U.S. data center update

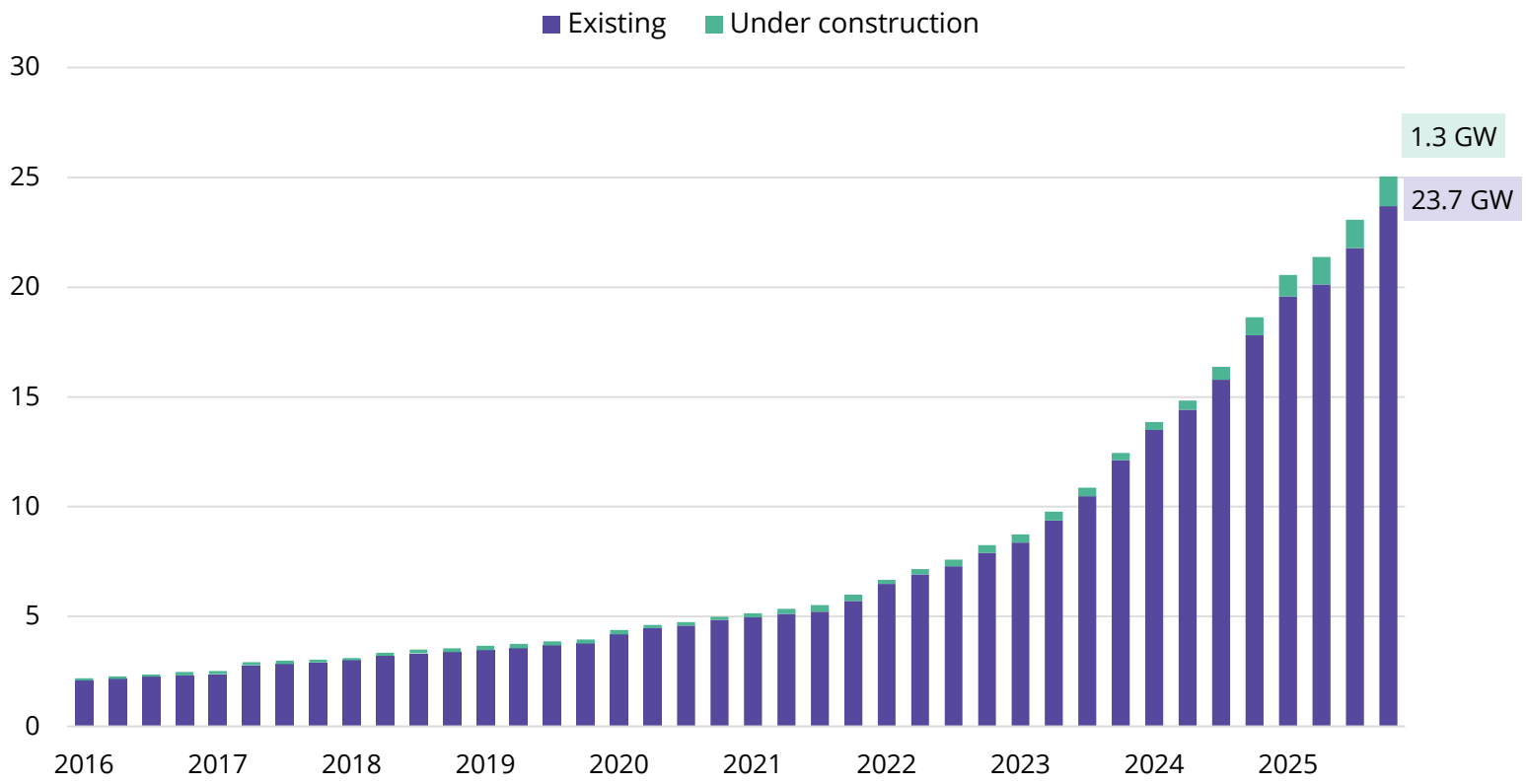
Leasing and Capital Markets trends

Q4 2025

AVISON YOUNG

Inventory

Colocation inventory by quarter (GW)



Note: Inventory includes U.S. primary and secondary markets.
Source: DatacenterHawk

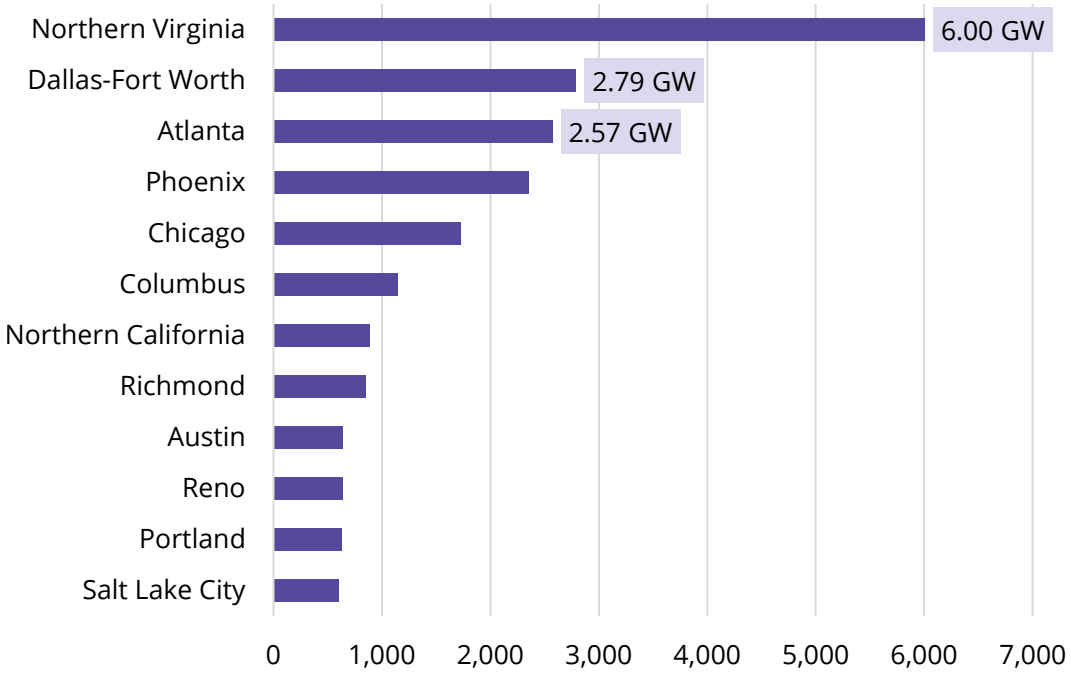
Colocation inventory rose 8.8% in Q4 2025, closing the year at 23.7 gigawatts (GW), equating to a 33% year-over-year (YoY) increase in existing inventory.

Following a minor decrease in new development in the first half of 2025, the second half recorded a notable uptick in new development, finishing Q4 2025 with 1.3 GW of new power actively under construction.

