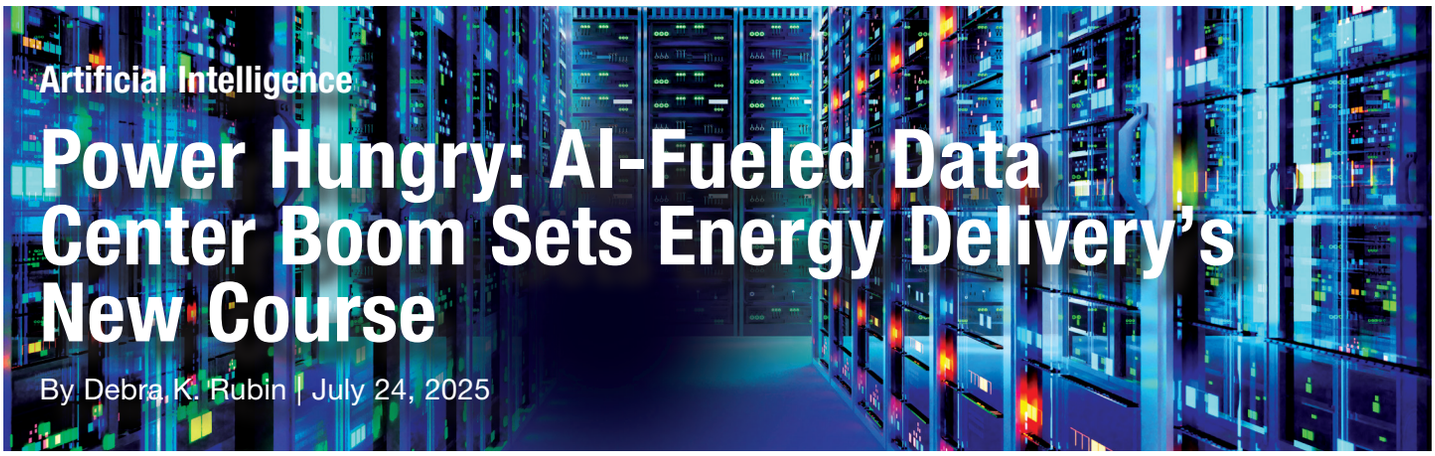
THE CONSTRUCTION RESOURCE FOR OVER 150 YEARS

ENR

Engineering News-Record



Artificial Intelligence

Power Hungry: AI-Fueled Data Center Boom Sets Energy Delivery's New Course

By Debra K. Rubin | July 24, 2025

The statistics seem astounding, as forecasts for artificial intelligence demand push expansion of data center and power infrastructure into hyper-blitz in the U.S., and beyond

The U.S. Energy Dept. said last year that data centers already account for more than 4% of U.S. electricity use, which could grow to 12% by 2028—akin to 580 billion kilowatt hours, with AI use comprising up to 40% of global data-center power demand by 2026, according to research in energy journal Joule. Consultant McKinsey & Co.'s outlook last year pointed to a rise in global capacity annual demand of up to 22% in 2030, amounting to between 171 and 219 GW. "This contrasts with the current demand of 60 GW, raising potential for a significant supply deficit," said the research firm. Moody's Investors predicts investment to exceed $2 trillion in the next four years.

But turning the investment into real infrastructure with adequate power wont be easy, industry practitioners say.

