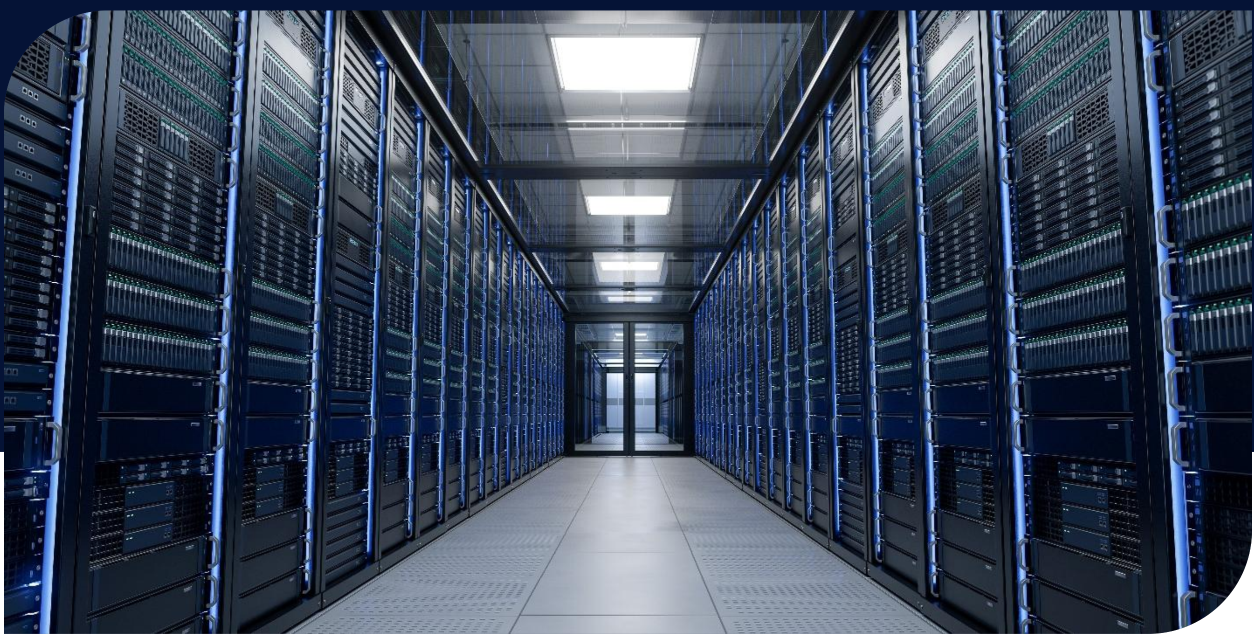
Teneo

How Data Centers are Remaking the U.S. Electricity Sector

Part One: The Surge in Demand

Teneo Insights | September 2025



The U.S. electricity sector is entering a period of transformation unlike anything in its history. At the heart of this change is the explosive growth of data centers, driven by AI, cloud computing and digital infrastructure. This is the first article in a three-part series, exploring how this unprecedented demand is reshaping the grid – from the unique characteristics of today's load growth to the innovations utilities and policymakers are deploying, and the long-term implications on prices, markets and regulation.

Utilities and independent power producers have been among the biggest beneficiaries of the excitement surrounding the "AI revolution." The surge in energy demand has created significant opportunities for investment in energy and energy infrastructure, spanning gas pipelines, power generation, transmission and substations.

Load growth, now measured in percentage points rather than basis points, has returned to the electricity sector. This latest wave of growth in the U.S. power system is unlike anything in the nation's history. The scale is truly unprecedented. During the zenith of U.S. power system expansion following World War II in the 1950s, power generation was growing at over 9% per year. Yet the absolute amount of new generation capacity added was less than 10 GW. By contrast, the U.S. Department of Energy now estimates capacity additions of ~70 GW or higher over the coming decade. Even more striking than the scale of these additions is the mismatch between the speed of interconnection demanded by data center providers (measured in months) and the deliberate pace typical of this highly capital-intensive, long-lived asset business—a pace that carries unintended consequences for both public safety and national security.

Growth at this scale and speed is driving significant innovations in the design, operation and regulation of the U.S. electricity sector. Additionally, it is putting significant pressure on the core principles that have underpinned electricity market design and utility ratemaking since the 1990s and earlier.

Below, we summarize some of the emerging changes driven by the surge in data center-driven demand growth in the U.S. These efforts focus on (a) accelerating interconnection through both demand- and supply-side measures and (b) mitigating system impacts through ratemaking reforms, cost allocation adjustments and changes to fundamental market design. We begin by isolating what's distinctive about this latest wave of demand growth, and we conclude by outlining what these changes could mean for the sector moving forward.

What's Different About Data Center Load Growth?

