04 May 2026

The sprint: Data centers are building a parallel energy system in the US

RaboResearch
rabobank.com/knowledge

Amit Mathrani
Executive Director, Senior Energy Transition Specialist
+1 929 697 5553

Summary

- Data centers are on track to add roughly 300TWh of incremental electricity demand in the US by 2030, equivalent to the combined annual consumption of New York and California. The five largest hyperscalers will spend between USD 745b and USD 775b on capital expenditure in 2026 alone, and the resulting power demand would represent the largest five-year surge to the US grid since the 1980s.
- Grid interconnection timelines of 36 to 84 months are structurally incompatible with data center build cycles of 12 to 24 months. That gap is pushing the sector behind the meter, with over 130GW of energy resources now proposed to serve planned data center projects across the US.
- Gas accounts for over 80 percent of the behind-the-meter pipeline, but the market is selecting for speed, not fuel. The technology stack runs from fuel cells at 90 days to heavy-duty turbines at three to seven years, and the mix reflects a deployment timeline, not an energy philosophy.
- Behind-the-meter power costs USD 100 to USD 165 per MWh versus USD 90 to USD 95 for grid alternatives. Although power is only 10 to 15 percent of a data center's total cost of ownership, it is the single binding constraint on whether a facility generates revenue.
- Hyperscalers are moving from buying power to owning it (for example, Alphabet's USD 4.75 billion acquisition of Intersect Power), raising the question of whether the sector is bridging to the grid or building a permanent alternative.
- The race to power AI data centers is a short-distance sprint to close an immediate infrastructure gap. The marathon that follows will test whether behind-the-meter solutions can endure economically, technically, and regulatorily.

The AI arms race has a power problem

The five largest cloud and artificial intelligence (AI) infrastructure companies plan to spend approximately USD 745b and USD 775b on capital expenditure in 2026, with three of the five raising their initial 2026 guidance during first quarter reporting. The collective figure is up from approximately USD 380b to USD 440b in 2025. Amazon appears to be leading the pack by spending approximately USD 200b, predominantly for Amazon Web Services (AWS).[1] Alphabet raised its guidance to USD 180b to USD 190b, up from USD 175b and USD 185b, citing the closing of its Intersect Power acquisition.[2] Meta raised its guidance to USD 125b to USD 145b, up from USD 115b to USD 135b, citing higher component pricing and additional data center costs to support future-yea capacity. [3] Microsoft now expects approximately USD 190b in calendar year 2026, including about USD 25b attributable to higher component pricing. [4] Oracle has allocated

---

[1] Amazon Q1 2026 earnings release and conference call, April 29, 2026

[2] Alphabet Q1 2026 earnings release and conference call, April 29, 2026

[3] Meta Platforms Q1 2026 earning release and conference call, April 29, 2026

[4] Microsoft Q3 FY2026 earning release and conference call, April 29, 2026

USD 50b.[5] Each reported in their most recent earnings call that demand for AI compute continues to outpace available capacity, and that they would deploy even more capital if construction timelines and equipment availability permitted.

But one constraint stands apart: power. A data center facility can be constructed in 12 to 24 months,[6] but the grid infrastructure to deliver electricity to the facility takes three to seven years. The companies setting the pace for AI innovation are concluding that they cannot wait for the grid to catch up. They are taking power procurement into their own hands, and in doing so, constructing a parallel energy system.

US electricity demand breaks two decades of stagnation

US electricity consumption is accelerating after nearly two decades of flat growth. Total consumption reached approximately ~4,200TWh in 2025, the highest on record, with retail sales increasing across all three end-use sectors: residential, commercial, and industrial.[7] As we outlined in a recent RaboResearch analysis, this breakout is being driven by three converging forces: data centers, the reshoring of advanced manufacturing, and electrification. By 2030, total consumption could climb by as much as 20 percent from 2025 levels, with system peak demand rising from roughly 760GW in 2025 to between 850GW and 930GW by the end of the decade.[8]

Data centers are the most significant near-term driver. Consumption reached an estimated 176TWh in 2025 and is forecast to grow to between 420TWh and 728TWh by 2030, depending on assumptions about interconnection success and the pace of AI adoption. Including the incremental demand of roughly 300TWh through 2030, that is equivalent to the combined annual electricity consumption of New York and California.