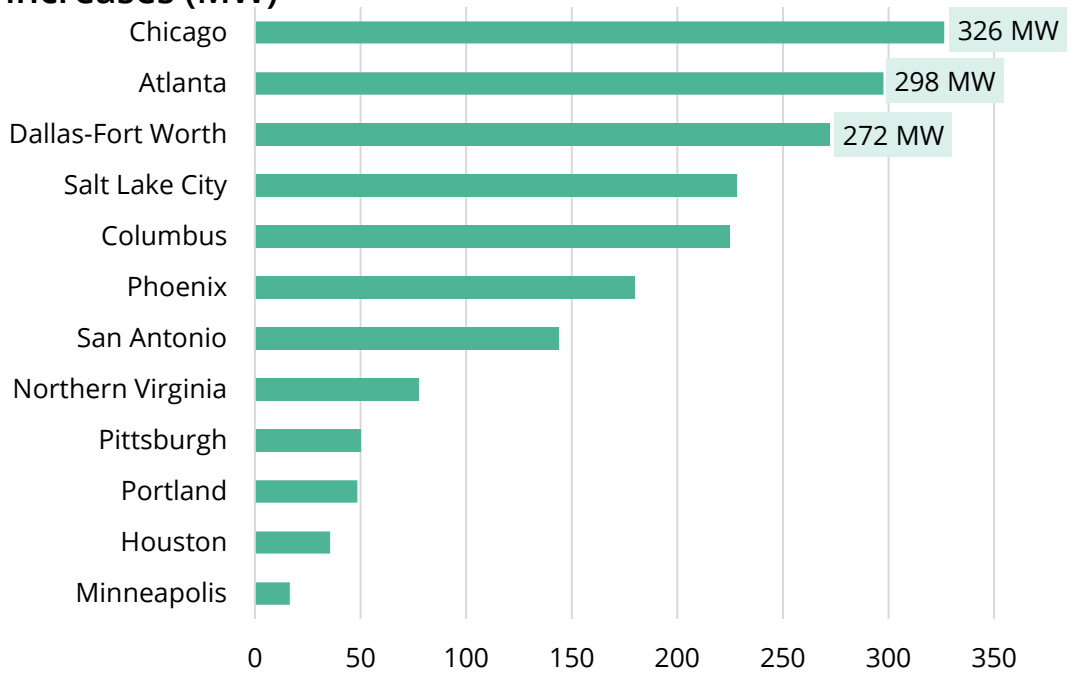
Inventory

In Q4 2025, Chicago led all U.S. markets in new deliveries, commissioning 326 megawatts (MW) of new power, with Atlanta close behind at 298 MW. On a percentage basis, the Pittsburgh market recorded a 100.3% increase — the only market of the quarter to record triple-digit inventory growth. Northern Virginia remains the standout leader in total existing inventory, while 2025 marked a record-breaking year of inventory growth across U.S. markets. Richmond continued to benefit from demand spilling over from Northern Virginia, delivering 848 MW of new commissioned power after beginning the year as a sparsely developed market.

Colocation inventory of largest U.S. markets (MW)



Markets with largest quarter-to-date (QTD) inventory increases (MW)

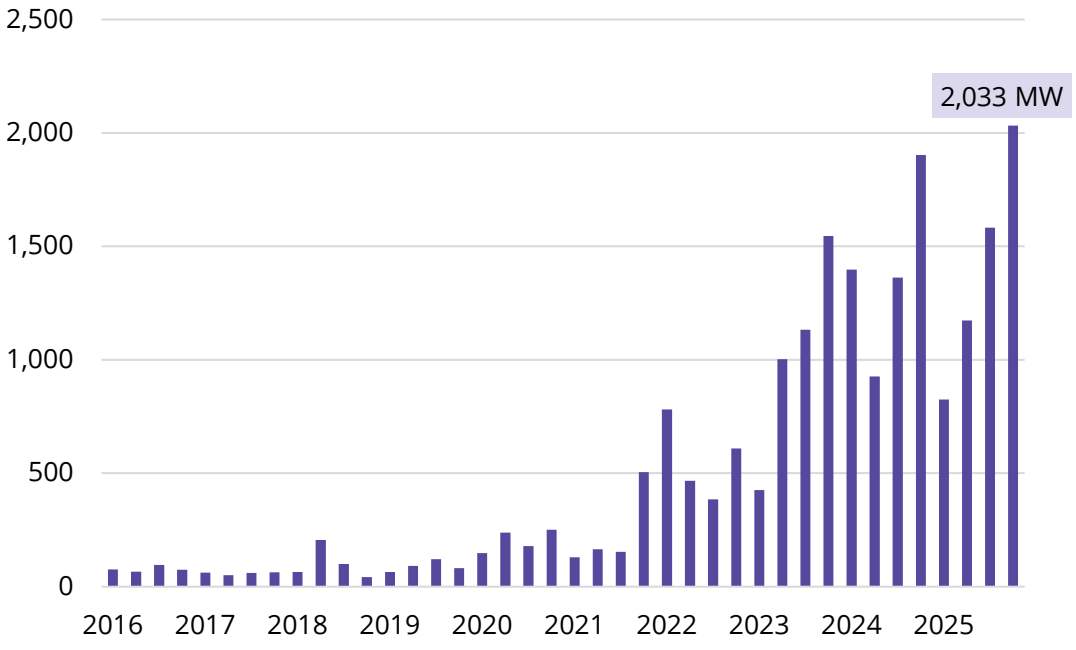


Note: Includes facilities that have delivered or have a committed tenant.
Source: DatacenterHawk

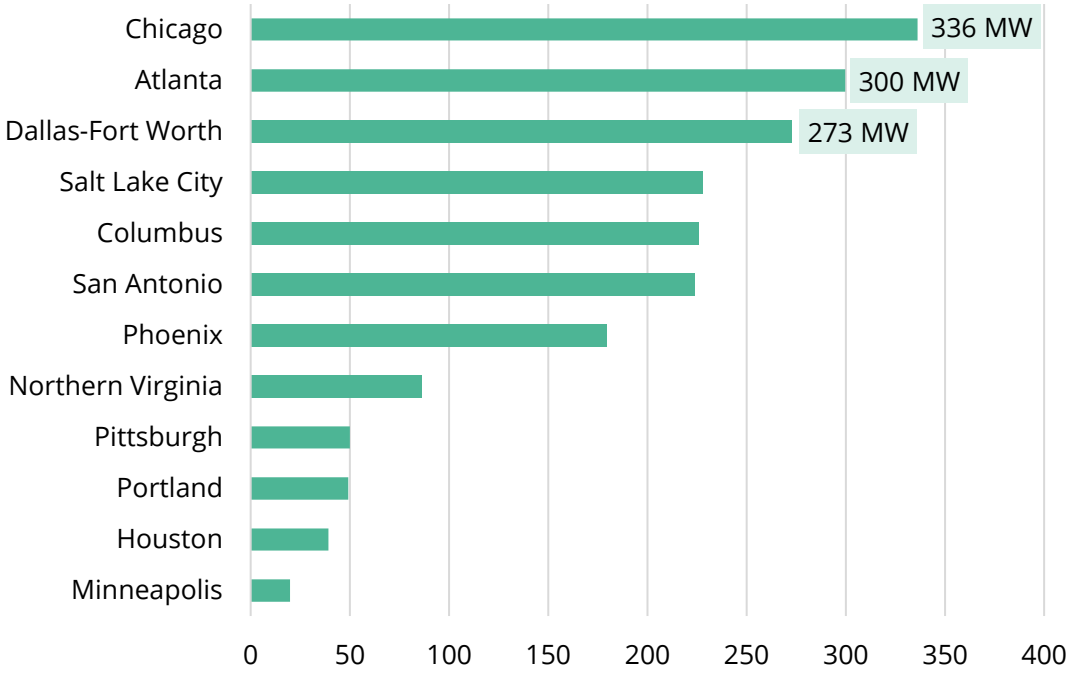
Absorption

Q4 2025 experienced the highest level of quarterly net absorption on record, with 2.03 GW in positive absorption across U.S. data center markets. Chicago, Atlanta, and Dallas-Fort Worth led absorption activity, with Atlanta and Chicago each surpassing 300 MW of positive absorption — an impressive feat for a single quarter. Pittsburgh, which began the year with relatively limited colocation inventory, closed out 2025 with 50 MW of positive absorption, a welcome sign for the emerging market.