After nearly two decades of modest electricity demand growth (+0.1% from 2005–2020; +1.7% from 2020–2026),[1] U.S. power system planners are now contending with a sharp, concentrated rise in large-load interconnection requests—driven primarily by cloud, high-performance computing, AI and crypto mining operations. Unlike traditional load growth, which tends to follow demographic or economic trends, this surge is driven by private-sector capital allocation decisions and digital infrastructure buildout, often decoupled from historical energy consumption patterns.

The demand influx expected over the coming decade is unprecedented in its sheer scale. And it is occurring in a regulatory environment that is much more challenging than faced by previous generations or in less litigious societies.[2] Arguably, the most significant characteristics include (a) its density, with implications for severe impacts on the overall grid; (b) the high value of the load and the associated opportunity cost of delays; and (c) the uncertainty of its ultimate magnitude given the high pace of technical change and the unproven rate of adoption.

Density

Unlike historical load growth, which was broadly distributed across regions, the current surge in demand—driven largely by hyperscale and AI-focused data centers—is intensely concentrated in a small number of locations that combine telecom connectivity, land availability and regulatory favorability. While the locations may vary across the U.S.—including Northern Virginia, Central Ohio, West Texas and the Texas Panhandle—the simultaneous arrival of 50–300 MW projects represents an unprecedented level of demand within a single transmission planning zone. This frequently strains local substations and feeders, accelerates the development of transmission and generation capacity and overwhelms state and ISO administrative processes.

[1] EIA. (2025). Monthly Energy Review and Short-Term Energy Outlook
[2] Mark Dunkelman (2025): Why Nothing Works: Who Killed Progress – And How to Bring It Back, Dan Wang (2025), Breakneck: China's Quest to Engineer the Future

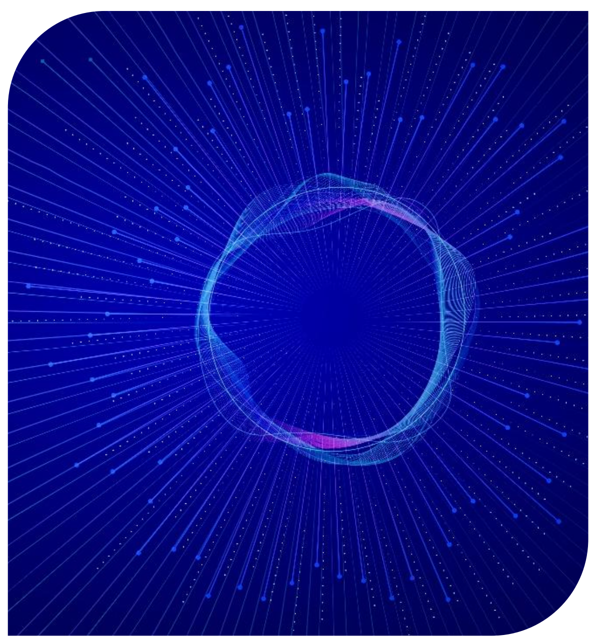
Urgency

At an economy-wide level, U.S. GDP generated per kWh of electricity consumed more than doubled from 1980 to 2020. However, over the past 50 years, the value produced by electricity varies dramatically by sector.[3] When measured by annual profit per kWh consumed, data center providers such as Meta and Google generate roughly 25 times more value than energy-intensive industries like steel producers such as Nucor. Alternatively, given the high capital intensity of new data centers (i.e., $10/w to $15/w for advanced designs), a day's delay in powering up a 100 MW data center costs between $2.7 million and $4.1 million in unproductive capital.[4] Combined with the intense competition for achieving advanced general intelligence—and, at a minimum, capturing market share in crypto and data center hosting—the urgency of reducing delays in securing access to power becomes clear. Developers are typically targeting interconnection within 24–36 months, dramatically shorter than the 5–10 years often required for associated transmission and distribution upgrades. For example, in Northern Virginia, Dominion Energy reports that 75% of substations in its data center zones are already at or near thermal limits, with less than 20 MW of headroom on average. This alone could push upgrade lead times to 8–10 years. In Texas, Oncor and AEP have received over 40 GW of large-load inquiries since 2023, much of it speculative and concentrated in the Panhandle region and West Texas.[5]

Uncertainty

As an asset class being deployed with unprecedented speed and strategic focus, data centers present unusual challenges for projecting future growth rates and load profiles. These difficulties stem from their heterogeneity, the inherent unpredictability of demand for the services (most notably AI-related offerings) and the rapid pace of efficiency improvements.