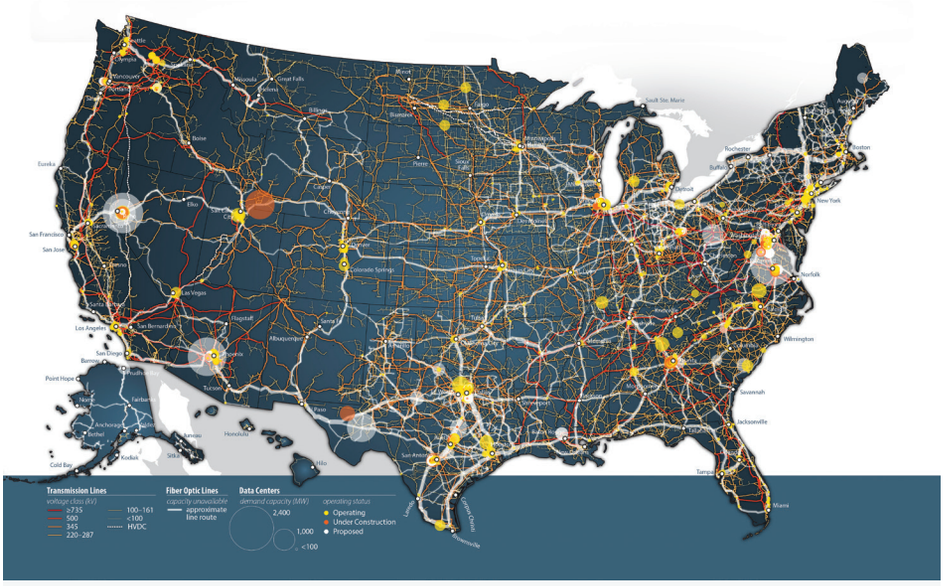
"Power load and transmission shortfalls are a significant challenge for the data center industry, especially as hyperscalers continue to accelerate timelines for getting facilities online," said Maja Rosenquist, executive vice president of Mortenson Construction, which has been involved in completing an estimated 2.4-million-sq-ft data center in Eagle Mountain, Utah that media report will serve tech giant Meta. "We're seeing projects that have broken ground but are experiencing slowdowns due to energy constraints."

She added that the surge in AI workloads "is forcing a rethink of cooling and power systems," causing project delay for redesign. "Labor and supply chain constraints continue to pad construction timelines, while higher costs are making speculative builds tougher to finance," said Rosenquist. "We expect more selective, strategic development as these headwinds persist."

Christine Wood, Burns & McDonnell vice president and data centers practice lead added that "while we are innovating and collaborating to squeeze more capacity from the grid, the reality is that the speed at which owners want to build and operate new data centers is often at odds with the pace of power and transmission infrastructure development." Without major changes in permitting,

Looking at US Data Center Infrastructure in 2025

U.S. National Renewable Energy Lab map shows, as of mid-2025, the geospatial relationship between U.S. data center sites, transmission infrastructure, fiber optic networks and metro areas, with color-coded transmission voltage classes and approximate fiber optic line routes. Varied circle sizes show local data center capacity.



Source: National Renewable Energy Laboratory (NREL). NREL/OT-6A20-94502. Map by Billy Roberts, NREL. NOTE: NREL did not produce or curate map datasets and is not responsible for their accuracy, completeness or reliability.

policy and technology adoption, she added, "the gap between demand and deliverable power supply will remain a critical constraint for the industry."

Colleague Keegan Odle, Burns & McDonnell vice president of ascending markets for its transmission and distribution group, stressed that "with everyone building big, we are quickly reaching the limits of what current infrastructure can support, and that means we'll have to get creative to meet demand."

'Manhattan Project'

U.S. Energy Secretary Chris Wright has termed the effort to expand centers and their power sources "the second Manhattan Project," comparing it to the massive atomic bomb development effort during World War II.

Said contractor Clayco CEO Anthony Johnson, noting a firm-sponsored survey in June that says 56% of respondents are concerned about a U.S. lag in building infrastructure to power AI innovation: "We are in the midst of a modern space race … to determine who will lead the AI and quantum computing revolution." Duke Austin, CEO of major energy sector contractor Quanta Services, told investors at a May 28 investment conference: "We sit at the nucleus of what I consider one of the biggest builds of infrastructure that we've seen." Wall Street analysts report the firm stock up 21% so far this year and as much as 870% since 2020, citing the AI-fueled power market transformation.

Meta was reported to be seeking to raise $29 billion from private capital for data centers in the U.S., while already building a $10-billion facility in Louisiana powered by three new natural gas power plants. On a podcast in May, CEO Mark Zuckerberg said the U.S. must rapidly expand data center investment or risk falling behind China in AI capacity, stressing the need "to focus on streamlining the ability to build centers and produce energy … or we'll be at a significant disadvantage."

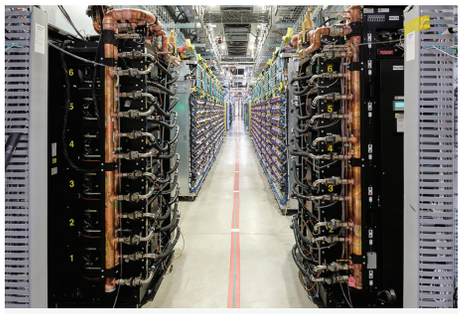
While both real and overhyped—and amid intense debate over cost, energy sources, power delivery method and emerging consequences—the data center push is gaining marketplace traction.