Colocation net absorption (MW)



Highest net absorption markets (Q4 2025)

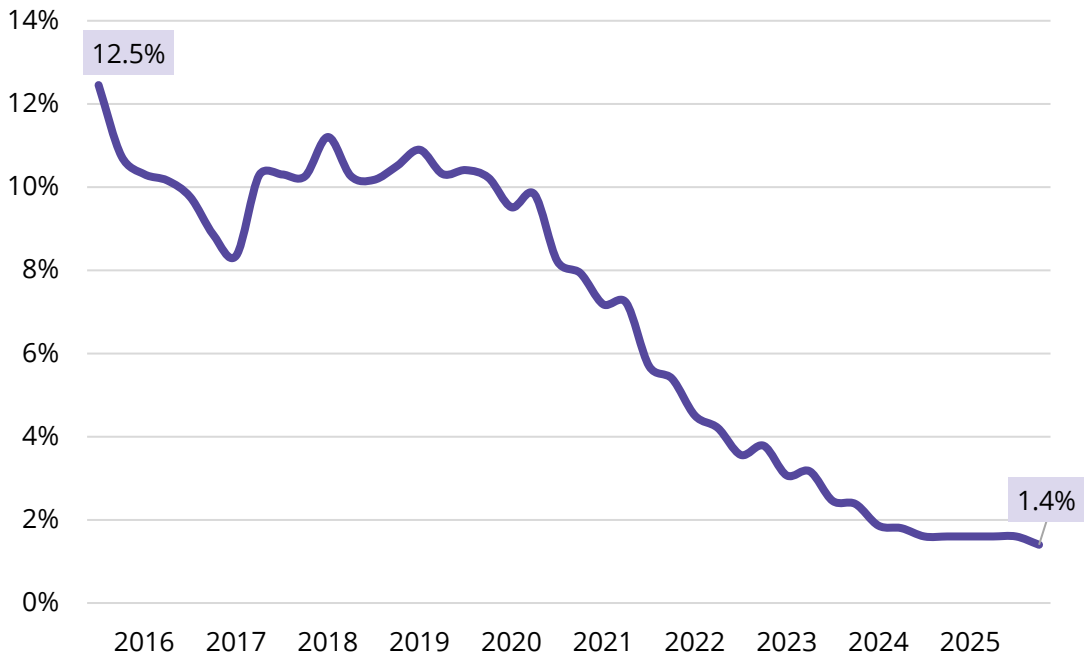


Source: DatacenterHawk

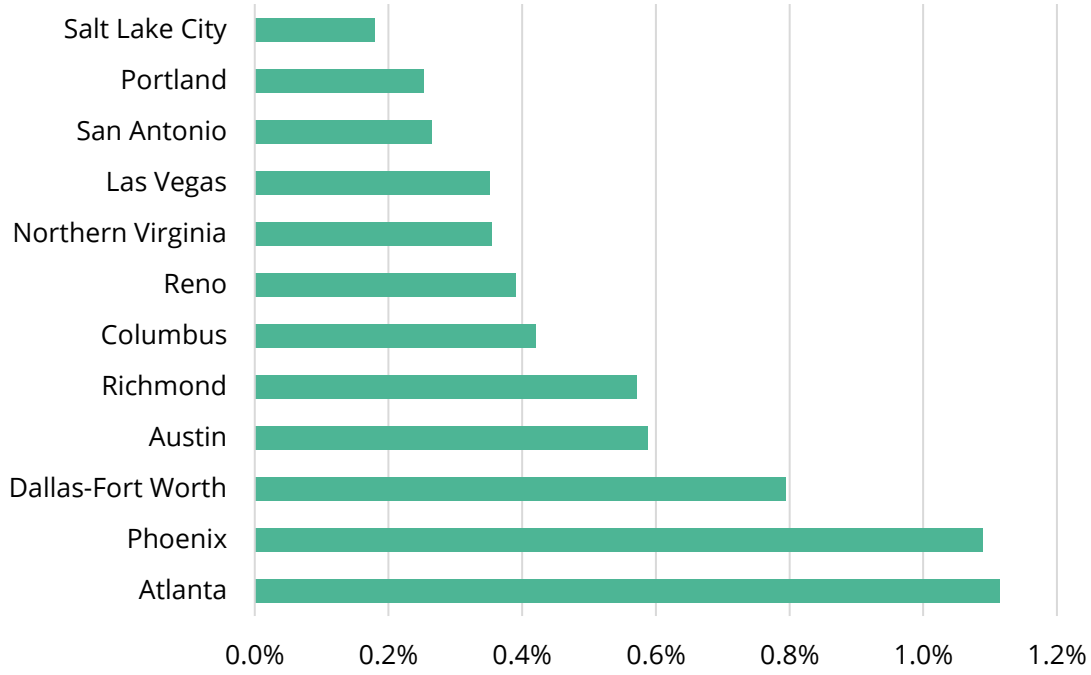
Vacancy

Data center demand continues to build strong momentum, with steady tenant demand across both established and emerging markets. While established markets like Northern Virginia and Columbus maintain historically low vacancy rates, that tightness is pushing more demand into emerging markets that offer high-quality turnkey facilities at much lower strike prices.

Vacancy rate by quarter



Lowest vacancy markets (Q4 2025)



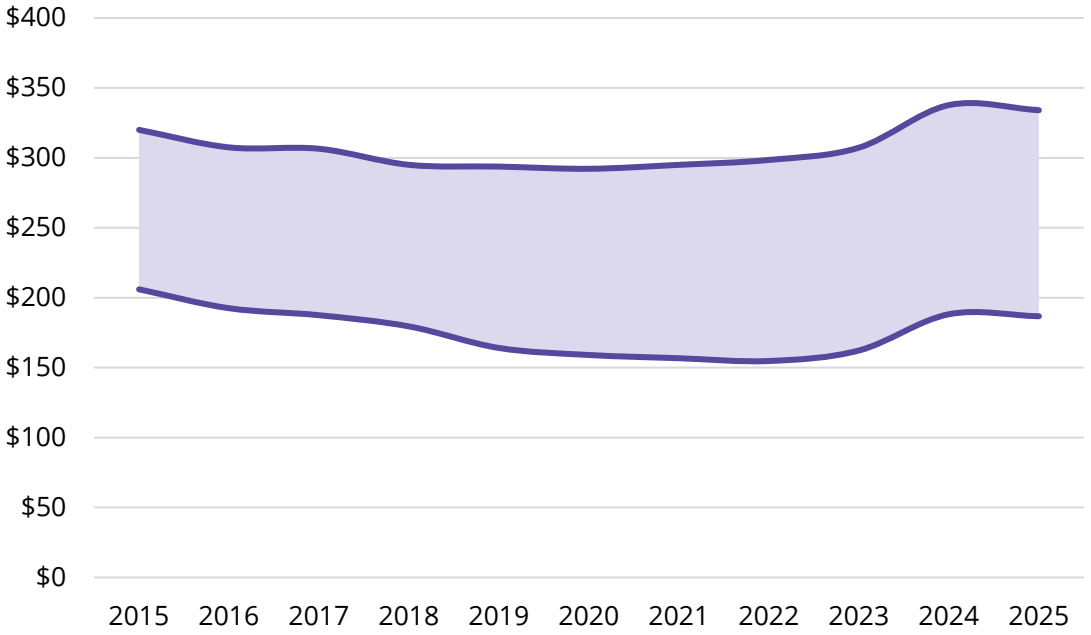
Top 12 North American markets; includes all facilities that have delivered or have a committed tenant.
Source: DatacenterHawk

Lease pricing

Turnkey colocation rents remained relatively flat to close the year, while pricing for large-scale requirements recorded a modest increase on both the top and bottom-line averages.