Meeting this demand will require a generation buildout of a magnitude not seen in decades. Industry forecasters point in the same direction (see figure 1). Across these estimates, the five-year increase implied through 2030 would be the largest observed since the 1980s.

---

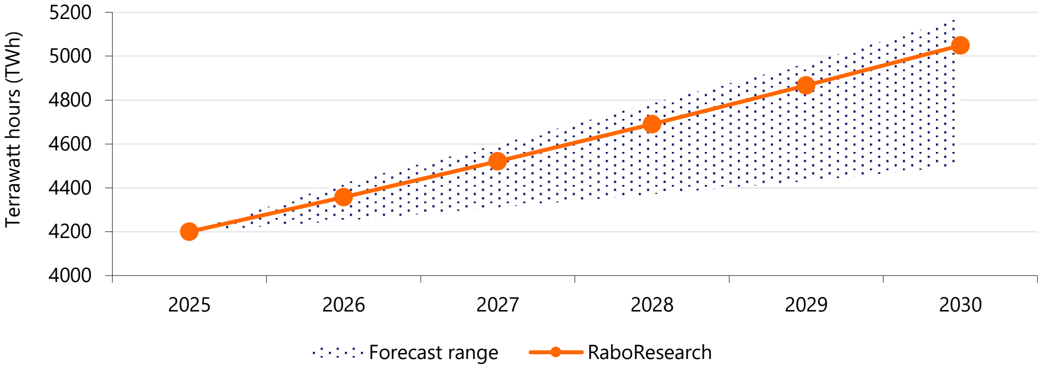
[5] Oracle Q3 FY2026 earnings release, March 10, 2026

[6] S&P Global Market Intelligence, 451 Research "2026 US Data Centers & Energy Report."

[7] Reuters, "US power use beat record highs as AI use surges", April 7, 2026

[8] RaboResearch, "The US scrambles to meet surging power demand by 2030." Amit Mathrani, January 28, 2026

Figure 1: US electricity demand forecasts through 2030 (TWh)

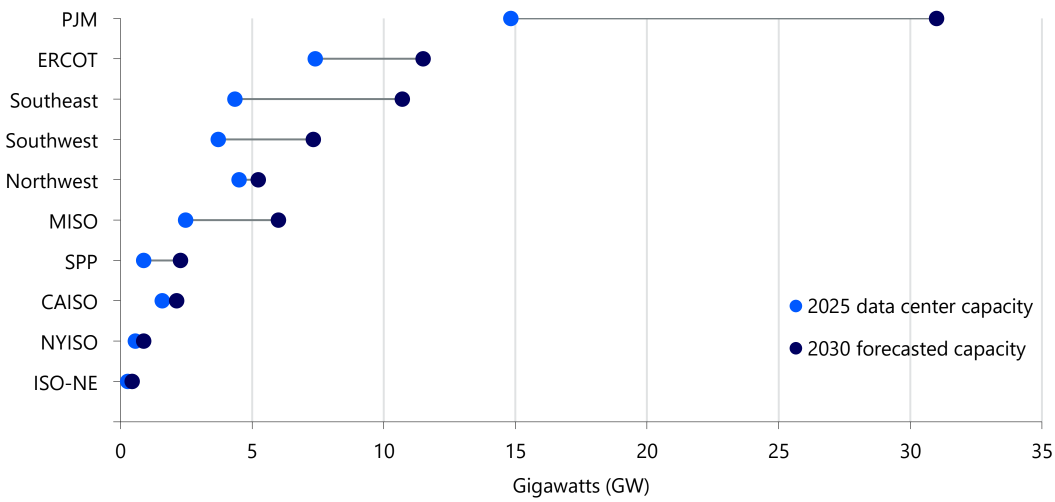


Source: US Energy Information Administration (EIA), S&P Global Commodity Insights, Edison Electric Institute (EEI), BloombergNEF (BNEF), McKinsey, RaboResearch 2026

This is not a temporary demand spike. While traditional cloud and enterprise workloads tend to run at relatively stable utilization rates, AI training clusters produce highly variable load profiles with sharp power ramps and dips at megawatt scale, and inference workloads fluctuate with user demand throughout the day. What is consistent across all data center types is the requirement for continuous, uninterrupted power availability, even if the actual draw varies.