Growing AI power demand is seen as the main catalyst for the first expansion of U.S. nuclear energy capacity in a generation. One of four President Donald Trump executive orders in May set to propel nuclear power resurgence, calls for having 10 large reactors under construction by 2030. He also has directed the U.S. Energy Dept. to designate AI data centers as critical defense facilities.

Intense data center growth in northern Virginia's "Data Center Alley," such as this planned Amazon facility in Ashburn, have caused concerns over environmental and community impacts. *Image courtesy Amazon Web Services*

mand, that's why baseload nuclear is super super important," said Bruce Flatt, CEO of infrastructure investor Brookfield Asset Management, which has a majority stake in nuclear technology innovator Westinghouse. "We are talking with firms to build sole-source nuclear plants for data centers. These will be an enormous export product."

Data center developers are "shopping around both within your community and next door," said Patricia Taylor, director of policy and research for the American Public Power Association. "Strategic locations are not found, they are engineered," said Masjood Jafri, Arcadis national energy transition strategic advisor and U.S. business development lead. "Performing rapid but robust due diligence in selecting project locations is key to maximize use of limited pre-investment resources as there is little recovery once major decisions are taken with limited or inaccurate information "



States such as Ohio, site of this Google data center, are changing power cost allocation rules for large power users. *Image courtesy Google*

'Action Plan'

With a goal to beat China in AI dominance, the Trump administration released on July 23 a 28-page "action plan" to expand data center and energy infrastructure, and boost U.S. technology exports and security. Claiming to use "every rule at our disposal," President Donald Trump said the strategy, backed

Startup Fermi America just unveiled plan for data center-power campus in Texas, said to cost $300 billion and provide 11 GW from fossil fuel and cleaner sources when fully completed. Its Initial 1-GW phase is aimed to be on line in 2026. *Image courtesy Fermi America*

by executive order, would streamline AI-related project permitting and environmental review and use federal lands to build data centers and expedite their grid connections.

The U.S. Energy said July 24 it selected 'sites at Idaho National Laboratory, Oak Ridge Reservation in Tennessee, Paducah Gaseous Diffusion Plant in Kentucky and Savannah River Site in South Carolina "to develop cutting edge AI data center and energy generation projects." The U.S. Environmental Protection Agency also is directed to identify brownfield and Superfund sites for potential data center reuse.

The plan reiterates the administration call to postpone shutdown of aging coal plants and related generation and push faster interconnection of new dispatchable power such as enhanced geothermal and nuclear fission and fusion. It aims to leverage backup power sources to bolster grid reliability during peak demand and reform the power market to "align financial incentives with the goal of grid stability," said S&P Global, but with no further detail.

Following the trend of Trump rhetoric and recent actions, the order fails to include renewable energy sources such as wind and solar under its provisions, even as data center developers need choices to meet projected power demands. Jeff Tench, executive vice president of Vantage Data Centers, told a Senate Energy and Natural Resources Committee hearing July 24. that the firm's "requirement is for more electrons. Vantage is relatively agnostic as to the source of those electrons." Recent administration rules that deter renewable energy projects "slow down the process of approving new generation or new transmission [and], would definitely be a negative for our business," he said.

"The Trump administration is blatantly seizing on AI and data center load growth as an excuse to justify [an] expensive fossil bailout agenda," Laurie Williams, director of the Sierra Club's Beyond Coal Campaign, said in a statement, terming it "a gross mismanagement of our country's energy needs."

Also proposed are steps to expedite permits to build data centers, semiconductor plants and energy infrastructure through categorical exclusions under the National Environmental Policy Act (NEPA) and streamlined rules under the federal Clean Air and Clean Water Acts and other environmental laws.

The administration said it could also set up a nationwide Section 404 permit under the clean water law to govern release of dredged or fill material into protected U.S. waters that "would not require a pre-construction notification and would cover development of sites consistent with the size of a modern AI data center," added S&P Global. It also eyes expanding use of the FAST-41 environmental review process for data center and data center energy projects.

Projects could qualify if they have at least a $500-million investment and will use more than 100 MW of electricity.

Also outlined is a push to boost skills and numbers in the AI-related workforce, with new training efforts by the US Labor and Education Depts, "including roles such as electricians, advanced HVAC technicians and a host of other high-paying occupations," the plan said. Under the strategy, however, federal agencies could rescind state regulations that block AI development and tie a state's oversight stance to future federal funding awards.