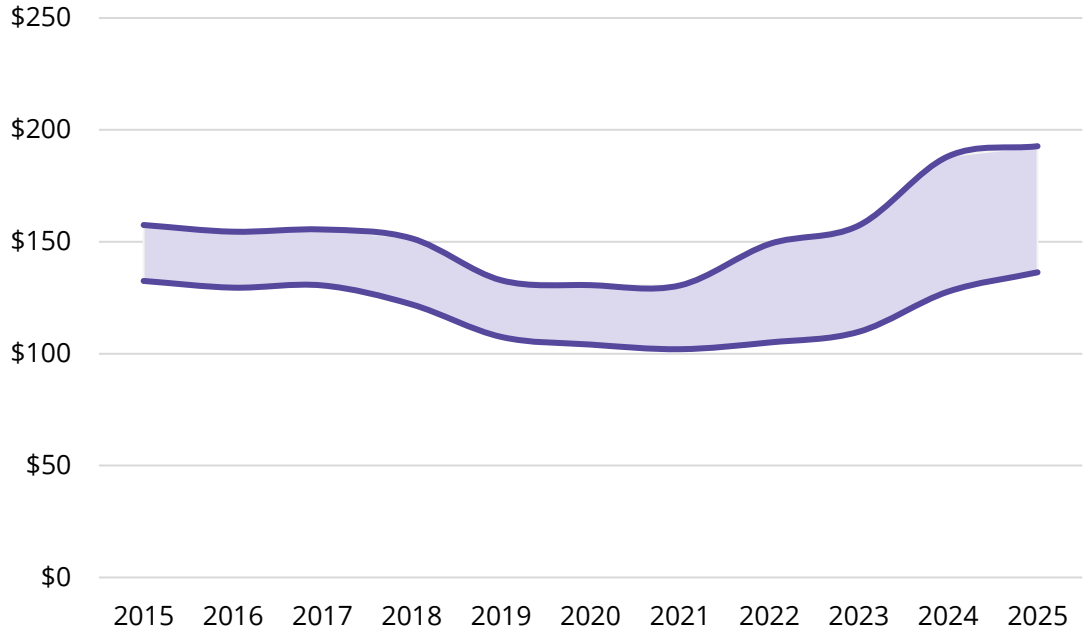
Average retail range ($/kW)

Leases under 250 kW



Average wholesale range ($/kW)

Leases between 250kW – 4 MW



Source: DatacenterHawk

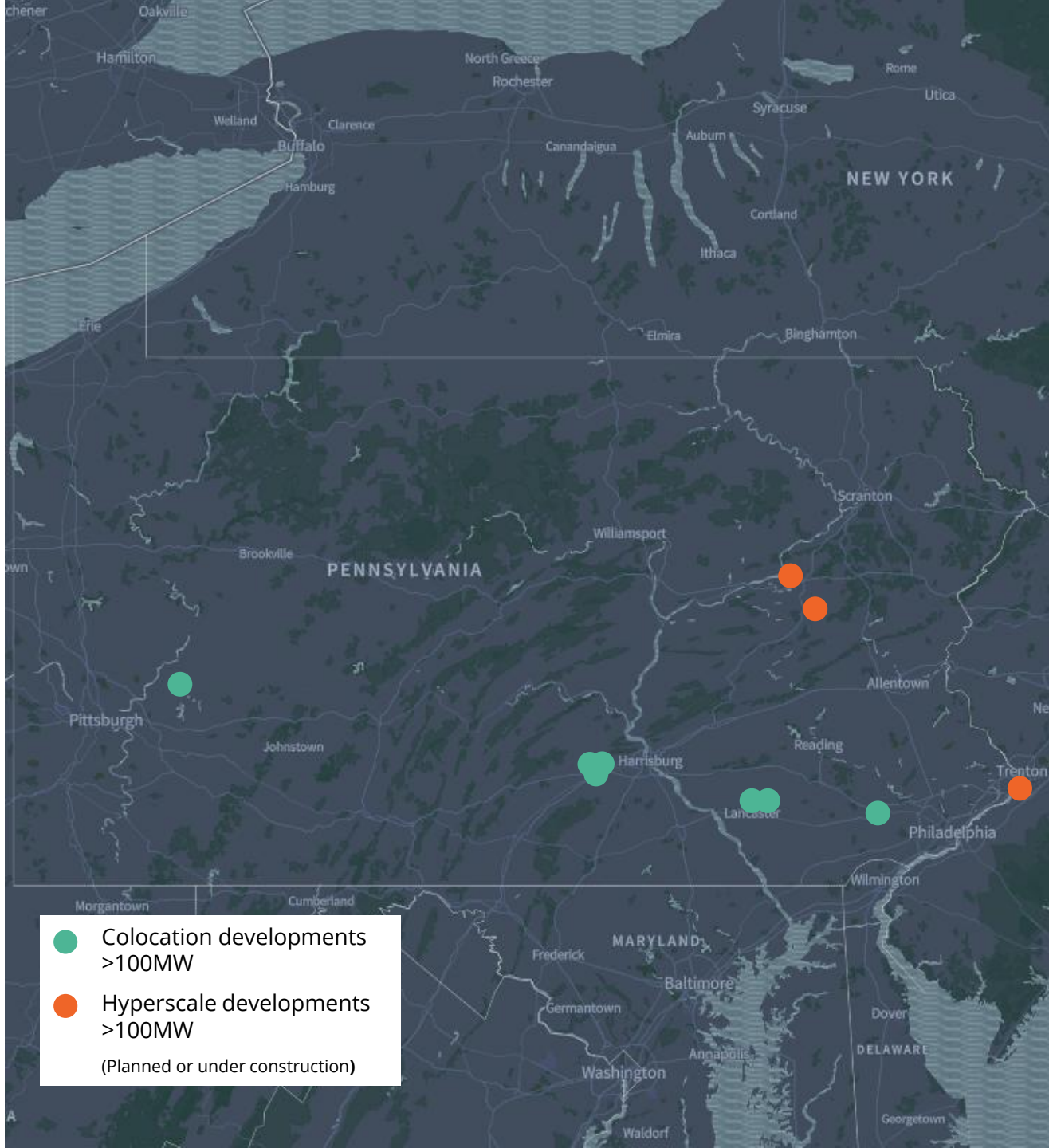
Market highlight: Pennsylvania

PA SB 939 (PN 1424): Artificial Intelligence and Data Center Act

In our Q3 2025 report, we highlighted how Pennsylvania is quickly becoming a landing spot for both hyperscaler and large-scale colocation projects. Bolstering this surge in development is the proposed PA State Bill 939, introduced in July 2025. The bill aims to achieve the following objectives: establish an Office of Transformation and Opportunity; create an AI, Data Center, and Emerging Technology Regulatory Sandbox Program; and implement a permitting and siting framework for High-Impact Data Centers, including incentives for projects that utilize the "Bring Your Own Generation (BYOG)" power strategy.

The bill also includes legislation for the Commonwealth Opportunity Zone, a state regulatory relief designation intended to make Pennsylvania more attractive for AI and data center investment.

If passed, the bill would streamline and accelerate the permitting process, establish a centralized resource for data center owners and developers to communicate with the state, and create incentives for the development of high-impact data centers across Pennsylvania.