The tension between demand and existing infrastructure is most visible in the electricity markets absorbing the major share of this load. Data center capacity in the US reached approximately 48GW across major electricity markets in 2025 and is forecast to grow to roughly 78GW by 2030 (see figure 2). PJM, ERCOT, and MISO collectively account for more than 60 percent of the projected national total. ERCOT's faster permitting environment has attracted rapid growth, while PJM and MISO face the most severe interconnection backlogs. Demand is clustering faster than infrastructure can expand, turning a national growth story into a set of regional capacity crises.

Figure 2: US data center capacity by electricity market, 2025-2030



Source: BNEF, RaboResearch 2026

Grid bottlenecks are forcing a strategic bifurcation

The US power grid was engineered for incremental load growth. It was not designed for hundreds of developers simultaneously requesting gigawatt-scale connections. According to Lawrence Berkeley National Lab, the average time from interconnection queue entry to commercial operation has stretched to approximately five years nationally, up from under two years in 2008.[9] In California, processing times exceed seven years. AEP Ohio, the American Electric Power subsidiary serving much of central Ohio, informed PJM Interconnection's Load Analysis Subcommittee in late 2025 that none of the 13GW of new data center load in its territory can be reliably served until Q4 2031.[10]

Renewables are deploying faster than any other resource

The US added 27.2GW of utility-scale solar and a record 15GW of battery storage in 2025, with solar and storage together accounting for 79 percent of all new electricity capacity.[11] But for data centers that require firm, uninterruptible power, intermittent resources alone do not close the gap. Orchestrating a solar-plus-storage system to deliver the 99.99 percent availability that a hyperscale facility demands is technically complex and capital-intensive. Compressed timelines for Income Tax Credits (ITC) and Production Tax Credits (PTC) qualification under the One Big Beautiful Bill Act (OBBBA), combined with Foreign Entity of Concern (FEOC) compliance requirements, are adding further uncertainty to the renewable pipeline beyond 2027. Renewables will play an important role in the long-term energy mix for data centers, but they are not a substitute for firm generation on the timeline the market requires.

There are supply constraints on the side of dispatchable energy solutions as well. Large gas-fired generation remains limited by long equipment lead times and sharply higher capital costs. The scarcity is manifesting in market prices. PJM's 2027-2028 capacity auction, held in December 2025, fell 6,623MW short of its 20 percent installed reserve margin target, the first time the entire RTO has failed to procure enough capacity to meet the one-event-in-ten-year reliability standard.[12] Prices hit the FERC-approved cap of USD 333.44 per megawatt-day, and PJM estimated that without the cap, the clearing price would have approached USD 530 per megawatt-day. Two years earlier, the same auction cleared at USD 28.92.

A strategic bifurcation is underway

One path is to wait for traditional grid service, accepting multi-year timelines and rising interconnection costs. The other is to take direct control by building or contracting generation behind the meter (BTM). As an industry where six months of delay translates into billions in lost revenue from idle Graphics Processing Units (GPU) capacity, the data center sector is not abandoning the grid, but is no longer willing to depend on it exclusively.

Gas dominates the pipeline of BTM power solutions

The US pipeline of data center projects exceeds 160GW of planned data center capacity, with more than 130GW of colocated energy resources announced to serve them. Gas-based generation accounts for more than 80 percent of that announced total. Not all of this announced capacity will move forward on the current timeline. After accounting for permitting status, firm equipment orders, secured fuel contracts, and realistic project phasing, S&P Global Energy

---

[9] Lawrence Berkeley National Lab, Queues, Costs, and Timelines for US Electricity Generation and Storage, 2025

[10] AEP Ohio filing to PJM Load Analysis Subcommittee, 2025

[11] EIA, New US electric generating capacity expected to reach a record high in 2026, February 2026

[12] Utility Dive, PJM capacity prices hit record high as grid operator falls short of reliability target, December 18, 2025

expects approximately 22GW to 32GW of on-site generation to be deployed by 2030, equal to roughly one-quarter of new data center capacity through the end of the decade.[13]

What matters most is not simply that gas dominates the pipeline, but why it dominates. Gas is not a single category. Heavy-duty turbines and combined-cycle plants carry lead times of 36 to 84 months, with new orders now facing wait times of up to seven years. These technologies will not energize new capacity in the near term. The equipment set to actually clear first is a different subset of the gas fleet: reciprocating engines, aeroderivatives, and smaller frame turbines that deploy in 12 to 24 months. The market is not selecting the most elegant long-term solution; it is selecting the fastest workable one. In an industry where months of delay can strand billions of dollars of IT hardware and deferred revenue, speed to energization has become more valuable than long-run efficiency. The technologies gaining share today are not necessarily those best suited to run a campus for decades. They are the ones that can be permitted, delivered, installed, and financed on the timeline AI infrastructure demands (see figure 3).

