COMPASS
Collaborative Optimization & Management of Power Allocation – Surface & Subsurface strategies

Bureau of Economic Geology

The University of Texas at Austin
Jackson School of Geosciences

The University of Texas at Austin
Center for Electromechanics
Cockrell School of Engineering

ON-SITE POWER GENERATION TECHNOLOGIES RESHAPING THE FUTURE OF DATA CENTERS

Prepared by COMPASS Research Consortium

A White Paper on the prospects, future, and comparison of different generation options, including grid connected, isolated, and Behind the Meter systems for data centers in Texas.

Contact Information

compass@beg.utexas.edu

10611 Exploration Way
Austin, TX 78758

Scan QR Code for More Information

COMPASS
Collaborative Optimization & Management of Power Allocation – Surface & Subsurface strategies

Bureau of Economic Geology

The University of Texas at Austin
Jackson School of Geosciences

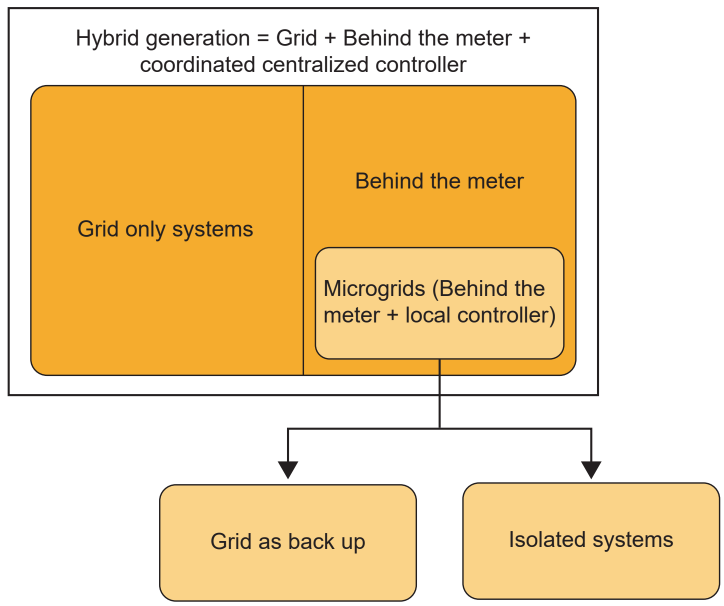
The University of Texas at Austin
Center for Electromechanics
Cockrell School of Engineering

Powering AI Data Center

The rapid growth of AI, cloud services, and digitalization is fundamentally reshaping the computing landscape and driving unprecedented increases in power demand across the data center sector.

Our white paper highlights the power requirements for data centers. We are also examining alternate power delivery solutions for data centers, comparing with grid connected systems.

Power generation options for data centers



Inside the Report

1. Driving of rising power demand of data centers
2. On-site generation options to power AI data centers
3. Texas power landscape and ERCOT grid constraints

Prepared by COMPASS Research Consortium

Contact Information

compass@beg.utexas.edu

10611 Exploration Way
Austin, TX 78758

Scan QR Code for More Information

On-site Power Generation Technologies Reshaping the Future of Data Centers

About This Report – A Message from COMPASS

At COMPASS, we are committed to advancing data-driven insights that support resilient infrastructure and sustainable resource management across Texas.

Rapid growth in artificial intelligence (AI), cloud computing, and high-performance digital services is driving a fundamental shift in electricity demand from large loads, particularly data centers. These facilities concentrate power consumption at a scale and speed that challenge traditional approaches to generation planning, transmission development and grid operations. As a result, large-load customers and data center developers are actively exploring alternative power supply solutions to reliably serve their operations.

Texas illustrates these dynamics clearly. The state has emerged as a major hub for data center development due to its favorable regulatory and economic environment, abundant energy resources, and competitive power markets. At the same time, ERCOT's system characteristics, interconnection timelines, and evolving regulatory requirements are shaping how large loads are planned and served. These conditions are prompting data center developers to reconsider conventional grid-dependent models and explore the need for innovative and diversified power strategies.

This white paper examines the growth in electricity demand from AI-driven data centers and reviews a range of power generation and supply options, including grid-connected systems, behind-the-meter generation, microgrids, and hybrid configurations. The analysis focuses on how these approaches interact with the Texas energy landscape, ERCOT operational realities, and emerging policy frameworks. Through this report, COMPASS reinforces its commitment to supporting the development of resilient, reliable, and sustainable digital infrastructure that balances industry growth with the long-term needs of Texas communities.

Prepared by COMPASS Research Affiliates Program at the University of Texas at Austin
Contact: compass@beg.utexas.edu

COMPASS
Collaborative Optimization & Management of Power Allocation – Surface & Subsurface strategies

List of Abbreviations

AALC	Air Assisted Liquid Cooling
AC	Alternating Current
AEO	Annual Energy Outlook
AI	Artificial Intelligence
BESS	Battery Energy Storage Systems
BTM	Behind the meter
BYOP	Bring Your Own Power
CAPEX	Capital Expenditure
CCGT	Combined-Cycle Natural Gas Turbine
CDN	Content Delivery Networks
CHP	Combined Heat and Power
CUE	Carbon Usage Effectiveness
DC	Direct Current
DG	Diesel Generators
DLC	Direct Liquid Cooling
DR	Demand Response
ERC	Emission Reduction Credit Program
ERCOT	Electric Reliability Council of Texas
ESG	Environmental Social and Governance
FEMP	Federal Energy Management Program
FFR	Fast Frequency Response
FFSS	Firm Fuel Supply Service
FFSSR	Firm Fuel Supply Service Resource
FFT	Fast Fourier Transform
FPGA	Field-Programmable Gate Arrays
GHG	Greenhouse Gas
GPU	Graphics Processing Unit
GW	Giga Watt
HBM	High-bandwidth memory
HVAC	Heating, Ventilation, and Air Conditioning
ITAD	IT Asset Disposition
LEL	Large Electronic Loads
LOLP	Loss of Load Probability
NSR	New Source Review
OPEX	Operating Expenditure
PBR	Permits By Rule
PPA	Power Purchase Agreement
PUE	Power Usage Effectiveness
PV	Photovoltaic panel (solar)
ROI	Return on Investment
TCEQ	Texas Commission on Environmental Quality
TPU	Tensor Processing Unit
TWh	Terawatt Hour
UPS	Uninterruptible Power Supplies
VPP	Virtual Power Plants
WUE	Water Usage Effectiveness

Glossary

ERCOT (Electric Reliability Council of Texas)
Independent system operator for the Texas region

Colocation data center
Facility where a business can rent space, power, cooling, and security to house its own servers and other computing hardware.

Hyperscale data center
Massive facility with power ratings ranging from 50MW to few GW, and high rack power densities, designed for extreme scalability to handle large-scale, high-demand workloads like cloud computing, big data analytics, and AI.

Enterprise data center
Dedicated facility that houses an organization's IT infrastructure, including servers, storage systems, and networking equipment, to support its data processing and storage needs.

Tensor Processing Unit (TSU)
Specialized ASIC (Application-Specific Integrated Circuit) designed by Google to accelerate machine learning workloads, especially deep neural networks, by efficiently handling the massive matrix multiplications central to these tasks.

Graphics Processing Unit (GPU)
Specialized electronic circuit designed to accelerate the creation and rendering of images, videos, and animations.

Field Programmable Gate Arrays (FPGA)
A customizable integrated circuit (IC) that users can program and reconfigure after manufacturing to create custom digital hardware.

Behind the meter (BTM) on-site generation
Customer owned generation located on the customer side of the utility meter.

Microgrid
A localized, self-contained electrical grid that can operate independently from the main power grid or be connected to it. They can be part of the BTM generation when located on the customers side of the utility meter and primarily serve the electrical loads of the customer.

Hybrid generation
Using power from different sources including the centralized grid, on-site generation using natural gas, nearby solar and wind facilities, combined heat and power (CHP), energy storage.

Power Utilization Effectiveness (PUE)
Ratio of total facility energy to IT equipment energy.

Water Utilization Effectiveness (WUE)

Annual site water usage in liters to the IT equipment energy usage in kilowatt-hours (kWh) during the same period.

Carbon Utilization Effectiveness (CUE)

Ratio of total carbon emissions to total IT energy consumption

Data center uptime (%)

The percentage of time a data center's systems and services are operational and accessible to users, measuring its reliability and availability. It is calculated as a ratio of total operational time to total time in a given period, often measured in "nines."

Executive Summary

Demand for data centers has accelerated dramatically over the past decade, driven by generative artificial intelligence (AI), cloud computing, high-performance digital services, and edge applications that process data closer to its source for real-time insights. This expansion has shifted the engineering focus from compute and storage capacity to securing reliable and high-quality power at scale. As of 2024, U.S. data centers represent approximately 54 GW of operational capacity and consume 233 TWh annually—roughly 44% of global data center electricity use [1]. This demand is projected to rise to 325–580 TWh (7%-12% of total U.S. electricity consumption in 2024) by 2028, translating to 74–132 GW of peak power requirements, assuming an average capacity utilization rate of 50% [2]. These projections reflect increasing Graphics Processing Unit (GPU) deployment, growing cooling loads, and more intensive AI workloads.

Managing the growing electricity demand from large AI data centers is a challenge given the current grid infrastructure in the U.S. In most states, the electricity system is planned under legacy assumptions about demand growth, load diversity, and generation mix. Large AI data centers, with rapid scaling of highly concentrated, inflexible loads, place strain on a system not designed for fluctuating (or spiky) characteristics.

Texas is emerging as a leading data center hub, supported by favorable siting conditions, business incentives, and a unique market structure under the Electric Reliability Council of Texas (ERCOT). In 2025, Virginia, North Dakota, Nebraska, Iowa, Oregon, Wyoming, Nevada, and Arizona, data center power consumption comprised between 7.5% and 25% of their total electricity consumption. We have an opportunity to learn from the experiences in these states to allow for efficient growth in Texas.

In Texas, ERCOT matches infrastructure to load demand with an objective to ensure grid reliability. With the rapid demand for data centers, Texas is now seeing load requests that exceed historical transmission and generation capacity margins. With the growth in demand, ERCOT's generation mix continues to shift—wind and solar provided 35% of supply in 2024. This changing landscape provides an opportunity for new strategies to deliver fast and resilient power.

On-site generation and behind-the-meter solutions, including microgrids, and distributed energy resources, are commonly adopted approaches to mitigate reliability concerns and accelerate time to power. These generation systems demonstrate that it is possible to improve power quality and operational resilience enabling partial or full separation from the centralized grid.

This white paper presents the current state of data center development in Texas and provides a structured review of power supply strategies, including on-site and hybrid generation models. Key insights include:

- Power availability is a long recognized dominant driver of site selection. It is less costly to transmit information than electrons. As AI and high-density computing accelerate load growth, developers increasingly choose locations based on the ability to secure reliable power at scale.

- Recent Texas policy reshape interconnection requirements, cost-allocation for transmission upgrades, and load-shedding rules for facilities above 75 MW. On-site and hybrid systems can help developers reduce infrastructure investment uncertainty and improve compliance with Texas Senate Bill 6, passed in the 89th legislative session.

- On-site systems improve resilience but bring new environmental and community considerations. Localized power generation can reduce strain on ERCOT and support grid stability, but also increases permitting, emissions, land use, water availability, and local impact challenges that must be integrated into long-term planning.

- Together, these findings stress that the next phase of data center growth—particularly in Texas—will depend not only on IT investment, but on innovative power system architectures capable of delivering reliable, scalable, cost-effective, and environmentally responsible energy at the grid.

Chapter 1
Rising demand for data centers

The accelerating expansion of the data center industry is placing significant upward pressure on electricity demand, thereby necessitating considerable additions to electric-generation capacity and grid infrastructure. Current and anticipated growth in data-center electricity consumption in the United States ([2], [3]) highlights the sector's increasing significance as a primary contributor to future electricity demand.