Source: DatacenterHawk, PA SB 939 (PN 1424): Artificial Intelligence and Data Center Act, Avison Young Market Intelligence

Capital market trends

AVISON YOUNG

Capital markets trends

More M&As

There was a notable uptick in M&A deals across the data center sector in the second half of 2025. Large investments by both public and private groups continued to bolster the sector's strong growth.

Power plays

Amid soaring demand for energy, developers and data center operators are moving upstream, targeting partnerships, joint ventures, and acquisitions directly within the power sector.

New players

As the sector matures, both private equity and institutional capital are increasingly entering the market. Their growing participation is expected to continue throughout 2026, capturing a larger share of data center ownership.

Speculative developers

Speculative developers hoping to capitalize on rising data center demand are encountering significantly higher barriers to entry than in other asset classes. Challenges such as power availability, community buy-in, and supply–chain constraints are pressure-testing many developers who are new to the sector.

Major M&A deals

December 2025

Alphabet announced a definitive agreement to acquire Intersect, a provider of data center and energy infrastructure solutions.

December 2025

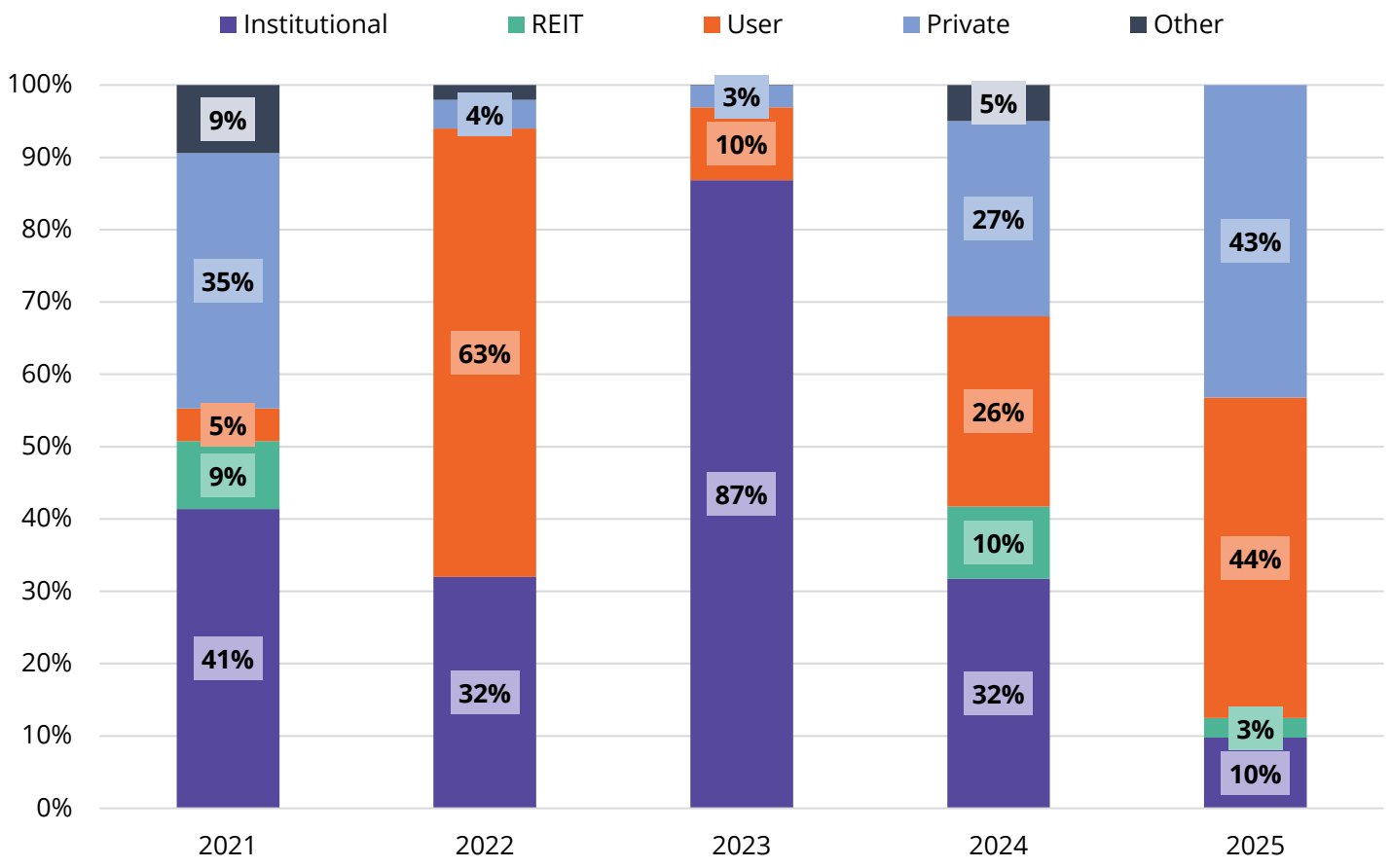
SoftBank Group announced a definitive agreement to acquire Digital Bridge, a global alternative asset manager dedicated to investing in digital infrastructure, including data centers, cell towers, fiber networks, and edge infrastructure.

December 2025

Vertiv Holdings Co., a global provider of critical digital infrastructure, announced the completion of its acquisition of Purge Rite Intermediate LLC, a leading provider of mechanical flushing, purging, and filtration services for data centers and other mission-critical facilities.