Data centers differ greatly in three key ways: their size, how sensitive they are to latency (and therefore how close they need to be to users) and how much flexibility they have to adjust their power consumption. The most consequential distinctions among these facilities are their uptime, load factor and utilization. Hyperscale data centers, operated by firms like Amazon Web Services, Google Cloud and Microsoft Azure, typically exhibit high utilization (>90%, depending on how much reserve capacity is built into the facility) and continuous 24/7 runtime. These facilities support cloud infrastructure and enterprise compute workloads that must be available on demand across global markets. They are often sited near high-voltage transmission and redundant fiber routes, and their interconnection requests are formalized well in advance.[6] Colocation or multi-tenant data centers, such as those operated by Equinix or Digital Realty, host a wide range of clients with varied load profiles. While the operator may have contractual control over power procurement, it rarely has direct authority to



[3] Increasing from ~$2.60 to ~$5.50 in constant (PPI) dollars from 1980 to 2020. EIA. (2025). Electricity Data Browser
[4] Thanks to Brett Galura for this approach to illustrating the value of expediting rapid interconnection
[5] Dominion Energy. (2023). *Northern Virginia Transmission Planning Update*; Oncor Electric Delivery. (2025). *ERCOT Load Interconnection Task Force Briefing Memo*
[6] RMI (2025): Fast, Flexible Solutions for Data Centers

curtail or shift demand. This limits the feasibility of load orchestration, despite nominal similarities to hyperscale architecture.[7] Enterprise-owned data centers—typically smaller facilities operated by corporations for internal IT needs—exhibit diverse duty cycles. Though often overlooked in forward planning, they still represent a significant portion of legacy load and can affect regional demand forecasts, especially in Tier 2 and Tier 3 metro areas.[8]

Cryptocurrency mining operations present a radically different profile: highly price-sensitive, intermittently connected and physically mobile. These facilities have demonstrated high degrees of dispatchability, particularly in ERCOT, where miners routinely curtail during peak price intervals or scarcity events.[9] The category attracting the greatest attention among policymakers and the press is dedicated AI and machine learning (ML) compute campuses, sometimes independent of traditional hyperscale architectures. These facilities operate at enormous scale, frequently exceeding 200–500 MW and are used for high-density training and inference workloads. Some exhibit flexibility through batch training schedulers, but their scale and unpredictable timing (e.g., alignment with model-release cycles) present the greatest overall challenge to long-term system planning.

Forecasts for AI-based demand have shifted repeatedly as AI adoption rates, workload mixes and hardware/software efficiency gains evolve. The energy intensity of AI training, inference and reasoning workloads can differ by an order of magnitude, and the ratio among them in deployed systems is still in flux. EPRI's 2024 analysis estimates that use-phase inference already accounts for roughly 60% of AI's total energy footprint, compared with 30% for training and 10% for model development, and that a single ChatGPT request consumes about 2.9 watt-hours, nearly 10 times the energy of a traditional web search.[10] While per-query efficiency may improve over time, EPRI and other researchers caution that aggregate demand could still rise if model sizes, release frequency or user adoption accelerate faster than efficiency gains. Conflicting projections, ranging from exponential growth to near-term plateau, leave utilities and system operators navigating between the risks of under-building critical infrastructure and stranding assets through premature overbuilding.

[7] Oregon PUC (2024), Docket UM 2257: Investigation into Large Load Classes and Tariff Structure
[8] LBNL (2024): 2024 US Data Center Energy Usage Report
[9] EIA (2024): Data Centers and Cryptocurrency Mining in Texas Drive Strong Power Demand Growth
[10] EPRI (2024): Powering Intelligence

Authors

Dan Gabaldon
Vice Chair, Management Consulting and Head of Energy

Bob Zabors
Vice Chaiman, Management Consulting and Head of Infrastructure & Innovation

How Data Centers are Remaking the U.S. Electricity Sector

This is the first article in a three-part thought leadership series from Teneo's Energy and Infrastructure team, exploring how the rapid expansion of data centers is transforming the U.S. electricity sector. Drawing on real-world developments, the series brings together perspectives on technology, regulation and market design. The next article will highlight how utilities, regulators and developers are deploying new tools and approaches to accelerate interconnections and manage system strain.

Teneo
The Global CEO Advisory Firm

Teneo is the global CEO advisory firm.

We partner with our clients globally to do great things for a better future.

Drawing upon our global team and expansive network of senior advisors, we provide advisory services across our five business segments on a stand-alone or fully integrated basis to help our clients solve complex business challenges. Our clients include a significant number of the Fortune 100 and FTSE 100, as well as other corporations, financial institutions and organizations.

Our full range of advisory services includes strategic communications, investor relations, financial transactions and restructuring, management consulting, physical and cyber risk, organizational design, board and executive search, geopolitics and government affairs, corporate governance and ESG.

The firm has more than 1,700 employees located in 45+ offices around the world.

teneo.com