Texas plays a central role in this trajectory. The state leads the nation in energy production—accounting for nearly one-quarter of total U.S. primary energy output and approximately 17% of national renewable generation [4]—and thus offers highly favorable conditions for continued data-center siting and expansion. In addition, Texas has historically attracted commercial and industrial growth due to its supportive policy environment, including the absence of a state income tax, ample land availability, and comparatively streamlined regulatory processes. These characteristics have positioned Texas as the second-largest data-center market in the United States, currently hosting approximately 413 facilities [5] with an estimated connected load of 7.5 GW.

In 2023, the ERCOT supplied roughly 22 TWh of electricity to data centers, representing approximately 4.6% of statewide electricity consumption [3]. Projections across recent studies indicate that this load is likely to increase sharply over the coming decade, potentially exceeding 10% of total electricity demand by 2030 (Fig. 1). Such growth would effectively double the sector's current share and emphasize the increasingly prominent role of data centers as a rapidly expanding load class with consequential implications for power-system planning, resource adequacy, and transmission infrastructure.

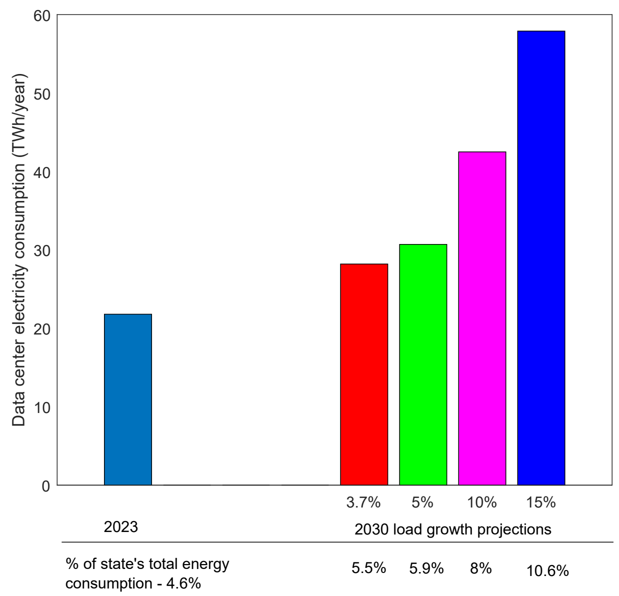


Fig. 1. Data center load forecast for 2030 for Texas [3]

The projected growth in data-center electricity demand is driven primarily by the rapid deployment of high-density computing infrastructure supporting artificial intelligence (AI) and cloud-based applications. These advanced facilities require substantially greater power per unit of compute capacity, consuming up to eight times more electricity than traditional data-center architectures. Meeting this escalating demand will necessitate the addition of substantial new, reliable power-generation capacity.

Reliable access to substantial electric power now drives most data-center siting decisions, shaping performance, reliability, and total cost [6]. Other major factors include infrastructure quality, network connectivity, environmental conditions, workforce availability, and proximity to end users.

The following sections of this report examine the implications of different power-sourcing strategies for data-center operations and assess the emerging challenges associated with powering current and future facilities in Texas.

Chapter 2
Powering AI data centers

The acceleration of digitalization and artificial intelligence (AI) is generating unprecedented demand for computing resources and electrical power. AI-oriented data centers increasingly depend on specialized high-performance processors, including Graphics Processing Units (GPUs) and Tensor Processing Units (TPUs), to support large-scale model training and real-time inference workloads. These computational requirements necessitate dense, low-latency architectures and highly reliable power-supply and thermal-management systems to ensure continuous and efficient operation [6].

Present status, the rise of Artificial Intelligence (AI) scale power demand and its challenges

Hyperscale and AI-focused facilities represent a rapidly expanding class of high-density, high-availability loads with distinct operational characteristics compared to traditional enterprise data centers. Understanding the underlying computational drivers, associated power and cooling requirements, and reliability considerations are essential for evaluating the implications of this growth for future electricity infrastructure and resource planning.

a) Training vs. Inference Workloads

The advent of hyperscale data centers—highly scalable facilities engineered to support large and complex digital workloads—has enabled rapid advancements in AI. AI applications such as conversational agents, autonomous vehicles, and real-time analytics systems require substantial dedicated computational resources across both training and inference phases.

Inference data centers apply pre-trained models to new data inputs to generate real-time predictions or decisions, whereas training data centers focus on model development by processing massive datasets to extract learned patterns. Training state-of-the-art models such as GPT-4 or Google's Gemini involves optimizing trillions of parameters and typically requires deployment of thousands of high-performance accelerators, including GPUs, TPUs, and Field-Programmable Gate Arrays (FPGAs). These components consume significantly more power than conventional CPUs due to their parallel-processing architectures and high core densities. Continuous retraining and fine-tuning further elevate sustained energy consumption.

b) Power and Cooling Requirements

The high-density hardware required for AI workloads demands advanced thermal-management systems. Although AI-optimized data centers may employ fewer racks than conventional designs for equivalent power capacity, their accelerated power density produces greater thermal loads, necessitating enhanced cooling infrastructure. System characteristics such as high-bandwidth memory (HBM) and high-speed interconnects further contribute to increased heat generation.

To meet these requirements, data centers require extensive infrastructure upgrades, including expanded electrical capacity, modernized transmission and distribution equipment, and advanced cooling systems. Aging facilities become less reliable and energy-efficient over time, increasing downtime risk and operating costs. Large-scale data storage and transfer consume significant energy and produce substantial heat; in less efficient enterprise environments, cooling alone may account for up to 30% of total energy consumption, underscoring its critical role in evaluating power demand [7], [8].

c) AI power demand characteristics

Fig. 2 illustrates the projected breakdown of data-center power consumption between AI and non-AI workloads for 2024 and 2030 [9]. AI-related workloads are expected to increase significantly, rising from approximately 20% of total data-center power consumption in 2024 to an estimated 40% by 2030. Within this category, large-scale training workloads—characterized by highly parallelized compute operations and continuous iteration—are projected to account for nearly 80% of AI-specific electricity use. These projections emphasize the substantial and growing contribution of AI computing to overall data-center energy demand, highlighting the scale and urgency of addressing associated power and infrastructure requirements. Such accelerated growth presents major challenges for electric-grid planning and resource adequacy, requiring timely investment in new generation capacity, transmission expansion, and enhanced system flexibility to maintain reliability and meet rising load demands.

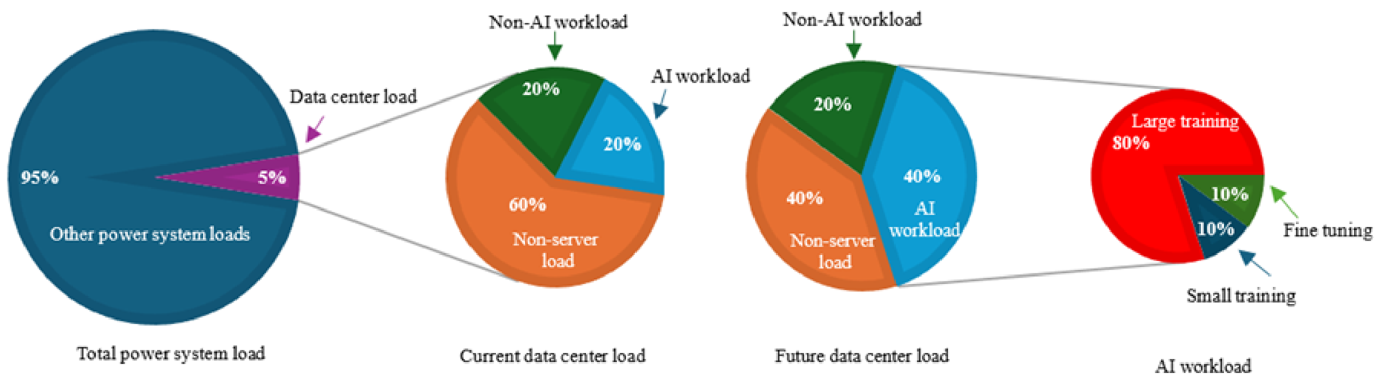


Fig. 2. Current and future breakdown of data center and AI workload [9]

Data centers have shifted the focus of grid dynamics from primarily supply-side considerations to demand-side variability. This results because the grid was optimized to provide power to loads that were primarily motors or heat. Data centers provide a primarily electronic load with novel temporal variations. Modern hyperscale facilities operate as dynamic electronic loads—referred to by ERCOT as Large Electronic Loads (LELs)—driven by rapid fluctuations in power consumption from high-performance IT servers and storage systems, particularly during AI model training. These facilities exhibit substantial short-term variability in power demand, characterized by large-magnitude swings over seconds to minutes, which impose significant stress on upstream generation and transmission infrastructure.

At hyperscale or giga-scale capacity levels, abrupt power swings occur during compute-intensive phases such as GPU synchronization, in contrast to more stable consumption during communication or data-transfer intervals. Increased frequency and duration of training cycles amplify the magnitude of these fluctuations. Fig. 3 [2] illustrates the variation in electrical demand during training of LLaMA-70B—a large language model developed by Meta for advanced coding applications—across eight compute nodes. The observed waveform demonstrates demand swings of up to 60% from idle to peak within short time windows, highlighting the inherent volatility associated with AI workloads.

Further evidence is provided in Fig. 4, which presents the Fast Fourier Transform (FFT) analysis of large versus small model training and fine-tuning workloads based on simulation studies by Ko and Zhu [9]. The frequency-domain representation reveals distinct peak magnitudes and occurrence patterns associated with different workload types, illustrating the scale, periodicity, and forcing characteristics that can induce grid-level oscillations. These characteristics directly affect ERCOT's operational planning, requiring expanded procurement of fast ancillary services (e.g., FFR), enhanced ramping capabilities, and more stringent interconnection requirements to maintain frequency stability under rapidly varying load conditions.

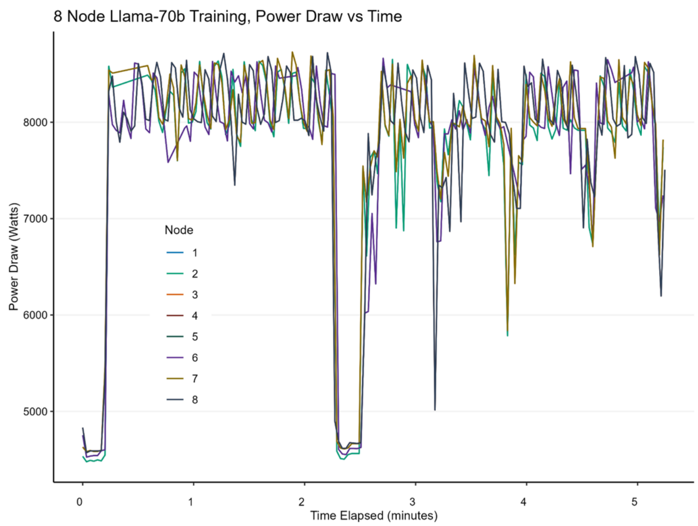


Fig. 3. Example of power demand during Llama-70B training across 8 nodes [2]

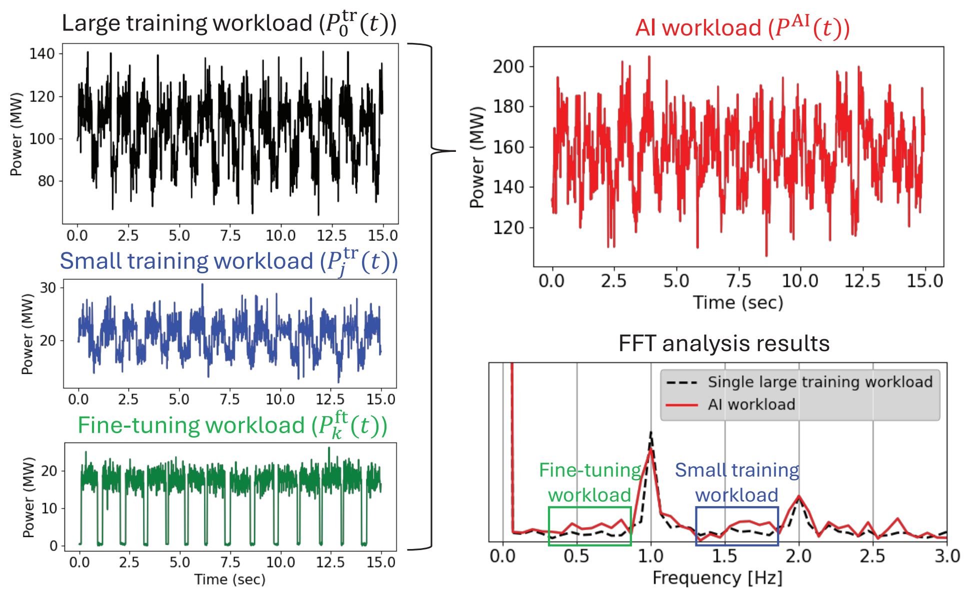


Fig. 4. AI workload power demand modeling [9]

Simulation studies have examined the sensitivity of grid frequency to increasing levels of data center load penetration, as illustrated in Fig. 5. The results demonstrate that as data center penetration levels increase from 7 GW (normal) to 14 GW (high), and subsequently to 21 GW (extra-high) [9], the magnitude of frequency oscillations correspondingly rises. At the highest penetration level, peak-to-peak frequency deviations approach approximately 0.5 Hz, indicating substantial system stress. Such deviations beyond prescribed regulatory limits can adversely affect generator performance, reduce equipment efficiency, accelerate component aging, and ultimately threaten overall system stability. These findings stress the need for transmission operators and planning authorities to incorporate load-induced frequency volatility into long-term resource adequacy strategies, ancillary service procurement, and interconnection standards.