Buyer composition

Data center buyer composition

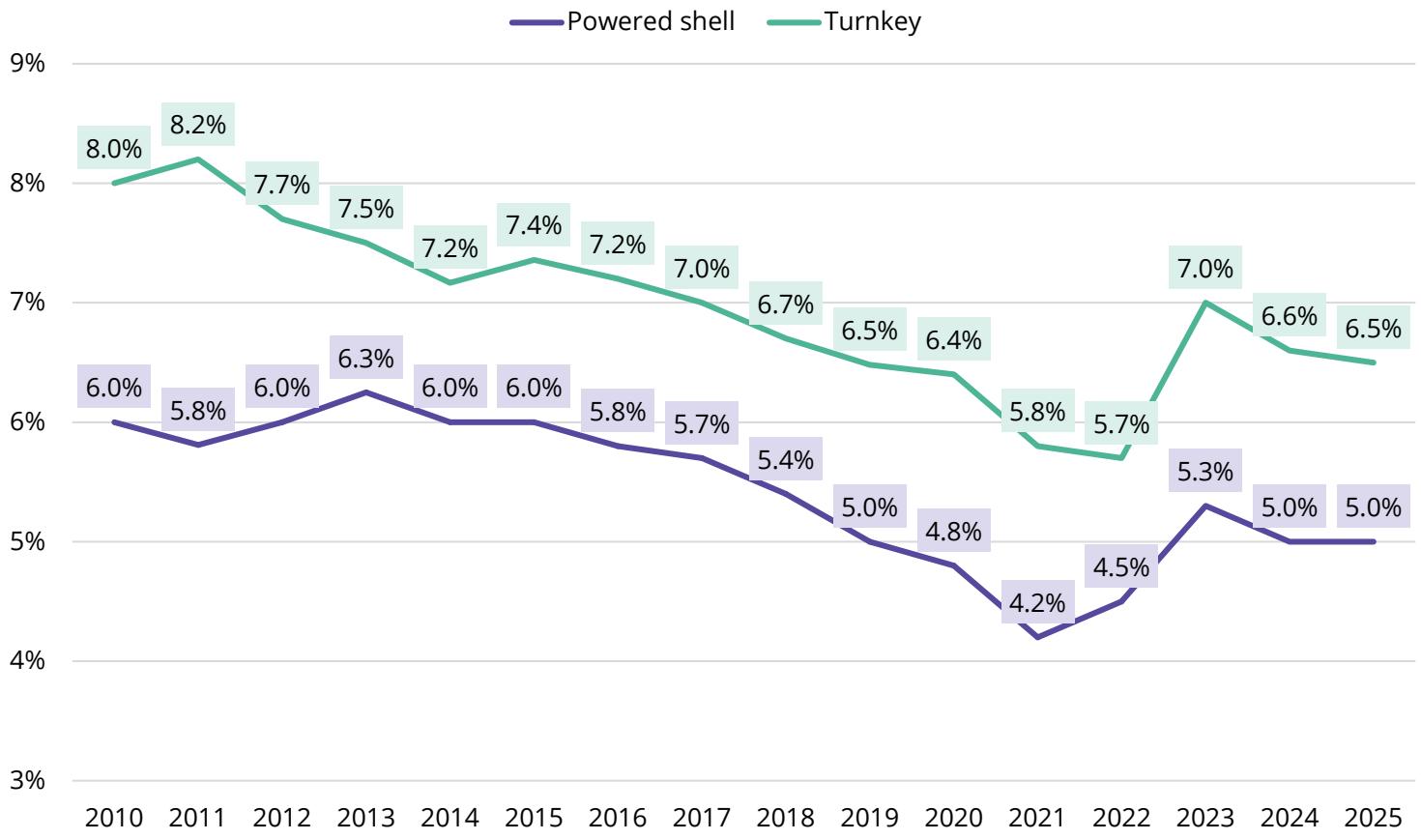


The buyer composition of data centers shifted throughout 2025 in response to the sharp rise in user demand for additional capacity. User buyers accounted for 44% of total sales volume, while private buyers followed closely at 43% of acquisitions, as more private firms sought to enter the expanding data center sector.

Source: Avison Young Market Intelligence

Cap rates

Typical data center cap rate



Note: Typical cap rates are not necessarily averages of all market transactions. To handle small sample sizes, we removed outliers and/or corrected for deal nuances that may impact cap rates.
Source: Avison Young Market Intelligence

Cap rates for data centers are holding in the mid-6% range for turnkey facilities and around 5% for powered shell properties. The Fed cut interest rates twice, by 25 basis points (bps) each time, bringing the current federal funds rate target range to 3.5%–3.75%.

Overall, declining interest rates are expected to spur increased investment in the data center sector, potentially bringing sidelined investors back into the market as they await lower costs of capital.

Looking ahead

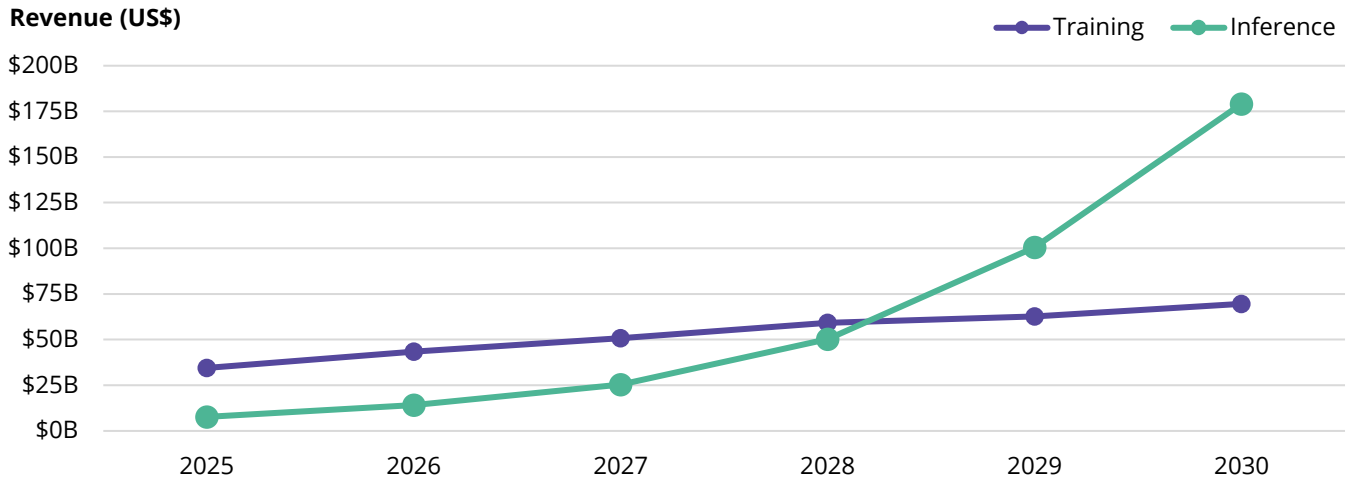
AVISON YOUNG

Neoclouds

Neoclouds (also known as GPU clouds) have emerged as a response to the sharp rise in demand for GPU intensive workflows, like AI training and inference. They are typically marketed as **GPU-as-a-service** and can comprise of thousands of GPU units bundled together with the ability to scale in real time as needed. A more cost-effective option compared to a generalized hyperscale solution, they are known to provide as close to the same compute performance as if the units were on-prem. As they support rapid provisioning, they are best for bare-metal or near bare-metal performance, high-throughput networking, and provide simpler/clearer pricing.

Neoclouds have become a clear development trend throughout the data center sector, as they are essentially AI-optimized capacity buyers/developers. They are driving demand for high power density, robust interconnect fabrics, and scalable-power sites with faster time-to-market, often using modular/prefabricated builds and colocation/wholesale deployment strategies.

Neocloud GPUaaS revenue by workload: 2025 to 2030

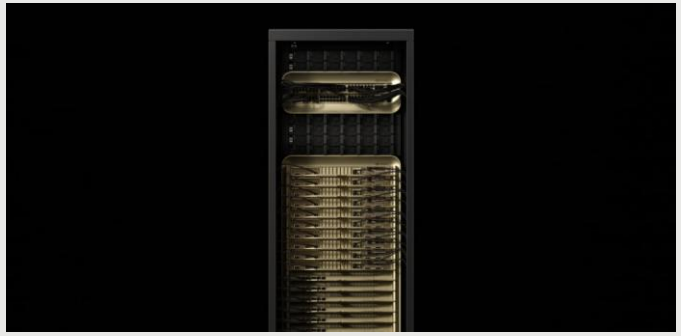


Neoclouds deployments

CoreWeave has Neoclouds deployed across multiple markets, including New Jersey and Texas. These dense clusters are mainly comprised of liquid-cooled H100/H200 NVIDIA GPUs, with the ability to upgrade to NVIDIA next-gen GPUs as they become available.

Lambda has deployed GPU clusters in Nevada and Texas, utilizing a broad range of NVIDIA chips, including A100, H100, H200, A6000, GH200, and GB200/300.

Crusoe Cloud has Neoclouds deployed near renewable energy hubs, utilizing energy-optimized scaling with NVIDIA H100/H200 chips.