This is producing a three-tier technology stack. Fuel cells can be deployed in as little as 90 days, making them the only near-term option for immediate energization. Reciprocating gas engines, aeroderivatives, renewables, and solar-plus-storage combinations deploy in 12 to 36 months, forming the medium-term backbone of the current buildout. Heavy-duty turbines, simple cycle gas plants, combined-cycle plants, and nuclear energy sit beyond three years (see figure 3).

Three broad deployment strategies are taking shape alongside this technology stack:

- **Colocation with existing or planned power plants**: Meta's Prometheus AI supercluster in New Albany, Ohio, is served by dedicated gas plants supplying 400MW of BTM baseload power. In Louisiana, Meta partnered with Entergy to develop 2.5GW of BTM gas plants for the Hyperion data center under a 15-year energy contract.
- **Self-installation of generation**: xAI deployed up to 35 mobile gas turbines to its Colossus data center in Memphis. Caterpillar is delivering hundreds of 2.5 MW gas engines to gigawatt-scale campuses across the country.[14]
- **Purpose-built microgrids and power islands**: Energy Abundance Development Corp. announced Data City, Texas, a 5GW fully off-grid data center hub near Laredo, Texas, where the data center and the power plant are conceived as a single integrated development.[15]

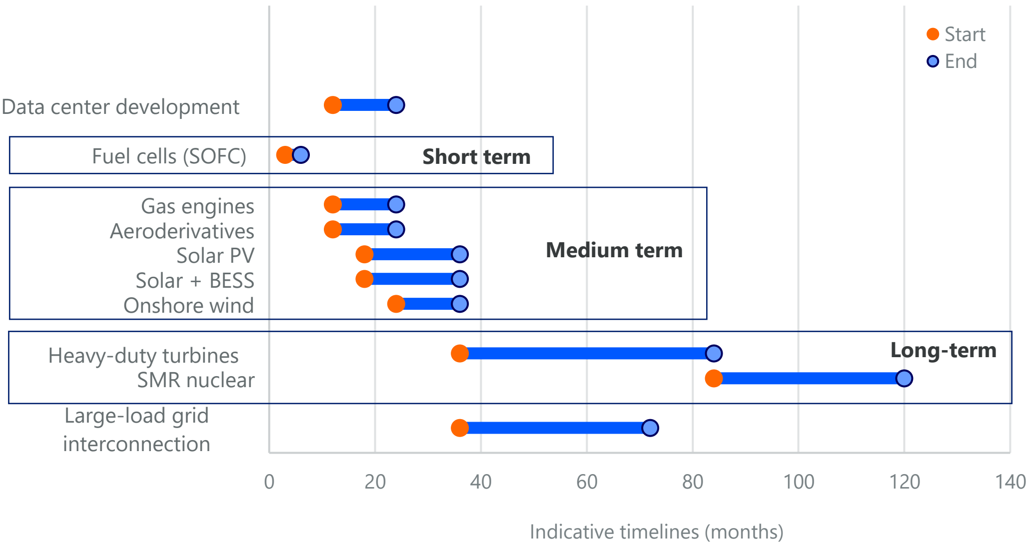
---

[13] S&P Global Energy, Colocating Data Centers with Energy Resources Help Plug the Power Supply-Demand Gap in North America, March 2026

[14] Yahoo News. Data Centers Are Turning to Gas Generators for Prime Power to Eliminate Long Lead Times for Grid Connections, October 28, 2025

[15] PR Newswire, Energy Abundance Announces Data City, Texas - The World's Largest Behind-the-Meter Data Center Hub Powered by 100% 24/7 Green Energy, March 21, 2025

Figure 3: Power solutions timeline by deployment horizon (months from firm order to commercial operation)



Source: BNEF, Lazard LCOE+, EIA, public disclosures

In the short term, fuel cells may offer the fastest path to energization

Bloom Energy's solid oxide fuel cells represent only 0.8GW of announced data center capacity, but they occupy a unique position in the technology stack. Because they generate electricity electrochemically rather than through combustion, they produce minimal noise and emissions, reducing the air permitting burden that slows gas-fired projects in more restrictive jurisdictions. Oracle has cited deployment timelines of roughly 90 days, far faster than any conventional alternative. Capital costs are higher, at approximately USD 3,000 to USD 4,000 per kW (see table 1), and manufacturing scale remains limited. But where permitting, not fuel supply, is the binding constraint, fuel cells can be the fastest viable route to energization.