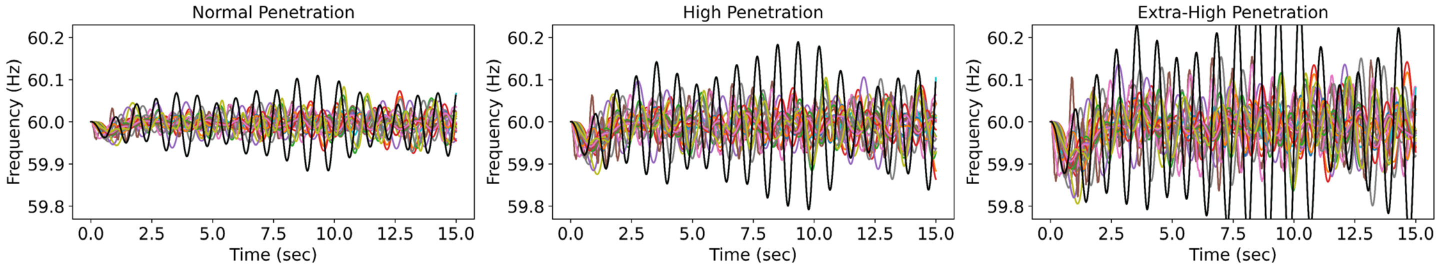


Fig. 5. Frequency variation for different data center load penetration levels [9]

d) Power quality and sub-synchronous oscillations

The unique characteristics of data center load profiles during training or inference workloads can pose significant local power quality challenges to nearby buses, including harmonics, voltage sags/swells, and flicker, as reported by utilities and customers. In addition, subsynchronous oscillations associated with pulsed or rapidly varying power demands have been reported, which may excite torsional modes in the shafts of nearby synchronous machines. To mitigate these issues, several developers are adopting fast-response power-electronic solutions and high-density energy storage systems, such as batteries interfaced through grid-forming inverters, which have been shown to provide effective damping of such oscillations. Power quality disturbances and forced oscillations arising from these operating conditions are currently the subject of intensive research in both industry, academia and regulatory agents. Regulatory frameworks to be included in the grid codes for fast load slopes are being contemplated by FERC.

e) Reliability and Resiliency

Because data centers must sustain continuous operation, disruptions resulting from grid instability, extreme weather, equipment failure, or fuel supply constraints pose risks to service continuity. As a result, maintaining energy resilience to mitigate these risks requires diversified energy sources, redundant backup generation, and advanced power-management architectures.

As articulated by David Bills (Microsoft), reliability reflects the outcome—ensuring uninterrupted service—while resiliency represents the capabilities that enable sustained operation despite disruptive events [10], [11]. Power-system reliability is commonly evaluated using Loss of Load Probability (LOLP), which describes the likelihood that electricity demand will exceed available generation capacity within a defined time. Data centers generally target LOLP levels approaching zero due to the necessity of maintaining continuous 24/7 operation. Industry standards classify facilities into Tier I through Tier IV based on redundancy design, fault tolerance, and guaranteed uptime levels [12] **(Table 1).**

Table 1: Tier classification of data centers

Tier	Description	Uptime (%)	Downtime hours in a year
I	Basic capacity	99.671%	28.8
II	Redundant Capacity Components	99.741%	22
III	Concurrently Maintainable	99.982%	95 minutes
IV	Adds fault tolerance to Tier 3	99.995%	26.3 minutes

Tier I data center sites face elevated exposure to risks associated with grid instability, extreme weather events, water scarcity, and cybersecurity threats. Grid stability is a particular concern when data centers rely directly on utility supply without substantial redundancy or on-site backup resources. The grid stability is a matter of cost. The loads needing extremely high reliability provide their own solutions rather than burdening the more outage tolerant users with the cost.

f) Sustainability and Cost

In addition to instantaneous load variability, an important metric for evaluating data center operational efficiency is Power Usage Effectiveness (PUE) ([13], [14]). PUE is defined as the ratio of total facility energy consumption—including IT servers, storage, networking equipment, cooling, lighting, and power distribution—to the energy consumed solely by IT equipment. A lower PUE indicates greater energy efficiency, as a smaller portion of total energy is expended on non-computational functions. Contemporary data centers typically report PUE values in the range of 1.2–1.5, while highly optimized hyperscale facilities are projected to reduce cooling-related energy consumption to below 7% of total usage.

Traditional and AI-oriented data centers are different in terms of their rack power density, measured in kW per rack. Conventional enterprise servers average approximately 15 kW/rack, whereas rack densities for emerging AI workloads commonly range from 60–120 kW/rack, and continue to trend upward [15]. Although increased rack density enables reduced physical footprint and greater compute concentration, it also necessitates significantly more advanced thermal management systems, alternative heat-rejection architectures, and more robust power delivery infrastructure. For example, a typical server dissipates approximately 1.5 kW of heat during standard operation, whereas AI servers equipped with high-performance GPUs generate five to six times that thermal output, requiring proportionally greater cooling energy and infrastructure support.

These escalating power density requirements and tightening PUE targets have direct implications for sustainability planning and grid resource adequacy, as they necessitate accelerated investments in high-efficiency cooling technologies, resilient power infrastructure, and coordinated long-term electricity planning to reliably support future AI-driven computational growth.

Meeting sustainability objectives—such as reducing energy consumption, lowering carbon emissions, conserving water resources, and minimizing waste—presents an additional challenge for data center operators. In the United States, the Energy Act of 2020 mandates that federal agencies evaluate the energy performance of their data centers at least once every four years [16], and the U.S. Department of Energy's Federal Energy Management Program (FEMP) provides technical guidance and tools to advance data center energy efficiency initiatives [17]. Beyond PUE, Carbon Usage Effectiveness (CUE) has emerged as a key sustainability metric, quantifying the greenhouse gas emissions associated with each unit of IT energy consumed within a facility [14]. Lower CUE values indicate a reduced carbon footprint and therefore a more carbon-efficient operation. In addition, Water Usage Effectiveness (WUE) is used to evaluate water sustainability and is defined as the ratio of annual site water consumption (in liters) to IT energy usage (in kWh) over the same period [14]. Total water consumption encompasses cooling systems, humidification, and other operational processes.

Typical WUE values vary widely, ranging from approximately 2.5 for evaporative cooling systems, down to between 0 and 2.5 for hybrid cooling configurations, and approaching 0.2 for highly optimized hyperscale facilities.

Operating costs represent a significant concern for data centers, given their requirement for continuous 24/7 operation. According to internal studies conducted by Schneider Electric, approximately 50% of a data center's operating expenditure is attributable to electricity consumption [18]. Considering that total cost of ownership for data centers is typically divided roughly equally between capital expenditures (CAPEX) and operating expenditures (OPEX), the magnitude of energy-related costs reinforces their critical influence on long-term financial sustainability. Energy prices are expected to increase in the near term due to supply–demand imbalances, volatility in fossil fuel markets, and constraints within existing transmission and distribution infrastructure. As a result, operators face increased pressure to reduce OPEX in the short term. Over a longer horizon, improvements in energy efficiency, greater deployment of cost-effective renewable energy resources, and continued technological innovation may moderate or stabilize energy pricing trends [19], although substantial uncertainty remains (**Fig. 6**).

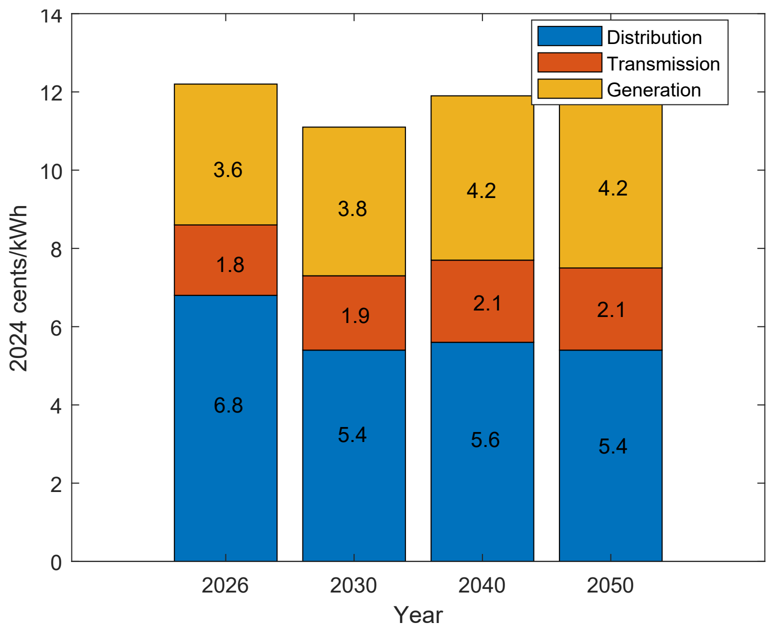


Fig. 6. Electricity prices by service category (AEO 2023 reference case) [19]

While advanced cooling and heat-rejection solutions—such as next-generation HVAC technologies—can improve efficiency and lower PUE, operators must ultimately secure energy at a cost lower than prevailing grid prices to maintain competitive operating margins. These considerations highlight the necessity for AI-driven data centers to adopt power strategies that are not only reliable and resilient, but also cost-efficient and environmentally sustainable. Dependence solely on grid power or traditional backup systems is insufficient to meet emerging requirements for continuous, high-performance operation. Instead, diversified and integrated energy portfolios are essential to support sustained computational growth while achieving economic and sustainability objectives.

How to cater to the increasing power demand of AI data centers?

Data centers may be powered through exclusively grid-connected systems, through on-site generation resources—often operating in a behind the meter (BTM) configuration—or through a hybrid approach that integrates both strategies. Data center developers determine the optimal configuration based on the site and is influenced by factors such as its geographic location, their risk tolerance, and sustainability objectives, regulatory considerations of the location, and financial resources of the developer.

The legacy approach to powering data centers relies on the public electric grid, supplemented by uninterruptible power supplies (UPS) and diesel generators (DGs) for emergency backup (Fig. 7). However, diesel-based backup systems are increasingly regarded as suboptimal due to reliability concerns associated with startup delays, limited operational duration, high maintenance and fuel costs, noise and emission impacts, and constrained scalability. These limitations draw attention to the need for more resilient and sustainable backup architectures capable of supporting continuous high-density AI workloads.