"Bluntly, there is no acceptable version of an AI moratorium," said more than 140 groups opposing the state oversight curbs, including unions and legal, education and technology advocates in a July 23 letter to Trump. They also lobbied Congressional lawmakers against a similar provision for a 10-year restriction on state AI regulation that was added to the budget reconciliation bill but removed in a 99-1 Senate vote before its passage.

All plan components are set to be executed in the next six to 12 months, Michael Krastios, director of the White House Office of Science and Technology, said at a press briefing.

The strategy "will accelerate infrastructure readiness so AI can be built and used here, and help students and workers with skills needed to win in an AI-powered global economy," Fred Humphries, Microsoft corporate vice president of U.S. government affairs, posted on social media.

'Hyper Scaling'

The plan follows a July 15 "Energy and Innovation Summit" in Pittsburgh organized by Pennsylvania Senate Republican David McCormick, which featured administration heads as well as academic, investment, technology and energy sector chiefs, who touted up to $90 billion in recent and new commitments to data center and energy infrastructure in the state. These include $20 billion from Amazon Web Services to build out infrastructure and $15 billion from FirstEnergy to expand and fortify power grids.

"Our intention is to raise $30 billion of capital and raise another associated $100 billion of debt working with Google and all the other hyperscalers," Larry Fink, CEO of BlackRock, told the conference.

In April, Homer City Generating Station—once Pennsylvania's largest coal plant—announced it would convert to be a 4.6-GW natural gas facility to power a planned $10-billion data center campus of 3,200 acres, equipped with seven gas turbines from GE Vernova. The project sponsor is a private investor group, and Kiewit Corp. is design-builder.

Hyperscalers have announced other efforts to move quickly and in a big way. The Stargate initiative, touted by Trump in January with key backers including technology giants OpenAI and Oracle and sector investor SoftBank, is aimed to invest $500 billion through 2029 to build data centers and infrastructure. But the effort has since reduced its scale and speed, with developer Crusoe and contractor DPR building two centers in Abilene, Texas, the first set to operate this year and the second in mid-2026,

This month, startup Fermi America, backed by former U.S. Energy Secretary Rick Perry and other investors, unveiled HyperGrid, a planned data center and power complex on 5,800 acres in Amarillo, Texas, that is owned by Texas Tech University. According to Fermi America co-founder and investor Toby Neugebauer: "We're going to create the largest private grid in the world." The development is said to cost $300 billion and will provide 11 GW of power from natural gas, solar energy and eventually from four Westinghouse AP-1000 nuclear reactors when fully completed. Its Initial 1-GW phase is aimed to be on line in 2026, with early site engineering finished, said the firm.

A Fermi America spokeswoman said it will reveal names of nuclear energy veterans hired as experts and release procurement detail "in the near future," but she did not disclose project investors or investment totals. The reactor building plan was submitted to the U.S. Nuclear Regulatory Commission on June 17 and "accepted for review in record time," the spokeswoman said.

Data center demand has created development hubs across the U.S., such as in northern Virginia. Dubbed "Data Center Alley," and said to be the world's largest concentration of the facilities, it continues to grow. Large state utility Dominion Energy, which hosts 13% of global data center capacity, reported in its May quarterly results that data related power projects in development nearly doubled to 40.2 GW in December from 21.4 GW last July. The utility company projects 85% demand growth

over the next 15 years.

Dominion CEO Robert Blue still expects that demand growth—even after Chinese AI company DeepSeek released earlier this year a more energy-efficient AI model that industry observers speculated could reduce data center space and power needs, and as U.S. economic uncertainty from Trump tariffs further affect AI infrastructure development scope and pace. "We're seeing continued appetite for additional data center capacity in our service territory," he said. Developers "want to go fast … That's their business, that's always been their business," Blue said. "I don't see any reason why that's going to change."

Energy Builds Ahead

Natural gas is the dominant form of energy being used to power the data center surge. S&P Global Ratings estimates that gas demand from centers could increase by 3-6 billion cu ft per day.

Williams Co., a major U.S. gas pipeline developer, is building gas-powered electric generation for Meta's data center campus in Ohio's developing hub outside of Columbus. Its estimated $1.6-billion investment in multiple plants in the area is the company's first venture into power generation. Williams said it could provide 1 GW of power for data centers by 2027. The firm "will be a big beneficiary of fast-rising data center power load," CEO Alan Armstrong told investors on a first-quarter results call in May.