Renewables, engines, and aeroderivatives are the bridge over the medium term

Renewables and storage strengthen the transition path

Solar PV and onshore wind deploy on 18-to-36-month timelines at costs of USD 1,150 to USD 1,600 per kW and USD 1,900 to USD 2,300 per kW respectively, making them among the lowest-cost generation resources available (see table 1). When paired with battery storage in a solar-plus-Battery Energy Storage Systems (BESS) configuration, these systems can also provide dispatchable capacity, manage peak grid withdrawals, and support microgrid designs. Solar, wind, and solar-plus-storage combinations account for roughly 22GW of the colocated pipeline.[16] Alphabet's $4.75 billion acquisition of Intersect Power, announced in December 2025 and closed in March 2026, illustrates the evolving logic. The transaction includes the co-located data center and power campus already under construction in Haskell County, Texas, along with multiple gigawatts of additional energy and data center projects in Intersect's pipeline. The value is not just the

---

[16] S&P Global Commodity Insights, 451 Research, 2026 US Data Centers & Energy Report

renewable electrons. It is control of a site, an interconnection pathway, and a platform that can support a broader portfolio of power resources over time.

Gas engines lead on speed and cost

Gas-fired reciprocating engines have become the clearest embodiment of the market's speed-over-efficiency logic. Manufacturers such as Caterpillar, Wartsila, and INNIO Jenbacher account for 18.3GW of announced capacity. Units typically range from 2.5MW to 35MW, with all-in capital costs of roughly USD 1,700 to USD 2,000 per kW and procurement timelines of 12 to 24 months (see table 1). The Monarch Data Center in West Virginia, a planned 2GW campus with Microsoft as lead tenant, ordered more than a gigawatt of Caterpillar engine capacity.[17] Engines are modular, scalable, and deployable quickly. Once a permanent grid connection is established, they can revert to backup duty, displacing diesel capacity the site would need anyway. The trade-off is operational complexity. A campus built around 2.5MW units can require hundreds of engines, meaning maintenance and long-run efficiency will be meaningful constraints.

Aeroderivatives balance scale with flexibility

Aeroderivative turbines occupy the middle ground between modular engines and large central-station gas plants. Adapted from aerospace platforms, they typically offer 25MW to 50MW per unit, strong ramping capability, and lead times of one to two years. Their flexibility makes them well suited to large campuses that need more scale than engines can provide without waiting for the full heavy-duty supply chain to loosen. Crusoe's Abilene site for Oracle and OpenAI illustrates this approach, combining GE Vernova LM2500XPRESS aeroderivatives with Caterpillar Titan 350 turbines to create a fleet-based on-site generation strategy.

For the long term, firm power and zero-carbon generation

Heavy-duty turbines and combined-cycle plants position for scale

Heavy-duty gas turbines and combined-cycle plants remain the more efficient long-run answer where developers can secure equipment and accept longer lead times. Heavy-duty gas turbines from GE Vernova, Siemens Energy, and Mitsubishi Heavy Industries make up the largest announced category in the pipeline at 27.2GW, with another 14.2GW tied to combined-cycle capacity. But they are poorly suited to the immediacy of the current AI build cycle. GE Vernova ended 2025 with an 83GW gas turbine backlog stretching into 2029, while capital costs for new combined-cycle plants have risen to roughly USD 2.3b to USD 2.5b per GW.[18] [19]

The mismatch between what the market wants and what the heavy-duty supply chain can deliver is creating room for new entrants. South Korea's Doosan Enerbility secured its first US data center turbine orders in late 2025, supplying DGT6-class turbines to xAI for the Colossus expansion, with a total of 12 units now under contract for US delivery. When incumbent OEMs cannot supply enough equipment on time, the market begins to widen the competitive field.