Grid connected operation offers several advantages, including access to large power supplies, and cost management strategies enabled through market participation mechanisms such as Demand Response (DR). Nonetheless, large-scale interconnection of data centers can intensify grid congestion, reduce hosting capacity, and increase local marginal prices. To address those setbacks, it will necessitate substantial upfront transmission and distribution upgrades—costs that are passed to consumers. In addition, interconnection timelines can extend across multiple years due to regulatory permitting and utility approval processes. This makes grid expansion challenging to align with the rapid deployment schedules of hyperscale facilities.

Demand Response (DR)

A grid management tool where consumers (homes, businesses, data centers) voluntarily reduce electricity use during peak demand or grid stress, often in exchange for payment, to prevent blackouts, lower costs, and support clean energy goals

Major hyperscale developers—including Google, Meta, and Microsoft—are actively pursuing alternative system architectures to mitigate escalating power requirements. These include transitions from traditional alternating current (AC) distribution to direct current (DC) architectures, which reduce conversion losses and cooling burdens [20], as well as the adoption of Bring Your Own Power (BYOP) models that incorporate on-site generation. On-site power generation can be behind-the-meter (BTM) systems, and hybrid configurations

. Behind the meter generation reduces dependency on centralized grids. They are customer owned generation located on the customer side of the utility meter. They minimize the need for additional transmission and distribution grid infrastructure and accelerate deployment by avoiding long interconnection queues. In smaller BTM configurations, the grid typically remains the primary supply, while larger systems may utilize the grid primarily as backup.

Behind the meter (BTM)

Customer owned generation located on the customer side of the utility meter.

Microgrid

A microgrid is a subset of the BTM architecture in which a group of interconnected loads and distributed energy resources operates as a single controllable entity with respect to the grid.

Microgrids represent a larger scale form of BTM architecture, integrating on-site generation (e.g., natural gas, renewables), energy storage, local loads, and control systems capable of islanded operation using a local controller (Fig. 8) [21] [22]. Microgrids enhance energy security, reduce exposure to grid outages, and provide faster access to power by circumventing transmission bottlenecks. They can also serve broader community infrastructure, extending beyond single-facility applications.

Islanding

When a section of the electrical grid becomes isolated from the main network but continues to be powered by a local energy source, like solar panels or a microgrid.

Hybrid generation systems have been used successfully in applications like military bases and university campuses for decades. They combine centralized grid power with on-site natural gas units, renewable resources, and battery energy storage with a coordinated centralized controller. These systems address limitations associated with reliance on a single source and support a balance of economic, reliability, and sustainability objectives. Designing hybrid systems requires comprehensive energy system design optimization, as cost, carbon footprint, startup schedules, regulatory constraints, land availability, and incentives vary significantly by site. For example, natural gas may offer lower cost and higher dispatchability where accessible, while solar, wind, battery storage, and hydrogen provide low-carbon alternatives at higher capital cost [23]. Complex hybrid systems may dynamically switch among supply options based on load profiles, market conditions, and operational priorities.

Hybrid generation

Using power from different sources, including the centralized grid, on-site generation using natural gas, nearby solar and wind facilities, combined heat and power (CHP), and energy storage. A hybrid generation can include BTM or Microgrid and the grid. In contrast to BTM or Microgrid, in a hybrid generation design, the utilization of the grid versus an on-site generation portfolio (there can be multiple types) is determined by a centralized controller.

Emerging technology pathways are also gaining traction as hyperscale developers explore more dispatchable and lower-carbon solutions. Examples include natural gas paired with carbon capture (Google and Crusoe), geothermal systems (Meta), and small modular nuclear reactors (Microsoft's initiative at Three Mile Island) [24], [25]. As energy availability becomes a critical constraint on AI expansion, diversification of technological solutions is expected to accelerate.

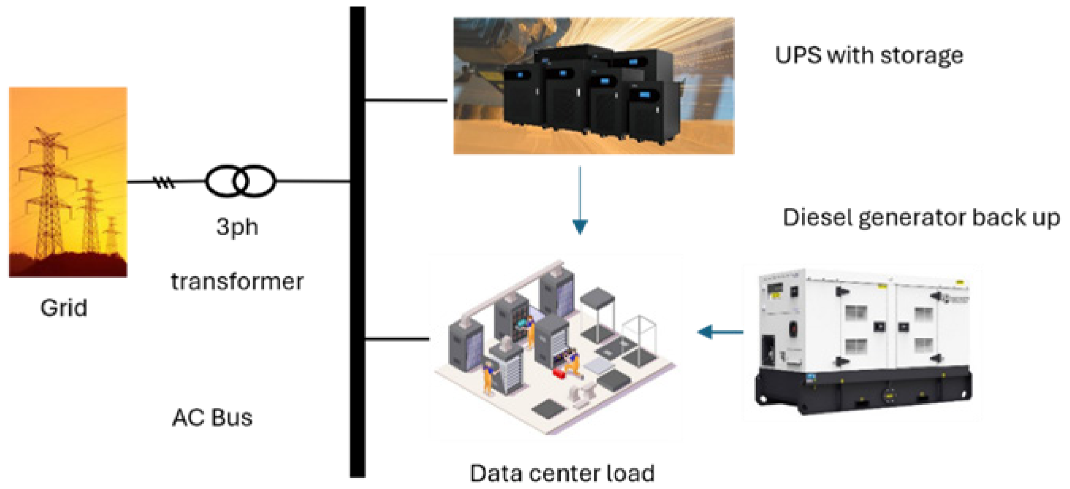


Fig. 7. The traditional data center power system with DG and UPS

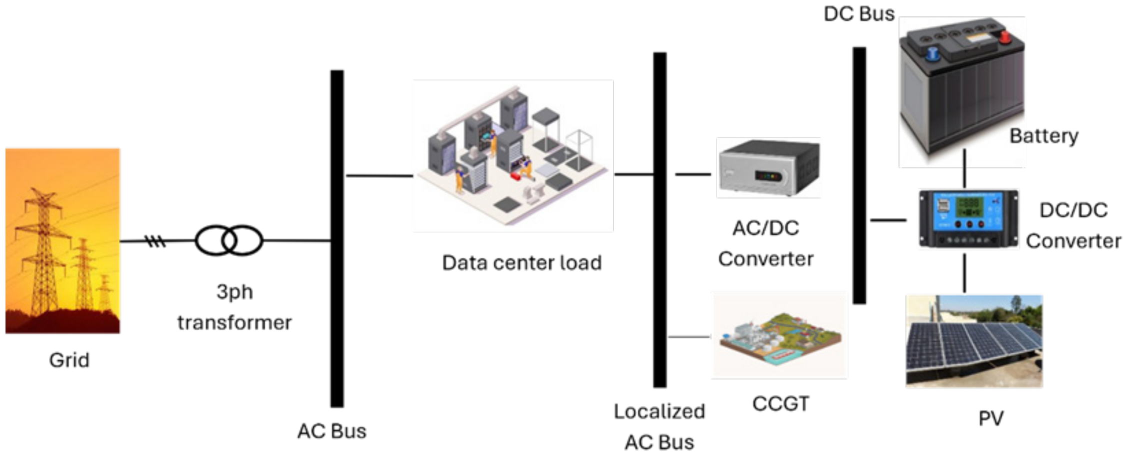


Fig. 8. Behind the meter (BTM) grid connected on-site power generation

AI-optimized data centers increasingly rely on advanced thermal management solutions—including liquid cooling, direct-to-chip cooling, immersion systems, and next-generation HVAC—to maintain safe operating temperatures and prevent hardware failure [26]. Traditional air cooling can consume up to 40% of total facility energy, whereas liquid cooling approaches can reduce overall energy consumption by 10% or more, enabling higher power densities and improved sustainability.

Comparing grid only, BTM and microgrids and hybrid generation for data centers

The choice between grid only, BTM, microgrids and hybrid generation as data center power source involves tradeoffs in cost, deployment speed, sustainability and reliability. Each of the three power supply options presents its own set of benefits and limitations, as summarized in **Table 2** and **Fig. 9**.

Table 2: Comparison summary of grid only connection, BTM, and microgrids and hybrid generation for data centers

Comparison Feature	Grid only	BTM and microgrids	Hybrid generation
Definition	Completely connected to the centralized grid	A self-contained electrical network that generates and manages its own power on-site, behind the utility's main meter. Excess generation can be fed back to the grid. Microgrids are advanced BTM systems, with limited or no access to the grid and can supply a community.	Combines different power sources, and grid. BTM generation and storage, within an energy system with an advanced control system. Actively connected to the grid, however the main generation is from the hybrid sources.
Reliability	Low, Susceptible to grid outages	High, Independent Source	Extremely high, combines grid with back up sources
Upfront Cost	Lowest, avoids generation infrastructure costs	Highest, including full power plants and maintenance costs	High, including cost of multiple hybrid systems and its control

Energy cost	Variable, costs are subject to utility rates, which depends on fuel price, time of use and demand	Low, you can decide your own rates and avoid fluctuating utility rates	Mixed, costs are determined by onsite and power purchased from grid.
Transmission and distribution costs	High, customers pay for the operation, maintenance and upgrades of the transmission and distribution grid	Very low, as the power is produced and consumed at the same location	Low, though the system is connected to the grid, a significant transmission infrastructure cost is avoided by generating your own power
Operation and maintenance cost	None, entirely handled by utility	High, the consumer is responsible for the maintenance and repairs of the generating equipment and transmission	Moderate to high, maintaining onsite equipment, but some grid connection costs can be avoided.
Time to power	Longest, due to interconnection queues.	Fastest, avoids grid interconnection delays and utility queues	Moderate to fast, depending on the complexity of the hybrid system
Operational control	Minimal control as reliant on utility	Highest, full control over the power quality and supply	High, control over generation mix, and storage. Flexibility to interact with the grid.
Sustainability	Dependent on the generation mix of the utility grid	Highly dependent on the onsite fuel source and storage	Potential for integrating renewables, to reduce environmental footprint.
Complexity	Lowest as utility manages the infrastructure	High, requires expertise in energy production and management	Highest, as managing multiple sources, grid, and storage.
Primary Vulnerability	Exposure to centralized fossil fuel-based generation and large scale transmission	Dependance on a single technology, subject to inconsistent performance or obsolescence.	High initial complexity and capital cost, needing sophisticated management
Regulatory risk	High exposure to policy shifts, including carbon pricing, fossil fuel phase-outs, and utility regulation	Moderate exposure to policy regulations only those affecting distributed generation and grid interconnection	Low exposure due to diversified assets, affected by policies on storage and renewables.
Financial risk	High risk of cost volatility due to fossil fuel market changes and potential for stranded utility assets	Depends on a single technology and its cost effectiveness over time	Diversified financial risk protects volatility of a single asset type.
Water Usage	High for thermal plants and low for renewable plants, depends on the generation mix. Can be vulnerable to regional water shortages.	Varies by technology. Solar plants require less water for operation, while fuel cells need minimal water for operation and maintenance. Independent from water related grid vulnerabilities	Low to moderate, combines the low-water-use benefits of onsite renewable generation with the grid's power, reducing the need for onsite thermal generators that require more water. Resilience against water related grid issues.