Utility company PPL Corp. has a new joint venture with investor Blackstone Infrastructure—also announced at the Pittsburgh energy summit—to develop gas-fired power plants for data centers in Pennsylvania and in the 13-state PJM Interconnection grid region. The grid operator's forecasted load is 167 GW by winter of 2030 and 200 GW by 2035, "almost exclusively being driven by data centers," said Matt Green, TRC Cos. senior vice president who head its integrated energy planning and advisory practice. He says "estimates are still quite conservative, since 50 gigawatts of data center projects are currently in queue."

Terra Data Center in San Jose, Calif., a planned 295,000-sq-ft facility, is being developed with natural gas fuel cells as the primary fuel source and set to be a near-net zero facility completely off the grid, said Jafri of Arcadis. "It is expected that more and more data centers will have baseload power being provided by the co-located power generation solution." he noted, adding that San Jose faciity would have absorption chillers that use heat from the fuel cells to generate

Susquehanna nuclear plant in Pennsylvania, owned by Talen Energy, aims to have future co-located Amazon data centers. The owner also said July 25 that is buying a nearby combined cycle gas plant and one in Ohio totaling 2.9 GW for $3.5 billion to supply other planned data centers.
Image courtesy of Amazon Web Services

Data center development experts debated gas power dependence at a recent Reuters energy conference. Joseph Vellone, CEO of ChargeScape, a virtual power plant software developer, acknowledged that despite gas power plant expense, data center sponsors say its reliability "now is more important. There will be a lot of gas in the next decade or two." Market research firm Morningstar said in a new report that gas would power 60% of new center demand, and renewable energy 25%.

John DeAngelis, Meta's head of clean technology innovation, called for an industry effort at the Reuters event to reduce energy intensity of data centers. "We have to prioritize that … and do it first," he said. "What is the cheapest megawatt hour? It's the one never consumed."

Going Greener

Geothermal energy is a growing renewables contender, with tech firms seeking 24/7 lower-carbon power for data centers and office campuses. In May, the Nevada Public Utilities Commission approved what a Google spokesman said was an unprecedented energy supply agreement with geothermal developer Fervo Energy to add 115 MW of enhanced geothermal power delivered by Nevada Energy. "This represents nearly 30 times growth in geothermal capacity … since [Google] announced its successful commercial pilot with Fervo in 2023," he said.

The ruling also approves the first U.S. Clean Transition Tariff, a new power supply rate allowing utilities and customers to enter long-term agreements to invest in new grid-scale clean energy

Mortenson Construction crew builds Eagle Mountain, Utah, data facility expansion for Meta.
Photo courtesy Mortenson Construction

Google can "receive exclusive electric service … from [Fervo's] 115-MW Corsac Station Enhanced Geothermal project, once it reaches commercial operation." The energy supply period is set to last for 15 years. Corsac Station is a single-phase development set to begin in 2026 and commercially operate in 2027, according to reports.

In an announced June deal, Meta said geothermal energy producer XGS Energy would provide up to 150 MW to support its New Mexico data center operations. In exchange, the tech firm will invest about $1 billion and create 3,000 construction jobs, said Gov. Michelle Lujan Grisham (D). Sage Geosystems also said it is developing a 150-MW geothermal energy project for the tech firm to offer a new power option for data systems east of the Rocky Mountains "where it has not been possible before." Phase one will deliver 8 MW of power by 2027, and phase two will deliver up to 150 MW, with an option for 200 MW more, said Sage CEO Cindy Taff.

The U.S. Bureau of Land Manage-

ment said in July it is undoing by September "obsolete" geothermal site leasing rules for projects, which also have a much longer window, until 2033, to qualify for federal tax credits under the new budget bill. The agency said it also plans to hold its next competitive geothermal lease sale Aug. 26 in California of 13 parcels across 23,000 acres, with comparable and larger lease sales to follow this year in Idaho and Nevada.

Google and Brookfield Asset Management also announced at the energy summit a deal to invest $3 billion in hydropower for data centers and other needs. Albany, N.Y.-based developer Soluna Holdings Inc., which buys excess power from wind, hydro and solar energy plants not sold to power grids for its co-located data centers, said it expects to have at least 800 MW of facilities operating or under construction by next year.

Alberta-based Enbridge said July 23 it has a new long-term contract to supply Meta data center operations in Texas with 100% renewable energy from the energy firm's utility-scale 600-MW solar power facility near San Antonio. Enbridge said it reached final investment decision on the Clear Fork project, with an estimated project cost of $900 million and set to operate in summer 2027. Meta recently said it anticipates adding 9.8 GW of renewable energy to local grids in the U.S. by the end of 2025.