---

[17] Power Engineering Factor This, Caterpillar engines to support 2 GW of onsite power at West Virginia data center campus tied to Microsoft, NVIDIA

[18] GE Vernova expects to end 2025 with an 80GW gas turbine backlog that stretches into 2029

[19] S&P Global Market Intelligence, US gas-fired turbine wait times as much as seven years; costs up sharply, May 2025

Nuclear remains a long-term horizon option

Nuclear remains the longest-term horizon answer in the stack. Small modular reactors attract outsized attention because they appear to offer what hyperscalers ultimately want: firm, zero-carbon power at scale. But the delivery timeline is still measured in the next decade, not the next procurement cycle. Total small modular reactor (SMR) capacity expected before 2033 remains under 2GW. TerraPower has secured a key permit for its Natrium project in Wyoming, but commercial operation is still expected no sooner than the early 2030s. The more immediate nuclear activity is occurring through long-dated PPAs tied to existing plants, such as Meta's agreement with Constellation Energy and Oklo's framework with Switch.

Table 1: Behind-the-meter technology cost and deployment comparisons

Technology	*All-in-costs (USD/kw)*	*Lead time (months*)*	*Pipeline (GW)*	*Key advantage*
Fuel-cells	3,000 – 4,000	3 – 6	~1	No combustion, fast permitting
Solar	1,150 – 1,600	18 – 36	9	Lowest cost, fast deployment
Wind	1,900 – 2,300	24 – 36	2	Higher capacity factor than solar
Battery storage	500 – 1,300	16 – 24	11	Peak shaving, flexibility, backup
Gas engines	1,500 – 2,900**	12 – 24	18.3	Speed, cost, modularity
Aeroderivatives	2,000 – 2,500	12 – 24	5.3	Scale, load-flexibility
Industrial engines	2,000 – 2,300	18 – 36	9.3	Mid-scale, firm power
Heavy-duty turbines (i.e., CCGT)	2,300 – 2,500	36 – 84	41.4	Efficiency at scale, lowest heat rate
SMR nuclear***	6,000 – 10,000	84 – 120	2	Zero emissions, baseload
Turbine****	N/A	N/A	20	Equipment class not specified
Other*****	N/A	N/A	13.4	Includes mixed-use and unknown

Note: Months* - months from firm order to commercial operation
Gas engines** - Reflect ~2.5MW-class natural gas reciprocating engines for BTM, island capable installations, range includes genset plus typical inside-fence balance of plant (BOP), switchgear/paralleling controls, enclosure / acoustics, civil / EPC / commissioning, and standard permitting; excludes fuel and major offsite utility upgrades. OEM base scope typically starts with engine / alternator / controls, and BOP added depending on site and emissions requirements.
SMR Nuclear*** - Small Modular Reactor. Pre-commercial technology in the US as of early 2026. Cost and lead-time ranges reflect first-of-a-kind deployment projections. Total US SMR capacity expected to reach commercial operation before 2033 remains under 2GW, with earliest commercial operation estimated in the early 2030s.
Turbines**** - Approximately 20GW of announced turbine-based capacity has not disclosed equipment class (heavy-duty, industrial, or aeroderivative).
Other***** - An additional 13.4GW reflects mixed-use projects or projects where no equipment information is publicly available.
Source: BloombergNEF, Bloom Energy, Lazard LCOE+ 2025, GE Vernova, EIA, public company disclosures, and RaboResearch 2026

The pattern in table 1 is instructive. The technologies deploying fastest – fuel cells, solar, storage, and gas engines – carry higher per-kW capital costs relative to their thermal efficiency or require pairing with firm generation to meet availability requirements. The most efficient long-run options – combined-cycle plants and nuclear – have multi-year lead times that make them impractical for near-term deployment. The market is paying a premium for speed. RaboResearch estimates that BTM gas generation for data centers costs approximately USD 100 to USD 165 per MWh depending on overbuild assumptions, bridge duration, and gas price, compared to USD 90 to

USD 95 per MWh for grid-connected alternatives.[20] As long as the grid timeline exceeds the AI deployment timeline, this cost structure will persist.