Land Use	Vast and dispersed, power plants require significant land varying by fuel type. A large amount of land is dedicated to transmission lines, and substations increasing the land cumulative land footprint.	Small and concentrated, Localized land use near the point of consumption. Has complete control over land use.	Combined, blend of on-site and grid connected, using less land compared to a completely stand alone system. Optimized land use by sizing on-site generation to meet a portion of demand, using the grid for the rest, avoiding a massive on-site footprint.
Permitting complexity	Requires federal, state, and local approvals, plus a complex utility interconnection agreement. The process can take years, with frequent delays.	It depends on local land use, zoning and building codes. Projects can go through fast track approval process in most areas	Combined requirements for onsite and stand alone generation, plus additional permits for integrating different energy sources and storage.
Community perception	Generally viewed as reliable, consistent, affordable, and highly accessible in urban and sub urban areas. Less land, water footprint and high regional balance.	Considered less reliable as it depends on user defined generation and storage, depending on weather conditions. High upfront cost, but having long term cost certainty, cleaner energy sources with local operational independence.	Perceived as the most reliable option with advantages of both the other options. Highest upfront cost due to multiple sources and their integration. Positive environmental impact and high degree of local control. Most resilient of all the three options in climate change, natural disasters and grid security.
Scalability	Least scalable due to the centralized grid and transmission infrastructure constraints	Moderately scalable as the power source and transmission facilities are independent of the grid.	Highly scalable as it combines both options, depending on the availability of grid for back up generation.

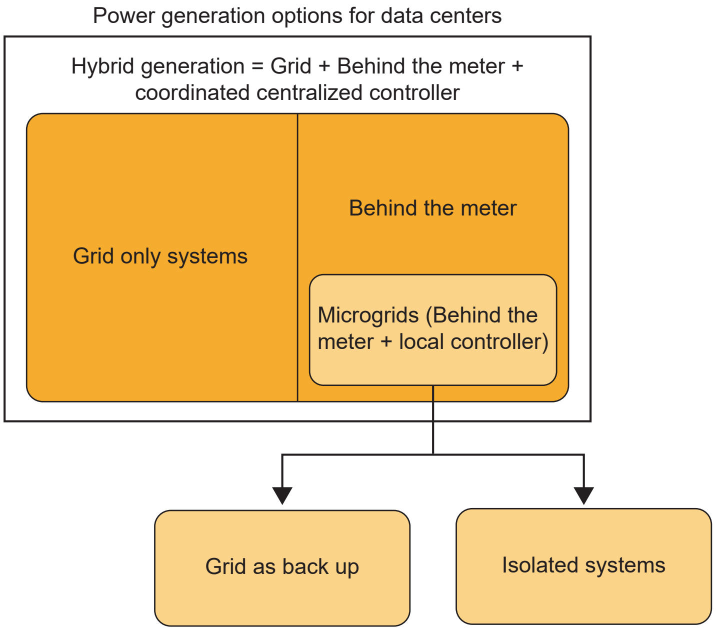


Fig. 9. Power supply options for data centers

According to a recent report by Bloom Energy [27] , approximately 38% of data centers in the United States are projected to utilize off-grid hybrid generation systems by 2030, increasing to an estimated 50% by 2035. Notably, the study also indicates that 27% of data centers are expected to depend exclusively on on-site power generation as their primary energy source, reflecting a significant shift away from conventional grid-centric architectures. Globally, an estimated 4.8 GW of on-site generation capacity has already been announced for deployment by 2030, underscoring the accelerating transition toward distributed and self-sufficient energy models.

The predicted hybrid growth is consistent with loads needing reliability significantly exceeding the norm on a system cover the cost of that unique need. Going with completely on-site power may create collateral issues. Communities may request limited or emergency power access to the behind the meter power as a siting requirement. Second the Department of Energy recently published a paper showing behind-the-meter generation as one of the top contributors to electricity price increases.

Selecting the optimal power supply strategy for a given data center ultimately requires balancing multiple criteria, including power accessibility and availability, sustainability objectives, reliability and resiliency requirements, and underlying technical and economic constraints. Consequently, no single configuration is universally optimal; rather, decisions must be tailored to site-specific conditions and long-term operational priorities.

Chapter 3

On-site generation for data centers in Texas

The Texas landscape and ERCOT constraints

Meeting the accelerating power requirements of AI-driven data centers will require substantial capital investment in new transmission lines, substations, and associated infrastructure necessary to integrate these large loads into the existing electric grid. Within the ERCOT system, the growing interconnection queue has contributed to increased transmission and substation costs, greater grid congestion, and heightened operational strain. The projected ERCOT large-load interconnection queue for 2025–2030 (Fig. 10) [28] remains highly dynamic, undergoing significant monthly revisions as new projects are proposed and existing applications are modified or withdrawn.

In response to these challenges, legislative and policy actions—such as Texas Senate Bill 6 (SB6)—alongside regulatory mechanisms including mandatory DR programs, are being advanced to enhance grid resource adequacy and resilience. Additionally, procedural and technical improvements, such as accelerated interconnection processes, advanced load forecasting techniques, and dynamic grid modeling, aim to reduce queue backlogs and improve long-term system reliability. In addition, FERC has been instructed to address the issue. Although the FERC requirements do not affect Texas, they can provide insight as to further improvements Texas might employ.

The combined effects of rising demand, long interconnection timelines, reluctance to use emerging control strategies, and network bottlenecks have made access to grid-supplied power a critical barrier to new data center development in Texas. Consequently, these conditions have accelerated the adoption of established behind-the-meter (BTM) generation technologies and hybrid power architectures as nearterm strategies to ensure timely, reliable, and resilient power availability. A range of additional factors further support the shift toward alternative power solutions, as outlined below.

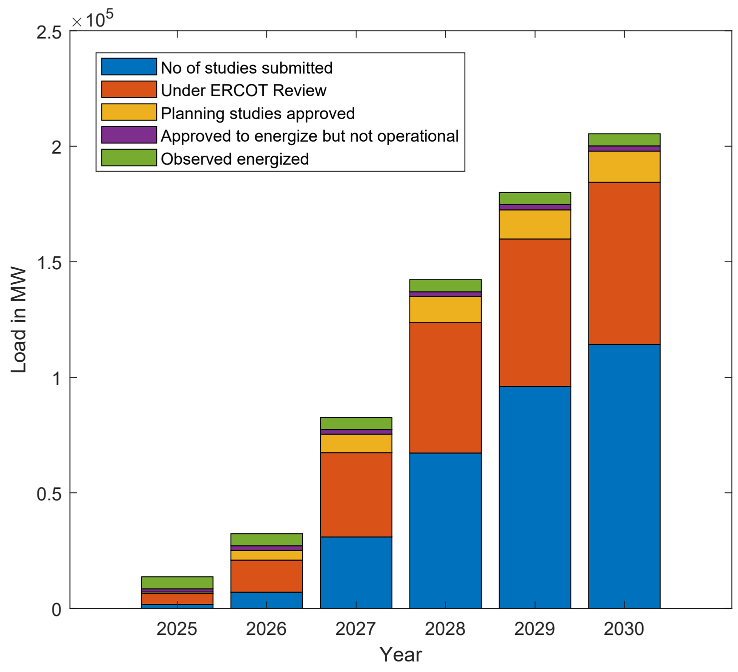


Fig. 10. ERCOT large load interconnection queue (as of October 2025) [27]

a) Time to power

Time to power is the time taken for the new facility to receive necessary power. The large load (loads > 75 MW including data centers) interconnection queue of the ERCOT grid is 13 GW in 2025 and is expected to rise to 205 GW in 2030 as of October 2025 (Fig. 10). Integrating on-site generation will improve the time to power by allowing data centers to bypass lengthy grid interconnections and permitting processes of the ERCOT grid. Data centers with on-site generation can avoid waiting can avoid waiting for equipment delivery and construction of new transmission infrastructure, such as high-voltage transmission lines. Accelerated access to power will help data centers come online faster and secure a competitive advantage in the market, with associated economic benefits. They also reduce strain on the existing centralized grid. However, hybrid generation systems are complex and may require more time than BTM systems, though they remain faster than relying solely on grid-connected power. These trends directly impact ERCOT's operational planning, requiring more adaptive forecasting, resource adequacy planning, and real-time operational strategies to manage increasingly distributed and hybrid supply configurations.

b) Reliability

ERCOT operates a grid that is largely islanded and not synchronized with the Eastern or Western U.S. interconnections. The Texas electricity production and usage is larger than the smallest thirty states combined. Thus, it has good capability to resolve supply–demand fluctuations internally. As elsewhere, data centers in Texas must prioritize redundant and flexible power strategies to maintain near-continuous uptime requirements (approaching 99.9999%). Extreme weather events can create peak demand conditions that exceed available supply. This forces system operators to implement emergency measures including voluntary curtailment and, in severe cases, load shedding. Data centers exceeding 75 MW are classified as large loads by ERCOT and may be required to participate in curtailment programs or operate on-site backup generation to support grid stability during stress conditions. In this context, on-site power plants offer a dedicated and resilient source of power that can sustain operations independently of the grid. When coupled with fast-responding energy storage and dispatchable generation, hybrid systems can handle both short-duration spikes and sustained high-load needs, significantly reducing overall Loss of Load Probability (LOLP) while enabling continuous operation even during grid disruptions.

c) Reduced energy cost

On-site generation can significantly reduce operating costs because it eliminates transmission and distribution losses associated with traditional grid supply. Industrial customers currently pay approximately 6.6¢/kWh on average for ERCOT grid power [29]. On-site generation helps mitigate exposure to extreme price spikes that characterize the Texas market during periods of peak demand, such as summer heat waves and extreme winter cold. Scarcity pricing in ERCOT's energy-only market structure incentivizes generation investment and can provide a major revenue opportunity for data centers capable of exporting excess energy during high-price intervals. Although constructing on-site renewable and storage assets such as solar, wind, geothermal, and batteries involves substantial upfront investment, ongoing operating energy costs are significantly lower because these systems avoid dependence on volatile fossil-fuel markets. Ultimately, the Return on Investment (ROI)—accounting for construction, operation, and maintenance—determines the economic and strategic value for the data center operator. These economic trends also shape ERCOT's long-term grid planning strategies by increasing the need to integrate distributed, customer-sited generation into resource adequacy models and demand forecasting frameworks.

d) Enhanced sustainability

Data centers are under increasing scrutiny for their impact on local communities, requiring strict compliance with permitting and environmental standards. They must demonstrate sustainability in air quality, noise, and broader environmental impacts. This necessitates air permitting—the process of obtaining government approval for air emissions from equipment such as cooling towers and emergency backup generators. In Texas, the Texas Commission on Environmental Quality (TCEQ) prescribes three air permitting pathways for data centers based on potential emissions: Permits by Rule (PBR), Standard Permits, and New Source Review (NSR) [30], [31]. On-site and hybrid generation using renewable sources can help companies meet greenhouse gas (GHG) standards and align with sustainability commitments. Many data centers offset energy costs by entering into behind-the-meter Power Purchase Agreements (PPAs), sourcing electricity directly from their owned solar or wind facilities rather than from the grid [4]. Hybrid agreements increasingly bundle solar with battery storage, enabling operators to use clean power during ERCOT peak periods. Virtual Power Plants (VPPs)—networks of distributed energy resources controlled by software and aggregated to operate like a single power plant—are emerging as additional revenue opportunities for PPA participants. These developments also have significant implications for ERCOT's long-term grid planning, as growing penetration of distributed and customer-sited resources requires enhanced visibility, coordination, and operational forecasting to ensure system reliability and resource adequacy.

e) Grid stability

The fluctuations and momentous variations in the grid can be curtailed by on-site and hybrid generation with medium-term storage options such as battery energy storage systems (BESS). They act as stabilizing forces on the larger power grid, meeting the varying and spiky loads demanded by data centers (Figs. 3, 4, and 5). This increases operational flexibility and grid stability. Hybrid systems are more flexible than purely on-site systems, as operations rely on advanced controllers that manage the interconnection of the central grid with other power sources to ensure continuous, disturbance-free load fulfillment, utilizing storage systems whenever needed. Storage with greater power density (such as hydrogen for fuel cells or supercapacitors) can enable more efficient use of space within data center facilities, increasing revenue potential. Additional solutions such as thermal storage systems and combined heat and power (CHP) cogeneration systems can also serve facility heating and cooling needs.