NVIDIA GB300 nvl72, rack-scale and liquid cooled

Source: ABI Research (MD-NEO0101)

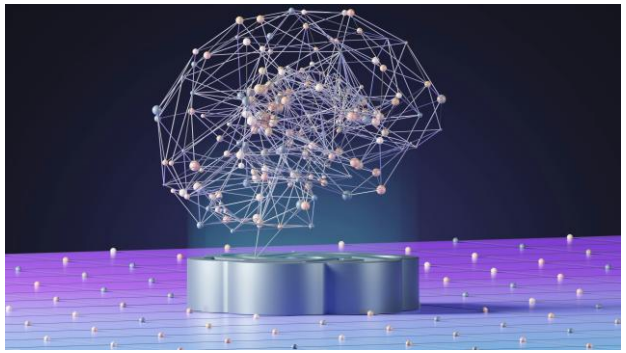
AI training → AI inference

Many industry experts expect AI to make up a significant — and in some cases dominant — share of a data center's overall workload. As more AI models transition from training to inference, the demand for low-latency, edge deployments has shined a light on how data centers are actually being utilized. While traditional workloads account for the majority of current data center workload capacity, AI inference is forecasted to overtake AI training as the main workload within AI data centers, equating to more than 50% of total AI compute capacity, and 30–40% of total data center demand.

This next evolution in data center workloads is forecasted to increase data center supply and power capacity from 30 GW in 2025 to 90 GW or more by 2030, equating to a compound annual growth rate (CAGR) of 22%. This surge in power demand is not expected to slow in the short term, as more LLMs continue to evolve from training to inference workloads.

Finding the balance between power intensive and low-latency workloads

As developers plan for future demand, it's critical to consider how AI workload will evolve, and specifically how the power required for these workloads will change. AI training benefits from the massive amount of power hyperscale that data centers can provide at a moment's notice, with less importance around latency and user interaction. AI inference, however, has different priorities, as low latency is essential to successfully deploying inference workloads.



Gruve raises $50 million to help close the gap in power as AI shifts from training to inference

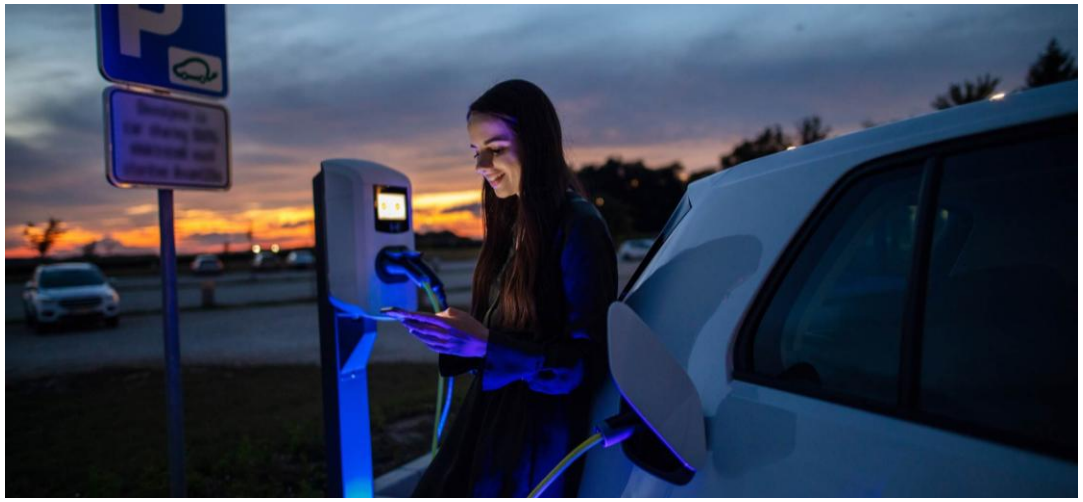
Gruve, an AI infrastructure and services startup, is taking aim at the power-gap problem head on, recently raising $50 million to combat the power constraints currently limiting AI inference workloads. The startup partners with colocation operators, such as Lineage and Open Colo, to create what they are calling an **Inference Infrastructure Fabric.** This is achieved by combining already-deployed power capacity across major metro markets with software to route AI inference requests to the "best" available site. This solution considers latency, cost, and power availability to route each inference request to the optimal location.

Sources: McKinsey & Company: The next big shift in AI workloads and hyperscaler strategies
Business Insider: Gruve raises $50 million to solve what its CEO calls AI's biggest problem: power

Edge data centers

As AI applications shift from the cloud to real-time environments, edge computing is becoming a major growth driver for data centers. The rollout of 5G standalone networks is fueling demand for regional, smaller, and mid-scale facilities located closer to end users, enabling ultra-low latency for workloads such as autonomous vehicles, industrial automation, and the expanding Internet of Things (IoT).

Telecom providers and digital infrastructure firms are expanding partnerships to deploy distributed edge nodes at network aggregation points. These facilities support AI inference and regional data replication, easing bandwidth strain on hyperscale hubs and improving response times. The next phase of growth will center on AI inference at the edge, where GPUs and specialized accelerators are embedded within telecom infrastructure to process data in milliseconds, bringing compute power closer to the user than ever before.



American Tower – edge expansion and data center growth

In Q3 2025, American Tower reported 7.7% revenue growth, driven by robust 5G network upgrades and surging demand in its CoreSite data center portfolio, where property revenue rose 14% YoY. The company highlighted record leasing activity from AI-related workloads, including inference, machine learning, and GPU-as-a-service deployments.

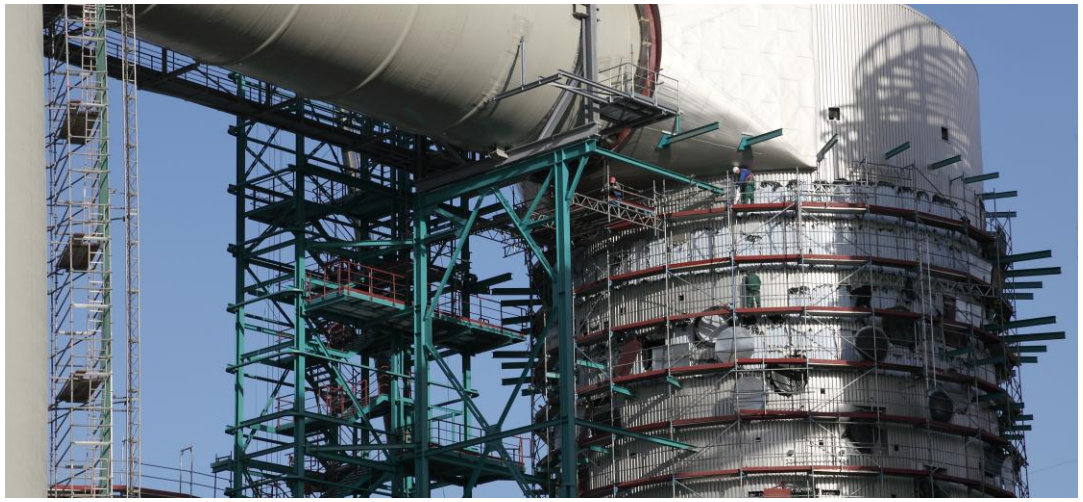
Beyond CoreSite's 30 U.S. facilities, American Tower is expanding its Access Edge program with small data centers co-located at tower sites in cities such as Raleigh, Austin, and Atlanta, positioning the firm to capture low-latency AI and hybrid-cloud demand at the network edge.

Source: Data Center Dynamics: American Tower reports revenue growth of 7.7%, fueled by telco upgrades and data center growth

On-site generation

As data center development accelerates, delays in power generation and transmission from local utilities have become a major constraint, limiting scalability and speed to delivery. In response, developers are increasingly pursuing self-generation and behind-the-meter solutions, such as on-site microgrids to maintain control over power timelines and mitigate grid dependency. Though capital-intensive, these systems allow operators to scale power on their own schedules rather than waiting years for utility upgrades.