Speed to power is reshaping ownership of energy infrastructure

More than 20GW of initial-phase colocated capacity is already in advanced development, with equipment procured, fuel supply secured, and permits approved. Power represents only 10 to 15 percent of a data center's total cost of ownership, yet it has become the single binding constraint on whether a facility generates revenue. That BTM premium is manageable when idle GPU clusters represent USD 3b to USD 5b in annual depreciation and lost revenue. This calculus is driving a strategic shift that makes procuring power infrastructure a core competency. Alphabet's USD 4.75b acquisition of Intersect Power in December 2025 included multiple gigawatts of energy and data center projects, signalling that hyperscalers are moving from buying power to owning the power pipeline.

The market is sprinting to bridge the gap between AI compute demand and existing infrastructure. Near-to-medium term technologies can energize a data center within the window it takes to build one, and the bridge to grid connection appears viable on a five-to-seven-year horizon. The harder questions sit with the 41GW of heavy-duty and combined-cycle capacity being built for longer durations, where the investment thesis is less about bridging to the grid and more about whether the relationship between data centers and centralized power has permanently changed. Ultimately, the marathon will be determined by three uncertainties: the longevity of the equipment, the regulatory viability of sustained BTM operation, and the bankability of generation assets that may outgrow their original purpose well before the end of their technical life.

---

[20] RaboResearch estimate. Derived from publicly available inputs: Lazard LCOE+ 2025 gas peaker range (USD 149 to USD 251/MWh), BloombergNEF equipment cost data (USD 1,500 to USD 2,000/kW for reciprocating engines), S&P Global Commodity Insights natural gas price outlook (USD 4.00 to USD 4.50/MMBtu Henry Hub near-term), EIA heat rate data for reciprocating engines (8,000-9,000 BTU/kWh), and industry-standard assumptions for overbuild (40-65%), 10-year amortization, and 75-90% capacity factor.

Imprint

RaboResearch
rabobank.com/knowledge

Amit Mathrani	Executive Director, Senior Energy Transition Specialist	amit.mathrani@rabobank.com +1 929 697 5553

Disclaimer

This publication is issued by Coöperatieve Rabobank U.A., registered in Amsterdam, The Netherlands, and/or any one or more of its affiliates and related bodies corporate (jointly and individually: "**Rabobank**"). Coöperatieve Rabobank U.A. is authorised and regulated by De Nederlandsche Bank and the Netherlands Authority for the Financial Markets. Rabobank London Branch is authorised by the Prudential Regulation Authority ("**PRA**") and subject to regulation by the Financial Conduct Authority and limited regulation by the PRA. Details about the extent of our regulation by the PRA are available from us on request. Registered in England and Wales No. BR002630. An overview of all locations from where Rabobank issues research publications and the (other) relevant local regulators can be found here: https://www.rabobank.com/knowledge/raboresearch-locations

The information and opinions contained in this document are indicative and for discussion purposes only. No rights may be derived from any transactions described and/or commercial ideas contained in this document. This document is for information purposes only and is not, and should not be construed as, an offer, invitation or recommendation. This document shall not form the basis of, or cannot be relied upon in connection with, any contract or commitment by Rabobank to enter into any agreement or transaction. The contents of this publication are general in nature and do not take into account your personal objectives, financial situation or needs. The information in this document is not intended, and should not be understood, as an advice (including, without limitation, an advice within the meaning of article 1:1 and article 4:23 of the Dutch Financial Supervision Act). You should consider the appropriateness of the information and statements having regard to your specific circumstances and obtain financial, legal and/or tax advice as appropriate. This document is based on public information. The information and opinions contained in this document have been compiled or arrived at from sources believed to be reliable, but no representation or warranty, express or implied, is made as to their accuracy, completeness or correctness.

The information and statements herein are made in good faith and are only valid as at the date of publication of this document or marketing communication. Any opinions, forecasts or estimates herein constitute a judgement of Rabobank as at the date of this document, and there can be no assurance that future results or events will be consistent with any such opinions, forecasts or estimates. All opinions expressed in this document are subject to change without notice. To the extent permitted by law Rabobank does not accept any liability whatsoever for any loss or damage howsoever arising from any use of this document or its contents or otherwise arising in connection therewith.

This document may not be reproduced, distributed or published, in whole or in part, for any purpose, except with the prior written consent of Rabobank. The distribution of this document may be restricted by law in certain jurisdictions and recipients of this document should inform themselves about, and observe any such restrictions.

A summary of the methodologies used by Rabobank can be found on our website.