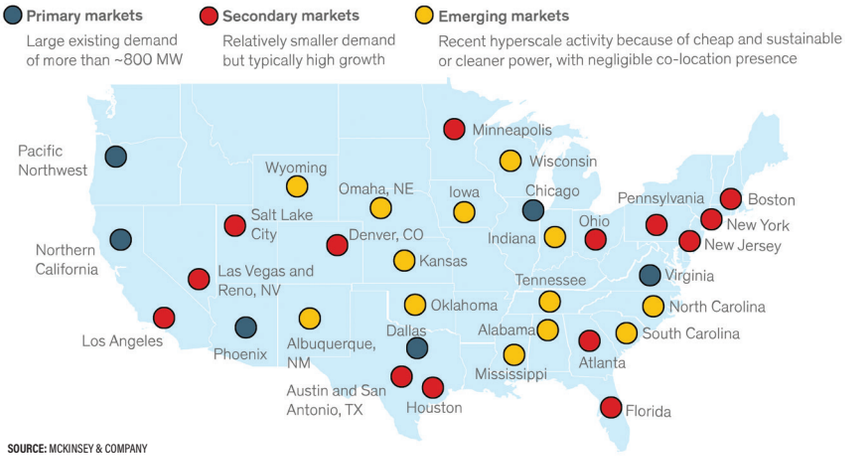
"Clear Fork demonstrates the growing demand for renewable power across North America from blue-chip companies who are involved in technology and data center operations," said Matthew Akman, Enbridge executive vice president and president of power. The announcement was the latest in recent large-scale renewable energy purchases by Meta.

"Renewables are advantaged by their attractive economics, quick time to power and zero carbon emissions," Morningstar analysts said. "We expect rising power demand from AI to drive incremental demand for new wind, solar and battery storage generation."

But impacts of Trump orders to curtail the sector through orders, directives to Congress and halted or curbed federal tax credits raise new workforce impact concerns for cleaner data center development, if growing renewable energy sector employment and training drops. "Even as domestic oil and gas production reached their highest levels ever in 2024, employment .... is down nearly 20% from five years ago ... mostly due to automation technologies," said a new survey by the Center for American Progress. "Oil and gas job growth is unlikely to compensate for the jobs that could be lost in the wind energy industry."

DATA CENTERS AND CHANGING POWER MARKETS

Three tiers of US energy markets



SOURCE: MCKINSEY & COMPANY

Nuclear power for data centers remains a longer term development, but states are also boosting the effort. Wisconsin Gov. Tony Evers (D) signed into law on July 3 legislation that speeds up the construction timeline for new reactors by directing a siting study at both existing power plants and new locations. The state is gaining as a data center hub, with Microsoft building a $3 billion data center in Mount Pleasant, and Meta Platforms Inc. planning a more than $1-billion facility in Beaver Dam, among others anticipated. A number of other states still have moritoriums on nuclear power construction, including California, Maine, Oregon, Vermont, New Jersey, Rhode Island and Hawaii. Minnesota legislators are seeking to lift the state ban.

Time-to-Power

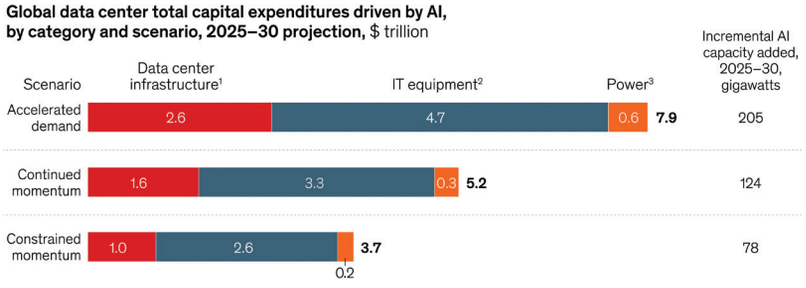
With more "time-to-power" constraints and longer grid connection waits, data center developers are adopting alternative energy strategies, particularly site "co-location" for direct power delivery from generating facilities, dubbed "behind the meter."