On-site generation presents an attractive alternative to utility-supplied grid power, particularly for co-location data centers and their tenants. However, relying solely on off-grid on-site generation requires significant upfront investment and surplus capacity to ensure redundancy, creating substantial cost barriers for organizations. Modular power systems offer expandable options, but limited availability can hinder scalability, and rapidly growing data centers may find it costly and complex to scale up relying solely on on-site generation. Dependence on a single source similarly exposes facilities to vulnerability during supply disruptions. Hybrid generation provides the advantage of increased flexibility, resilience, and potential cost savings by combining the benefits of on-site generation and grid-connected systems. For ERCOT planners, the increasing adoption of hybrid and distributed resources at large-load data centers necessitates updated planning models that account for controllable behind-the-meter assets, localized grid impacts, and improved forecasting of flexible resources, fundamentally reshaping transmission investment decisions and resource adequacy strategy.

Many existing hybrid facilities use contractors to manage the local power. Growth of local power management services promises to accelerate the development of higher quality systems.

Status of on-site generation for data centers in Texas

Since the process has been designed, funded, and staffed to respond to traditional growth rates, the rapid expansion of large-load requests has contributed to a substantial increase in the ERCOT interconnection queue. Approval will drive an expansion of transmission infrastructure costs to minimize additional strain on the system. These conditions necessitate major investments in new substations and high-voltage power lines to ensure adequate capacity and system reliability. If the increased capacity is used by increased load, the existing regulatory structure should minimize additional consumer cost.

Concerns regarding the management of this accelerated growth in power and infrastructure needs have prompted legislative action, including SB6 [32]. SB6 establishes a regulatory framework to oversee the increasing number of large energy consumers by requiring data centers and similarly large loads to assume an equitable share of transmission-related costs and participate in emergency power management programs. The legislation also mandates upfront fees for transmission studies, ensuring that grid planning and infrastructure expansion costs are not unfairly redistributed to other ratepayers.

In response to unprecedented energy demand and ongoing grid reliability challenges, many developers are pursuing dedicated on-site natural gas generation and co-located renewable resources, paired with bridging battery storage, to secure reliable and cost-effective power for data center operations. This approach offers accelerated deployment timelines, operational flexibility, and enhanced resiliency—addressing the prolonged delays associated with ERCOT grid interconnection processes while also supporting sustainability objectives.

Table 3 summarizes a selection of ongoing Texas data center projects that have adopted on-site or hybrid power generation strategies to address grid constraints and accelerate access to reliable power.

Table 3: Ongoing data center projects with on-site and hybrid generation in Texas [27]

Project	Location in Texas	Capacity	Onsite generation	Capacity
Tract	Caldwell County	2 GW	Natural Gas	360 MW
Texas Critical Data Centers (TCDC) with Power Forward Energy	Ector County	1 GW	Natural Gas	250 MW
Cloudburst and Energy Transfer near New Braunfels	Hays County	Not Available	Natural Gas	1.2 GW
Open AI-Stargate project in Abilene	Taylor County	4.5 GW	Natural Gas	360 MW
Data City at Laredo	Laredo	5 GW	wind, solar, batteries, and dual-fuel gas turbines initially running on Texas-produced natural gas and later shifting to 100% green hydrogen sourced from a 2TWh Hydrogen City salt dome storage facility.	5 GW
Soluna- Project Fei	Northern Texas	100 MW	Co located with a Solar Farm	240 MW
Soluna- Project Gladys	Southeast Texas	150MW	Co located with a Wind Farm	226 MW
HyperGrid	Amarillo	11 GW	nuclear, natural gas, wind, solar, and battery storage.	1 GW nuclear

Tier I data center markets in Texas—such as the Dallas–Fort Worth, Houston, and Central Texas (Austin/Round Rock) metro regions—host major operators including QTS, CyrusOne, Element Critical, and Vantage Data Centers, each running large multi-facility campuses. However, Tier I markets also face elevated risks related to grid instability, extreme weather events, water scarcity, and cybersecurity threats, which increases the complexity of reliable long-term power planning and operational resilience. These challenges have significant implications for ERCOT's transmission planning and interconnection processes, as growing concentration of large loads in Tier I markets intensifies pressure on local substations, transmission capacity, and reliability planning requirements.

Risks/ Challenges for on-site generation in Texas

Texas Senate Bill 6 and its implications: The integration of new large-load facilities into the Texas electric grid introduces substantial operational and infrastructure challenges which are associated with maintaining system reliability and stability. The rapid expansion of data centers and other high-demand users within the ERCOT footprint has been a primary driver behind the passing of SB6 [32], signed into law on June 21, 2025. SB6 establishes new financial, operational, and reliability requirements for ERCOT large-load customers, particularly those with demand exceeding 75 MW.

Under SB6, large-load customers are required to submit a minimum $100,000 upfront fee to cover initial transmission screening studies for grid interconnection. The legislation also mandates that large loads fund the necessary transmission and infrastructure upgrades directly rather than shifting costs to other ratepayers, including residential and distribution-level customers [33]. In addition, SB6 formally defines the conditions under which on-site backup generation may qualify as supporting grid reliability, specifying that on-site systems must be capable of serving at least 50% of the facility's load and must not export power back to the ERCOT grid.

Large-load projects subject to SB6 must also provide operational information to their interconnecting utility and to ERCOT and may be directed—given reasonable notice—to curtail load through demand response or activate on-site generation during periods of system stress. While SB6 introduces more equitable cost allocation and strengthens reliability provisions for both large and small customers, it also adds complexity and cost exposure for data centers exploring on-site or hybrid generation strategies. As a result, SB6 materially affects project economics, technology selection, and interconnection timelines for AI-driven hyperscale development within Texas. These heightened regulatory and financial obligations under SB6 are accelerating interest in on-site and hybrid power solutions—including microgrids, modular generation, and advanced demand-response architectures—as developers seek greater control over reliability, interconnection timelines, and long-term energy cost certainty within the ERCOT market.

Weather challenges and supply chain disruptions: One of the major challenges faced by the PUC-T is deciding opn the appropriate reliability investment to mitigate the effects of extreme weather events including hurricanes, tornadoes, and ice storms.. Texas uses Firm Fuel Supply Service (FFSS) [34] to ensure that certain generating sources have reliable fuel supply even under natural gas curtailments, thus maintaining grid stability. Generators must have firm gas storage agreements or firm transportation agreements certified by ERCOT (firm fuel contracts). The fuel supply must come from a pipeline that meets ERCOT's standards for a Qualifying Pipeline, and pipelines cannot serve human needs or participate in Demand Response (DR) programs, which could compromise reliability. The generator must be designated as a Firm Fuel Supply Service Resource (FFSSR), which is awarded through ERCOT's procurement process. ERCOT minimizes the risk of supply disruptions by including storage agreements that allow generators to stockpile natural gas in advance, ensuring reserves are available during emergencies. Natural gas power generation, which are part of data centers, must comply with FFSS standards to ensure they can supply demand during unforeseen circumstances. Disruptions in the complex supply chain of hardware, CPUs, GPUs, and generator components also threaten normal data center operations. These weather-driven vulnerabilities and supply chain risks increasingly motivate operators to adopt resilient microgrids, modular power systems, and hybrid generation strategies that reduce dependency on a fragile grid and enable more predictable planning.

As developers navigate these complexities and shift strategies to manage reliability, cost, and sustainability pressures—including the heightened risks from extreme weather events disrupting grid stability—market dynamics are accelerating changes in data center site selection trends. In 2025, the digital landscape is shifting away from congested Tier I markets facing power bottlenecks, soaring land and real estate costs, and tighter sustainability and Environmental Social and Governance (ESG) requirements. Hyperscale developers and colocation providers are actively scouting new territory and moving to next-generation locations, including Tier II and Tier III metros. These sites offer lower latency advantages driven by edge computing, content delivery networks (CDNs), AI inference, and streaming service demand. For example, Ashburn, Virginia—long known as the beating heart of "Data Center Alley"—has historically dominated the colocation and hyperscale market due to its proximity to federal agencies, the highest fiber density in the U.S., and a mature cloud ecosystem led by Amazon Web Services. However, other regions are now positioning themselves to become the next major hub. To compete, emerging markets must provide inexpensive land, favorable tax incentives, a reliable power grid with strong transmission capacity, proven uptime in extreme weather, efficient cooling, and access to major fiber routes. In Texas, these competitive advantages are accelerating the shift toward more advanced Tier II and Tier III (Table 1) locations such as Dallas, Fort Worth, Houston, and Austin. Dallas alone is a major Tier III data center market with a projected capacity of 2.01 GW in 2025 and 2.47 GW by 2030 [35]. The migration of hyperscale developers to Tier III and Tier IV locations has accelerated following Winter Storm Uri, driven by the appeal of multi-fuel power portfolios and stronger transmission interconnections. This transition has important implications for ERCOT's long-term grid planning, as the combination of regional growth and increasing weather volatility is reshaping infrastructure investment priorities and reinforcing demand for resilient solutions such as microgrids and hybrid power systems.

Air Permitting: As shown in **Table 3**, many ongoing data center projects rely on on-site natural gas–fueled power generation, which requires local air permitting to comply with environmental regulations. In Texas, air permitting for data centers is governed by the Texas Commission on Environmental Quality (TCEQ), which evaluates potential emissions from equipment such as gas turbines, reciprocating engines, cooling towers, and emergency generators. Depending on emissions levels and project scale, facilities must obtain one of three permit types: Permits By Rule (PBR), Standard Permits, or New Source Review (NSR) [30], [31] (**Table 4**). These permitting pathways define the regulatory requirements and review timelines for data center operators and play a critical role in planning and deploying on-site and hybrid power systems.

Table 4: Comparing three tiers of air quality permitting in Texas

Type of air permitting	Permits By Rule (PBR)	Standard Permits	New Source Review (NSR)
Stakeholders	For small, very specific emission sources and are the easiest and fastest to obtain, provided all conditions are met	For specific classes of moderate-emission facilities, offering a streamlined process compared to a case-by-case review, but still with fixed conditions	For larger or more complex sources that do not fit the criteria for a PBR or standard permit, require a detailed, case-specific engineering analysis, and include public participation
Major permitted emissions level	· 250 tons per year (tpy) of carbon monoxide (CO) or nitrogen oxides (NO_X) · 15 tpy of particulate matter with diameters of 10 microns or less (PM_{10}) · 10 tpy of particulate matter with diameters of 2.5 microns or less ($PM_{2.5}$) · **Low; for facilities that do not significantly contribute air contaminants**	For an electricity generating unit: · 250 tons per year (tpy) of carbon monoxide (CO) or nitrogen oxides (NO_X) · The equipment must produce less than six pounds of emissions per hour and ten tons per year. · **Moderate; for sources with higher emissions than PBRs but cannot trigger major NSR**.	· 100 or 250 tpy of a regulated pollutant · **High/Major Source; for facilities whose emissions exceed PBR and standard permit limits or "significant emission rates"**
Permitting process	Simplest; it only requires demonstrating compliance with pre-established conditions outlined in specific rules (e.g., Texas 30 TAC Chapter 106)	Streamlined; uses a set of pre-determined requirements for specific, well-characterized facility classes.	Case-by-case, detailed technical review
Renewal Cycle	Ongoing authorization if conditions are met	Typically, a ten-year renewal cycle	Typically, a ten-year renewal cycle

However, navigating these permitting pathways can introduce schedule uncertainty, and delays in receiving air permits directly affect time to power, making permitting risk a critical consideration for developers pursuing rapid deployment of on-site and hybrid generation systems.

Data centers can also utilize Emission Reduction Credits (ERCs), which function as market-based regulatory instruments designed to support compliance with air quality standards. ERCs represent verifiable reductions in air pollutants achieved at existing facilities, and they can be purchased by new or expanding emission sources to offset their projected emissions, thereby enabling permitting approval while maintaining regional air quality requirements. In Texas, these credits are administered by the Texas Commission on Environmental Quality (TCEQ) through the Emission Reduction Credit Program (ERC) [39]. By providing a flexible compliance pathway, ERCs can mitigate permitting barriers for large-scale data centers deploying on-site generation; however, the availability and cost of ERCs introduce additional planning and financial considerations that may affect project schedules, overall feasibility, and long-term operational strategy. From an ERCOT planning and ESG perspective, access to ERCs supports the development of lower-emission on-site and hybrid generation resources, enabling data centers to align with regional decarbonization goals while maintaining operational reliability and contributing to broader grid resilience initiatives.

