Units (GPUs)**computer chips that perform rapid graphics-related calculations in parallel to smoothly render digital images.
- Smart grids- an electricity network that uses digital technologies, sensors, and software to better match the supply and demand of electricity in real-time while minimizing costs and maintaining the stability and reliability of the grid.
- Intermittency Issues- inconsistency in energy production by renewable sources, often due to natural weather variability.
- Long-duration energy storage systems- technologies that capture and store excess energy, and when needed, can use this energy to

models have further amplified demand, with energy consumption of GPUs used for AI surging exponentially from 400 watts per unit in 2022 to 700 watts in 2023 and reaching 1,200 watts in 2024.

However, this surge in electricity demand may be overestimated. Advances in energy and processing efficiency can help mitigate demand. For example, Google recently developed Tensor Processing Units, custom chips that enhance energy efficiency during AI processing by utilizing AI to manage cooling and optimize data center operations. **Smart grids** are also gaining popularity, utilizing Phasor Measurement Units to assess grid stability, automatically report outages, automate feeder rerouting, and access battery storage during high electricity demand.

Grid Capacity Challenges and Potential Solutions

The United States' electricity grid consists of more than 9,200 electric power plants, which output approximately 1 million megawatts of electricity over 600,000 miles of transmission lines. However, data centers tend to be concentrated in specific regions of the U.S., resulting in localized demand surges. Northern Virginia, known as Data Center Alley, has the world's largest concentration of data centers, with 70% of global traffic passing through them. The Eastern Interconnection grid that serves Northern Virginia is already struggling to keep up with increasing demand. In 2022, utility company Dominion Energy said it might not be able to meet energy demand in areas densely populated with centers. As a result, Virginians' utility bills are expected to increase by a little over 2.5% annually for the next 15 years, as demand is projected to double, not only from AI but also from the growing use of electric vehicles and population expansion. produce electricity for 10 hours or more.

- **Dispatchable-** a power source that can be adjusted instantly by grid operators to match supply with electricity demand.
- **On-site mini-grids** a set of small-scale electricity generators and energy storage systems deployed at or near the point of use, operating independently from the national transmission grid.
- **Rate Class-** a category of electricity consumers assigned specific pricing based on their energy consumption levels. In this context, it refers to a higher charge classification for facilities using 20 MW or more.
- Nuclear Regulatory Commission- an independent agency created by Congress in 1974 to ensure the safe use of radioactive materials for beneficial civilian purposes while protecting people and the environment.

Leading AI companies are investing in renewable energy sources to supplement the energy they receive from the U.S. grid. In May 2024, Microsoft acquired more than 10.5 GW of wind and solar projects from Brookfield Renewable to power its data centers with 24/7 carbon-free energy—the largest corporate power purchase agreement of its kind. Google is also the sole customer of the Charlottesville-based Apex Clean Energy wind farm, using all 189 MW it generates to power data centers along the East Coast. However, solar and wind power sources have **intermittency issues**, necessitating the use of backup sources, such as natural gas or **long-duration energy storage systems**, during periods of low generation.

On the other hand, nuclear power plants operate at maximum power more than 92% of the time, higher than any other energy source, including natural gas. Their high-density uranium fuel enables greater energy output per unit and is **dispatchable**, improving efficiency. These advantages make nuclear power an appealing investment for technology companies. Microsoft signed a 20-year power purchase agreement for the nuclear facility on Three Mile Island, which is expected to generate 837 MW of energy for its data centers in Chicago, Virginia, and Ohio. Additionally, the company secured a 24/7 nuclear power deal with Constellation to supply its Boydton data center.

However, these renewable and nuclear energy sources will not be operational for several years. Dominion Energy estimates that its projects in Virginia will take another 3 to 5 years to complete, and Three Mile Island is set to come online in 2028. This timeframe is comparable to an average-sized natural gas plant, typically taking 2 to 4 years to construct and become operational. These delays threaten consumers, who will see their utility prices rise and power reliability decline until the new projects come online.

On-site mini-grids are a form of independent power generation designed to seamlessly transition a data center's energy supply from the main grid during peak demand. They integrate multiple energy sources, including renewables (wind and solar), small modular nuclear reactors, combined heat and power systems, and excess energy stored in batteries. Unlike building new infrastructure to support renewable energy projects, microgrids can be fully operational in just a few months. A medium-sized mini-grid can be installed in three months, producing approximately 26 megawatt-hours (MWh) of power (roughly enough to power 1,000 homes). While mini-grids present real advantages, building a new grid requires serious capital investment and is inaccessible for many smaller firms and data centers. However, mini-grids can potentially reduce data centers' high long-term utility costs by enabling them to generate their own electricity and disconnect from the U.S. grid during peak demand. They also support decarbonization efforts through the integration of renewable energy.

Looking Ahead

In the absence of aggressive federal action, some states are stepping in to address the growing energy demands of data centers. In January, Virginia's legislature introduced a bipartisan package aimed at protecting consumers from rising electricity costs and preserving the state's climate goals. The move followed a report from the Joint Legislative Audit and Review Commission, which found that data centers are responsible for nearly all of Virginia's recent increase in electricity demand. In Oregon, State Representative Pam Marsh (D) plans to reintroduce a bill that creates a separate **rate class** for data centers, ensuring their increased energy use does not result in high costs for residential consumers. New York State Senator Kristen Gonzalez (D) also plans to reintroduce the *Sustainable Data Centers Act*, requiring data centers to operate on 100% renewable energy by 2050, ensuring the state meets its net-zero emissions target that year.

Data centers are also a major source of economic investment, prompting some state leaders to voice concerns that increased regulations would lead companies to invest elsewhere. In his State of the Commonwealth speech, Virginia Governor Glenn Youngkin (R) cited data centers' critical role in creating 74,000 jobs and contributing more than \$9 billion to Virginia's economy. In her annual address, New York Governor Kathy Hochul relayed a similar sentiment, touting data center development as a key economic driver for New York.

The Trump administration has expressed its commitment to increasing investment in AI infrastructure and data centers in the United States. On January 22, the White House announced the creation of Project Stargate, an investment project led by OpenAI, SoftBank, and Oracle, which aims to invest \$500 billion in new data center infrastructure to power AI development. As the demand for AI-driven data centers continues to surge, balancing innovation with alternative energy solutions and regulatory measures will be crucial to ensuring both economic growth and environmental responsibility in the years to come.

Links to Other Resources

- Bloomberg Law- <u>States Propose Data Center Energy Guardrails as Demand Soars (1)</u>
- Columbia Center on Global Energy Policy- <u>Projecting the Electricity Demand Growth of Generative AI Large Language</u>
 <u>Models in the US</u>
- Data Center Dynamics- <u>All power is local: The nitty-gritty of the US energy grid</u>
- Department of Energy- <u>Grid Modernization and the Smart Grid | Department of Energy</u>
- Pillsbury Law- <u>Trump 2.0: AI and Data Centers in a Time of Legal and Technological Disruption</u>
- VPM- Data centers keep growing in Virginia and so does energy demand