Nuclear plant owner Constellation's landmark 20-year deal last September with Microsoft to purchase the 835-MW output of the former Three Mile Island nuclear operating unit in Pennsylvania that now is restarting, and a June agreement with Meta for the 1.1-GW output from its Clinton nuclear facility in Illinois, both will involve power to data center facilities through regional transmission providers. But the company also seeks direct power supply to data centers that are co-located on power generation sites, it has said.

"Colocation is a very good data center [construction] option because it results in lower cost and faster speed," Mike Kramer, Constellation vice president for data economy strategy, told Reuters

CAPITAL INVESTMENTS TO SUPPORT AI-RELATED DATA CENTER CAPACITY DEMAND COULD RANGE FROM ABOUT $3 TRILLION TO $8 TRILLION BY 2030

Global data center total capital expenditures driven by AI, by category and scenario, 2025–30 projection, $ trillion



Note: Figures may not sum to totals, because of rounding.
[1]Excludes IT services and software (eg, operating system, data center infrastructure management), since they require relatively low capex compared with other components.
[2]Includes server, storage, and network infrastructure. IT capex also accounts for replacing AI accelerators every 4 years.
[3]Assumes $2.2 billion–$3.2 billion/gigawatt (including power generation and transmission cost) to account for a range of power generation scenarios (eg, fully powered by gas, a combination of gas power and storage, and solar) and regional cost differences. Distribution cost is neglected, as most AI centers are expected to be >50 megawatt scale and connected to a transmission grid.
Source: McKinsey Data Center Capex TAM Model; McKinsey Data Center Demand Model

SOURCE: MCKINSEY & COMPANY

event attendees. "You don't have to build out transmission; there's savings overall," he said.

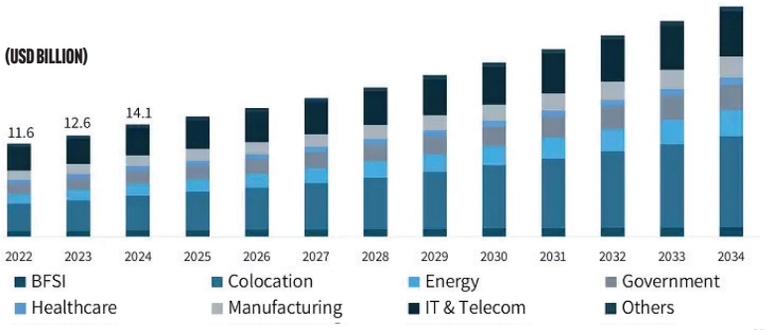
But policy disputes between the company and others involved in data center power, and utilities, grid operators and state and federal regulators over co-location issues related to supply logistics, needs, reliablity, security, costs and regulation are far from over. Constellation CEO Joe Dominguez in May "urged" the Federal Energy Regulatory Commission to quickly "clarity on the rules for behind-the-meter configurations and to provide latitude so there could be innovation here." He added: "Locating AI facilities proximate to large, clean and reliable power plants continues to make all the sense in the world. But critically, we do not need to have load colocated or behind the meter for us to achieve compelling pricing."

Ohio's state regulator on July 9 approved a landmark tariff structure requiring large new data center customers to pay for a minimum of 85% of their subscribed power use—regardless of actual consumption—for up to 12 years. The order also allows utility AEP Ohio to lift its halt moratorium on new service agreements in central Ohio. The order is "a well-balanced package that safeguards non-data center customers … while establishing a dependable and reasonable environment for data centers to continue to thrive in the state," said Public Utilities Commission of Ohio Chair Jenifer French.

Amazon, Google and industry trade group Data Center Coalition are set to debate Virginia utility Dominion at a Sept. 2 state hearing over its proposed new rate class for "high-load" customers, while some data center co-location issues could be on FERC's agenda also that month. Coalition President Josh Levi said the group is "committed to working corroboratively with local officials, policymakers and regulatory bodies at every level." But another group official urged regulators not to ultimately create something that ends up imposing a one-size-fits-all solution, or an inflexible solution."

Also raising more concern are water use risks of data centers, with the huge amounts extracted to cool equipment impacting residential and other customer drinking water supplies, especially in areas already at risk. "Data centers represent a new and rapidly growing demand on water resources," said Newsha Ajami, a water supply research expert at Lawrence Berkeley National Lab in California. "Many hyperscalers are investing to reduce their water footprint. But for some, water is an afterthought," she said, adding: "This is a fast-growing industry, so the challenge of balancing growth with sustainability is ongoing."