Equipment lead time: Natural gas plants, transmission, and substation equipment currently face procurement lead times of 46–48 weeks, driven by global supply chain disruptions, rising demand from renewable energy projects, and long manufacturing cycles for critical electrical components such as large power transformers and switchgear (used for switching and protection in transmission systems). When transformers are damaged, replacement units are not readily available, creating prolonged delays before full operational capability can be restored. As a result, although on-site generation can theoretically accelerate deployment by bypassing the ERCOT interconnection queue, the advantage of rapid project initiation can be significantly undermined by extended equipment delivery timelines. This increasing dependence on long-lead equipment introduces substantial schedule and cost uncertainty, reinforcing the need to incorporate lead-time risk into time-to-power strategies and resource adequacy planning.

Water consumption: Advanced cooling techniques for AI data centers such as direct liquid cooling (DLC) or air-assisted liquid cooling (AALC) can operate in open-loop or closed-loop configurations, where closed-loop systems recirculate coolant and reduce total water usage. Water requirements vary significantly depending on the cooling technology deployed. The per-MWh water consumption of different electricity-generating technologies in West Texas indicates that combined-cycle natural gas (CCGT) generation has the highest water footprint, followed by solar + BESS and wind + BESS systems (Fig. 10) [40]. This variation directly affects local water resources depending on the on-site technologies selected for data center power and cooling. Increased scrutiny of water availability can trigger permit delays or denials in water-constrained regions, particularly where aquifer depletion and drought risk are elevated. As water-use considerations become more central, data center developers must integrate water-efficient cooling strategies and diversification of generation technologies into long-term siting and resiliency decisions to ensure compliance and protect time to power schedules.

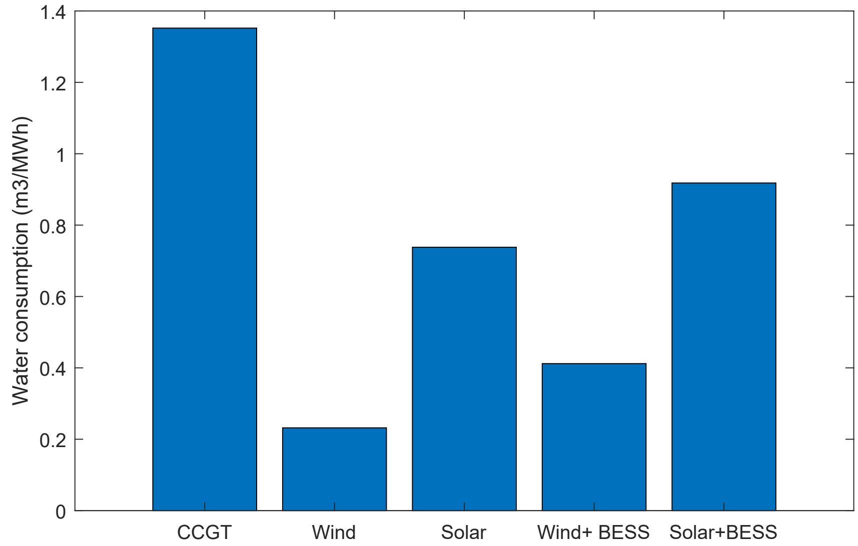


Fig. 11. Water consumption potential for the different generating technologies in West Texas [39].

Economic contributions, community and environmental impacts of data centers and on-site generation in Texas

Texas offers a conducive location for hyperscale and AI-focused data center development due to its solar, wind and natural gas abundance, availability of land, a strong fiber optic network, and affordable energy prices. Data centers create both short-term construction jobs and long-term jobs.

Long-term positions are for the ongoing operation and maintenance of the facility, but these numbers are often small, with some large data centers having fewer than 150 permanent employees. Texas data centers support a range of well-paid jobs, including IT, maintenance and security staff. It has been reported that Texas data centers account for 10% of the nation's data center workforce for Q2 2024 [41]. The data center industry contributed 485,000 jobs, $35 billion in labor income and $39 billion in GDP to Texas in 2023.

Beyond capital investment, real and personal property taxes (local sales and use taxes, which can be up to 2%) on data center facilities and equipment can provide funding for schools and emergency services. This creates a substantial economic footprint. Data centers require improvements to roads, power systems and fiber connectivity, benefiting not only the facility but also surrounding areas. Underutilized or abandoned sites can be repurposed as technologically advanced facilities. The presence of a data center will attract other technology-driven companies, resulting in a "halo effect." Sectors like e-commerce, finance and AI can have reduced latency and improved performance if set up near data centers. Texas is uniquely positioned to support the growth trajectory of AI, calling for strategic investment due to the state's business-friendly environment, energy diversity and support for innovative systems. However, the rising electrical demand is stressing the electrical grid, natural resources and the environment, leading to severe perception changes among communities.

The GHG footprint of a data center depends on the operating strategy, participation of renewables and storage, and the grid electricity mix. Carbon intensity is defined as the GHG emissions (kgCO_2eq or tonCO_2eq) per unit of energy (MWh) consumed. Virginia, with the highest number of colocation data centers, has the highest carbon intensity (199 tonCO_2eq/MWh), followed by Texas (117 tonCO_2eq/MWh). The total emissions calculated for the different types of data centers in the U.S. is 1223 tonCO_2eq/MWh ([42], [43] [44], [45]). This illustrates the significant environmental impact that Texas data centers have on the United States' overall carbon emissions, as more data centers are built in Texas.

On average, direct water consumption for data centers, including for cooling, accounts for roughly 25% of total water use, while indirect water consumption makes up the remaining 75%. Texas data centers consumed an estimated 49 billion gallons of water in 2025, a figure projected to grow to 399 billion gallons by 2030 [46].

Traditional recycling of data center e-waste involves many steps such as collection, sorting, processing and manufacturing of recycled materials. The volume, material composition and recycling complexity often make this extremely expensive and technically challenging. Hence, Texas data centers are required to dispose of e-waste through Secure IT Asset Disposition (ITAD) [47], partnering with specialized companies that handle secure transport, data destruction (wiping or physical destruction), and responsible recycling of servers, storage, and networking equipment.

Increasing land costs, decreasing local land footprint, disruption of habitats, and infrastructure conflicts are impacts of the huge land demand of data centers in Texas. Strategic siting can help mitigate these issues by positioning facilities near renewable sources. Industrial zones and tech parks are ideal data center sites, and colocation centers may be located in suburban business parks. Zoning frameworks are defined so that each data center type aligns appropriately with surrounding land use. Clear, forward-thinking zoning practices, if adopted by communities, will be best positioned to attract digital infrastructure investment.

The persistent noise from servers, HVAC systems and generators adversely affects data center staff, nearby communities, and local wildlife, prompting increased public concern and pushes for noise mitigation strategies. The average noise level around the server areas of a Texas data center can be up to 92 dB(A) from within the facility, typically at head level in a standing position [48], and ranging from around 35 dB to over 85 dB in nearby residences. Within the server racks, noise levels can reach up to 96 dB(A). OSHA and NIOSH enforce a threshold for required hearing protection at 85 dB(A) over an 8-hour time period, ensuring protection against noise-induced hearing loss [49].

The average construction cost of data centers in the U.S. ranges from $M15/MW to $M11.7/MW. Dallas and Austin data centers cost $M12.1/MW and San Antonio is slightly less at $M11.7/MW [50]. Energy, including that required for cooling, is a significant operational expense, often accounting for 30–60% of total costs in Dallas. Energy costs are highly location-dependent, influenced by grid mix, local regulations, grid infrastructure and stability. Cooling costs are weather and geography dependent. Operating costs for Texas data centers are significant, with large facilities potentially costing $10 million to $25 million annually, largely driven by electricity, infrastructure maintenance, hardware, and software.

As economic and environmental impacts scale rapidly, community acceptance and long-term sustainability concerns are shaping ERCOT planning policies and accelerating industry interest in resilient on-site and hybrid power solutions that minimize grid dependence, support ESG commitments, and preserve development timelines.

Conclusion

In Texas, there are increasing risks and challenges associated with data centers connecting large electrical loads to the ERCOT grid. The interconnection queue for large loads continues to grow, resulting in significant delays in obtaining grid capacity for new facilities. As a result, data center developers are actively seeking alternative power strategies to accelerate project timelines and reduce reliance on traditional utility connections. Texas Senate Bill 6 includes provisions for emergency curtailment of large loads and initiates evaluation of upfront transmission cost allocation for data centers, further motivating operators to pursue on-site or hybrid generation options. With some of the nation's highest solar irradiance and extensive wind resources, Texas offers substantial potential for low-cost, customer-sited renewable power.

For colocation providers and hyperscale facilities, on-site generation enables sustained operations regardless of grid status. Integrating solar, wind, combined heat and power (CHP), and medium- to long-duration storage provides the ability to optimize performance, reduce operational costs, and mitigate exposure to ERCOT market price volatility and peak-demand charges. These hybrid systems improve resilience during grid outages and enable coordinated control mechanisms that balance on-site resources with grid imports. By reducing dependency on centralized transmission infrastructure, on-site generation supports environmental, social and governance (ESG) objectives, enhances energy independence, and improves long-term cost efficiency.

Texas continues to attract a growing number of data center developments, particularly in Tier III and Tier IV markets such as Dallas, where construction costs remain lower than many other regions of the United States. While favorable capital conditions—such as available land, competitive construction costs, and strong renewable potential—support rapid growth, long-term operating economics are shaped by ERCOT interconnection requirements, environmental permitting, regulatory compliance, and workforce expenses. Therefore, developing a comprehensive energy strategy that aligns with ERCOT grid conditions and evolving policy requirements is critical. Crafting optimized hybrid or near–off-grid configurations can reduce schedule risk, improve resilience, and enhance profitability for operators across the Texas data center ecosystem

References

[1] International Energy Agency, “Electricity consumption from data centres by region (TWh),” 2025 Accessed: Sept. 10, 2025. [Online]. Available: https://www.iea.org/data-and-statistics/data-tools/energy-and-ai-observatory?tab=Energy+for+AI

[2] Shehabi, A., Smith, S.J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M.A.B., Holecek, B., Koomey, J., Masanet, E., Sartor, D, “2024 United States Data Center Energy Usage Report,” LBNL-2001637, 2024. Accessed: Sept. 10, 2025. [Online]. Available: https://eta-publications.lbl.gov/sites/default/files/2024-12/lbnl-2024-united-states-data-center-energy-usage-report.pdf?utm source=substack&utm_medium=email

[3] EPRI, “Powering Intelligence, Analyzing Artificial Intelligence and Data centre energy consumption,” 2025. Accessed: Sept. 10, 2025. [Online]. Available: https://www.epri.com/research products/3002028905

[4] EIA, “Electricity Data Browser,” 2025. Accessed: Sept. 10, 2025. [Online]. Available: https://www eia.gov/electricity/data/browser/

[5] Data Center Map, “Texas Data centers.” Accessed: Nov. 24, 2025. [Online]. Available: https://www.datacentermap.com/usa/texas/

[6] Mariam Arzumanyan, Edna Rodriguez Calzado, Ning Lin, Vaibhav Bahadur, Jani Das, Tingwei Lucy Ko, Lars Koesterke, “Geospatial suitability analysis for data center placement: A case study in Texas, USA,” Sustainable Cities and Society, vol. 131, p. 106687, 2025, doi: https://doi.org/10.1016/j.scs.2025.106687.

[7] RCR Wireless news, “Five reasons AI data centers require massive amounts of power,” 2025. Accessed: Sept. 15, 2025. [Online]. Available: https://www.rcrwireless.com/20250318/featured ai-data-centers-power

[8] IEA, “Share of electricity consumption by data centre and equipment type, 2024, IEA, Paris.” Accessed: Oct. 20, 2025. [Online]. Available: https://www.iea.org/data-and-statistics/charts/share of-electricity-consumption-by-data-centre-and-equipment-type-2024, Licence: CC BY 4.0

[9] Ko, M.S; Zhu, H, “Wide-Area Power System Oscillations from Large-Scale AI Workloads,” doi https://doi.org/10.48550/arXiv.2508.16457.

[10] D. Bills, “Reliability Series #1: Reliability vs. resilience.” Accessed: Sept. 20, 2025. [Online]. Available: https://www.microsoft.com/en-us/security/blog/2014/03/24/reliability-series-1-reliability-vs resilience/

[11] C. Baumann, “How Microgrids for Data Centers Increase Resilience, Optimize Costs, and Improve Sustainability,” 2020. Accessed: Sept. 20, 2025. [Online]. Available: https://i.nextmedia com.au/Reports/20230710114613_WP289_V1_EN.pdf

[12] Hawaii’s Cloud Solution, “Data Center Tiers Explained,” Servpac, 2023. Accessed: Sept. 21, 2025.[Online].Available:https://servpac.com/data-center-tiers-explained/#:~:text=Data%20Center%2 Classification,Tier%204:%20Faulty%20Tolerance

[13] Nlyte Software, “Fundamental Measures of Data Center Sustainability,” 2023.

[14] Association, T. G., “The Green Grid: Industry Consortium fordvancing energy efficiency in data centers,” 2007. Accessed: Sept. 20, 2025. [Online]. Available: The Green Grid website https:/ www.thegreengrid.org/.

[15] Delloro Group, “AI Power and Cooling Requirements Lift Data Center Physical Infrastructure Market to Over $46 Billion in 2028,” 2024. Accessed: Sept. 10, 2025. [Online]. Available: https:/ www.delloro.com/news/ai-power-and-cooling-requirements-lift-data-center-physical infrastructure-market-to-over-46-billion-in-2028/

[16] US Department of Energy, “Energy Act of 2020,” 2020. Accessed: Sept. 20, 2025.

[Online]. Available: https://www.directives.doe.gov/ipt_members_area/doe-o-436-1-departmental sustainability-ipt/background-documents/energy-act-of-2020
[17] US Department of Energy, "Federal Energy Management Program," 2025. Accessed: Sept. 20, 2025. [Online]. Available: https://www.energy.gov/femp/federal-energy-management program?nrg_redirect=267680
[18] N. Rasmussen, "Determining the total cost of Owneship for Data center and Network Room Infrastructure," APC by Schneider Electric. Accessed: Nov. 10, 2025. [Online]. Available: https:/ www.zones.com/images/pdf/apc_infrastruxure02_wp.pdf
[19] EIA, "Annual Energy Outlook 2025," Apr. 2025. Accessed: Sept. 20, 2025. [Online]. Available: https://www.eia.gov/outlooks/aeo/
[20] Bloom Energy, "2025 Data Center Power Report, From Gridlock To Growth: How Power Bottlenecks Are Reshaping Data Center Strategies," Jan. 2025. Accessed: Sept. 20, 2025. [Online]. Available: https://www.bloomenergy.com/wp-content/uploads/2025-Data-Center-Power-Report.pdf.
[21] B. Luyster, "Vertiv," What Is A Microgrid And How Does It Work. Accessed: Sept. 20, 2025. [Online]. Available: https://www.vertiv.com/en-us/about/news-and-insights/articles/blog-posts what-is-a-microgrid-and-how-does-it-work/
[22] Vertiv, "Strengthening Mission-Critical Microgrids with a Battery Energy Storage System," 2023. Accessed: Sept. 20, 2025. [Online]. Available: https://hub.techcentral.co.za/wp-content uploads/2024/03/vertiv-bess-data-center-microgrids-wp-en-na-sl-71129-web.pdf
[23] C. Thai and J. Brouwer, Scientific Reports, vol. 14, no. 2435.
[24] R. Tison, "Sage Geosystems to build new geothermal project to power Meta data centers." Accessed: Nov. 24, 2025. [Online]. Available: https://www.txgea.org/sage-geosystems-to-build new-geothermal-project-to-power-meta-data-centers/
[25] AP Newsroom, "Inside Three Mile Island's nuclear restart with Microsoft, now targeting 2027 launch, one year ahead of plan." Accessed: Nov. 24, 2025. [Online]. Available: https://newsroomap.org /editorial-photos-videos/detail?itemid=9766e3aeac3f4f39b06f581f43fee742&mediatype=video
[26] A n D Cable Products, "Why AI Needs Data Centers—and What It Means for Industry Growth," 2025. Accessed: Sept. 15, 2025. [Online]. Available: https://andcable.com/cable-management why-ai-needs-data-centers/#:~:text=What%20are%20the%20key%20reasons,handle%2 these%20large%2Dscale%20workloads
[27] Bloom Energy, "2025 Data Center Power Report-Mid-Year Pulse," June 2025.
[28] Public Utility Commission of Texas, "Large Load Interconnection Process Overview," Oct. 2025. Accessed: Oct. 25, 2025. [Online]. Available: https://interchange.puc.texas.gov Documents/55999_168_1550025.PDF?utm_source=substack&utm_medium=email
[29] T. Kavulla, "Remarks on Transmission Cost Allocation and Rate Design in ERCOT," NRG, Feb. 2025. Accessed: Nov. 15, 2025. [Online]. Available: https://www.nrg.com/assets/documents energy-policy/ercot-transmission-costs-and-rate-design-remarks-on-feb-25-2025.pdf#:~: text=For%20large%20industrial%20and%20commercial%20customers%2C%24CP,rates%2 that%20are%20based%20on%20energy%20consumption
[30] Texas Commission on Environmental Quality (TCEQ), "Fact Sheet - Air Permitting," TCEQ, 2021. Accessed: Sept. 20, 2025. [Online]. Available: https://www.tceq.texas.gov/assets/public permitting/air/factsheets/permit-factsheet.pdf
[31] Trinity Consultants, "Powering the Next Generation of Data Centers: Navigating Due Diligence and Air Permitting for Data Center Development," 2025. Accessed: Sept. 20, 2025. [Online]. Available: https://trinityconsultants.com/resources/powering-the-next-generation-of-data centers-navigating-due-diligence-and-air-permitting-for-data-center-development/
[32] 2025-2026, 89th Legislature, "TX SB6," 2025. Accessed: Sept. 10, 2025. [Online]. Available: https://legiscan.com/TX/bill/SB6/2025

[33] Aslam, W; Hytowitz, R, "Texas SB6 Explained: Addressing Large Load Impacts," EPRI, 2025. Accessed:Sept.20,2025.[Online].Available:https://www.epri.com/research/products/3002033410
[34] ERCOT, Firm Fuel Supply Service. Accessed: Nov. 15, 2025. [Online]. Available: https://www.ercot com/services/programs/firmfuelsupply
[35] Mordor Intelligence, "Dallas Data Center Market Size & Share Analysis - Growth Trends & Forecasts (2025 - 2030)." Accessed: Sept. 21, 2025. [Online]. Available: https://www.= mordorintelligence.com/industry-reports/dallas-data-center-market
[36] Texas Commission on Environmental Quality (TCEQ), "Permits By Rule," TCEQ. Accessed: Sept. 20, 2025. [Online]. Available: https://texas-sos.appianportalsgov.com/rules-and-meetings?chapt r=106&interface=VIEW_TAC&part=1&title=30
[37] Texas Commission on Environmental Quality (TCEQ), "Standard Air Permits," June 2025. Accessed: Sept. 20, 2025. [Online]. Available: https://www.tceq.texas.gov/permitting/air/newsourcereview standard
[38] Texas Commission on Environmental Quality (TCEQ), "Air Quality Standard Permit for Oil and Gas Handling and Production Facilities," Feb. 2018. Accessed: Sept. 21, 2025. [Online]. Available: https://www.tceq.texas.gov/permitting/air/newsourcereview/chemical/oil_and_gas_sp.html
[39] Texas Commission on Environmental Quality (TCEQ), "Emission Reduction Credit Program." Accessed: Sept. 20, 2025. [Online]. Available: https://www.tceq.texas.gov/airquality/banking/erc program.html
[40] Das, J., Ur Rehman, A., Verma, R., Gulen, G., & Young, M. H., "Comparative Life-Cycle Assessment of Electricity-Generation Technologies: West Texas Case Study," Energies, vol. 17, no. 5, p. 992, 2024, doi: https://doi.org/10.3390/en17050992.
[41] United States Census Bureau, "Quarterly Workforce Indicators (QWI) (Time Series: 1990 - present)." Accessed: Sept. 21, 2025. [Online]. Available: https://www.census.gov/data/developers/data-sets qwi.html
[42] Gianluca Guidi, Francesca Dominici, Jonathan Gilmour, Kevin Butler, Eric Bell, Scott Delaney, Falco J. Bargagli-Stoffi, "Environmental Burden of United States Data Centers in the Artificial Intelligence Era", doi: https://doi.org/10.48550/arXiv.2411.09786.
[43] Md Abu Bakar Siddik, Arman Shehabi and Landon Marston, "The environmental footprint of data centers in the United States".
[44] United States Environmental Protection Agency, "Emissions & Generation Resource Integrated Database (eGRID)." Accessed: Sept. 21, 2025. [Online]. Available: https://www.epa.gov/egrid
[45] "States Most Impacted by Data center pollution," Known Host Blog. Accessed: Sept. 21, 2025. [Online]. Available: https://www.knownhost.com/blog/states-most-impacted-by-data-center pollution/
[46] M. Spenrath and A. Datta, "POWERING TEXAS' DIGITAL ECONOMY: DATA CENTERS AND THE FUTURE OF THE GRID," Houston Advanced Research Center. Accessed: Nov. 14, 2025. [Online]. Available: https://harcresearch.org/wp-content/uploads/2025/05/Part-2.-Powering-Texas-Digital Economy-Final.pdf
[47] Texas Commission on Environmental Quality (TCEQ), "Electronics Recycling." [Online]. Available: https://www.tceq.texas.gov/p2/recycle/electronics
[48] D. Miljković, "Noise within a data center," presented at the 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2016, pp. 1145–1150. doi: 10.1109/MIPRO.2016.7522312.
[49] Tech Target, "Understanding the Impact of Data Center Noise Pollution, 2024," 2024. Accessed: Sept. 21, 2025. [Online]. Available: https://www.techtarget.com/searchdatacenter/tip Understanding-the-impact-of-data-center-noise-pollution
[50] Cushman and Wakefield, "2025 Data center development cost guide." [Online]. Available: https:/ www.cushmanwakefield.com/en/united-states/insights/data-center-development-cost-guide.

Contributions

Author	Role	Contributions
Dr. Jani Das	Lead Author	Original manuscript writing and editing, literature review and analysis
Dr. Ning Lin	Principal Investigator	Supervision, manuscript review and editing, validation
Dr. Yashvi Malhotra	Technical Contributor	Manuscript review and editing
Shipra Mishra	Technical Contributor	Manuscript review and editing
Dr. Mariam Arzumanyan	Technical Contributor	Manuscript review and editing
Dr. Pablo Paz	Technical Contributor	Manuscript review and editing

Acknowledgments

The authors would like to thank many colleagues, collaborators, and reviewers whose insights, analyses, and contributions have shaped this white paper. We are especially grateful to Dr. Lorena G. Moscardelli, Dr. Mark Shuster, David Chapman, and Mark Blount of the University of Texas at Austin for their support and guidance. We also thank Shannon Strank, Robert E Hebner and other colleagues at the Center for Electromechanics for their technical expertise and critical input. We would like to thank the Bureau's Media Production Services team for their support in designing the report. We also acknowledge editorial and review assistance, as well as the guidance and support in communicating this work. This report would not have been possible without the support of our extended research team and institutional partners.

Contact: compass@beg.utexas.edu

COMPASS
Collaborative Optimization & Management of Power Allocation – Surface & Subsurface strategies