DATA CENTER POWER MARKET SIZE, BY APPLICATION, 2022-2034



SOURCE: GLOBAL MARKET INSIGHTS

Chinese developer launches first commercial underwater data center project, 2.5 GW, powered by offshore wind off Shanghai, with 198 server racks set to operate in September, local media report adding that a planned second phase could boost capacity to 24 GW. Photo courtesy of HiCloud

Looking to Innovation

Related to future strategies in sector innovation, Jacobs and Nvidia are partnering to use digital twins to design, simulate and optimize data centers in physically accurate virtual environments, to boost reliability and early-issue detection, said Jacobs Executive Vice President Koti Vadlamudi. The National Renewable Energy Laboratory's Chip-to-Grid Consortium had 300 public- and private-sector attendees at its June meeting to strategize data and power initiatives "to address the problems that one stakeholder alone can't solve," said agency manager Bill Livingood.

Meanwhile, the American Bureau of Shipping and naval architecture specialist Herbert Engineering have released new research on the feasibility of floating nuclear-powered data centers. "This design is a 'modern', efficient, high-density data center that can employ advanced server-cooling technologies. It is modular and remotely constructed, and comes with its own dedicated grid-independent, fully fueled and carbon-free power source," said the firms.

China also began construction in June of an ocean-based data center powered by offshore wind off Shanghai, a major AI hub in the country, according to a Chinese media report. Local developer HiCloud said it uses pipes to pump seawater through a radiator on the back of server racks to absorb heat and remove it, with a China government assessment saying the project uses at least 30% less electricity than on-land data centers.

In the U.S moving forward, data center-generated congestion will increase across the transmission system, "particularly at interconnection points not designed to handle multi-gigawatt-scale loads," said Pam Cannon, TRC vice president of marketing for the utilities sector. "Optimistically, this ... also opens the door to innovation on the demand side. Technologies that enable load flexibility, such as demand response, interruptible load programs and on-site energy storage, all help to address these

issues, at least until new generation and transmission projects are built." While none of these are silver bullets, "with coordination, investment and foresight, they represent an encouraging shift toward more dynamic, data-informed and collaborative approaches to solving these challenges," she added.

"The stakes are high," said consultant McKinsey & Co. in an April report. "Overinvesting in data center infrastructure risks stranding assets, while underinvesting means falling behind. AI is a radically evolving space." Noted developer Compass Datacenters CEO Chris Crosby in a recent opinion: "There's a gold rush or dot-com mentality in our space right now. Speculative developers are coming out of the woodwork to flip a piece of powered land for profit. Through close collaboration, risk sharing and embracing innovative solutions like the co-serve model, we can address the power challenges of the digital age."

Engineer Helps Data Centers to Cool Off

The expectation that exceedingly denser equipment will grow more energy intensive and hotter to meet AI's computing demands is leading to changes to traditional cooling designs for data centers. Shumate Engineering, an MEP firm specializing in data center construction, is moving to address water and energy demands concurrently with its Hybrid-Dry/Adiabatic-Cooling system that combines dry cooling methods and an adiabatic process that changes air pressure to make it cooler. The system received a patent in July.

The system's components—chillers, pumps and hybrid adiabatic coolers—are all commercially available. What's new is how they function together to create two streams of water within one closed loop—90° F water for direct-to-chip cooling, which circulates liquid through a cold plate to draw away excess heat from servers; and 68° F water for typical air cooling.

When dry cooling cannot achieve the necessary chilled water temperature, an adiabatic pad is moistened, "act[ing] as a pre-cooler to the outdoor air being pulled across the dry coil," says R. Stephen Spinazzola, Shumate director of mission critical services and system designer.

For a data center in Ashburn, Va., that approach equates to a power use effectiveness ratio averaging 1.095, says the firm. Ratios, calculated by dividing a site's overall energy use by its IT equipment energy use, are typically compared against a power use effectiveness value of 1.0. The larger the value, the less efficient. The Ashburn site system works "in full economizer mode" with no mechanical cooling about 55% of the year and uses water for about 9% of the year, says Spinazzola. In 2023, annual average power use effectiveness for U.S. data centers was 1.4, says a Lawrence Berkeley Lab report.

The firm says its system reduces by half the energy required of a typical data center cooling system. "By dramatically reducing both energy and water consumption, we're enabling data center operators to scale up AI and high-density workloads sustainably and cost-effectively," said Daren Shumate, firm founder and managing principal. "It's an engineering solution that meets the moment."

By Corinne Grinapol