Many operators are integrating renewable sources, including solar, battery storage, and fuel cells, to align with ESG goals while improving reliability. At the same time, natural gas generation is emerging as a practical near-term bridge solution. In power-constrained markets like Texas, Ohio, Pennsylvania, and Alberta, abundant gas resources and existing pipeline networks make on-site generation a cost-effective and scalable alternative to traditional grid supply.



Brookfield and Bloom Energy – $5-billion AI infrastructure partnership

In October 2025, Brookfield and Bloom Energy announced a $5-billion strategic partnership to develop next-generation AI factories: data centers designed to integrate compute, power, and infrastructure from the ground up. The agreement makes Bloom Energy the preferred on-site power provider for Brookfield's global AI developments, marking the first investment under Brookfield's new AI Infrastructure strategy.

The partnership will deploy Bloom's advanced fuel cell systems to deliver scalable, low-carbon on-site power that can operate independently from the grid. Together, the companies aim to accelerate AI infrastructure delivery in power-constrained markets, with the first site expected to launch in Europe by the end of 2025.

Source: Bloom Energy Press Release

Evolving design standards

AI and HPC workloads are fundamentally reshaping data center design, driving power densities from traditional 5-10 kW racks to deployments exceeding 50 kW, and in AI clusters, often surpassing 100 kW. To manage these loads, operators are rapidly implementing advanced liquid-cooling systems such as direct-to-chip and rear-door heat exchangers, paired with higher-voltage power distribution for greater efficiency. Many new facilities now integrate both air- and liquid-cooling capabilities within the same floor spaces, intentionally designing overlapping capacities that exceed total load requirements to maintain flexibility for varying workloads. Designs increasingly feature reinforced slabs, taller clear heights, and modular layouts to support phased expansion, heavy GPU enclosures, and dense thermal loads — most of which now rely on overhead cooling and power delivery rather than raised floors. As a result, modern data centers operate less like traditional IT environments and more like industrial-scale energy systems, purpose-built for resilience, scalability, and sustained high-performance computing.

NVIDIA GB300 NVL72 – a new benchmark for rack design

NVIDIA's GB300 NVL72 racks represent the latest evolution in high-performance compute infrastructure and are now being delivered to major data center operators for large-scale AI training and inference workloads. Each cabinet draws roughly 130–140kW and relies entirely on liquid cooling to manage heat from its 72 Blackwell B300 GPUs and 36 Grace CPUs.

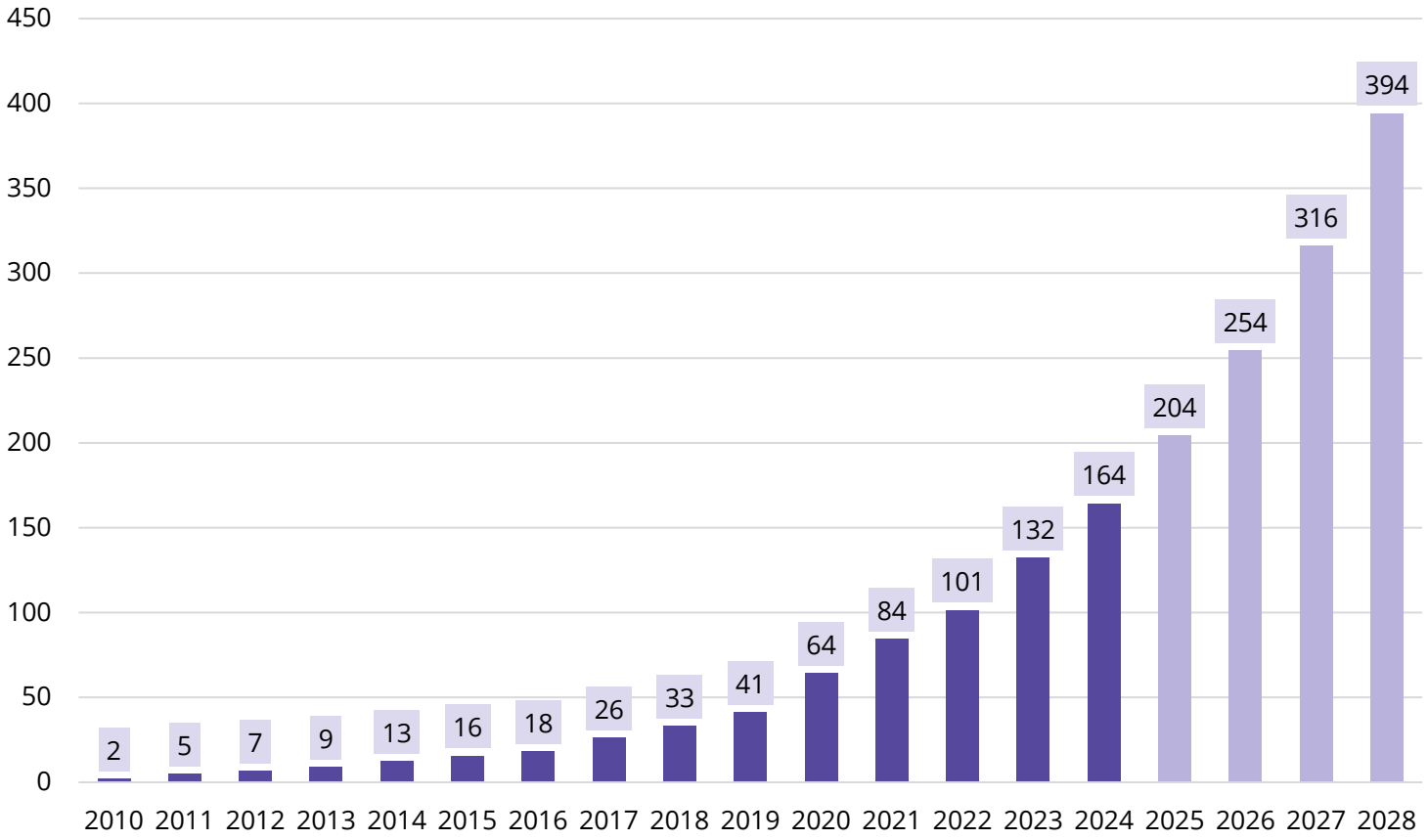
Each B300 GPU operates at 1,400W, and each Grace CPU operates at 500W. To put that into perspective, modern TVs typically operate below 100W and the average American household consumes around 800-1,000W.

Data center operators are now retrofitting facilities to accommodate these systems with higher-voltage power distribution, reinforced slabs, and dedicated liquid cooling loops.

Sources: NVIDIA Corporation: NVIDIA GB300 NVL72 – Designed for AI reasoning performance & efficiency
SemiAnalysis Newsletter: Nvidia's Christmas present: GB300 & B300 – Reasoning inference, Amazon, memory, supply chain

Demand drivers – digitalization

Data created and replicated worldwide (zettabytes)



The global datasphere (data created/replicated worldwide) is growing rapidly, projected to double from 2025 to 2028. Several new technologies have the potential to consume significantly more data than typical applications today, including:

Artificial intelligence (AI): AI and machine learning (ML) necessitate the collection, storage and computation of massive data sets

Robotics and autonomous vehicles (AVs): the training, inferencing, and retraining of robots and AVs create large quantities of data

Internet of things: previously "dumb" devices are now connected and generating data

Advancing technology: existing technologies require more data as sophistication increases (e.g., photo file size)

Cloud storage and computing: increasingly more enterprises are shifting data storage and computation to cloud and hybrid solutions

5G connectivity: enhanced mobile connection speeds have enabled new use cases that rely on quick communication with edge data centers (necessitating many more such edge data centers)

Source: IDC Global DataSphere Forecast, 2024-2028

Looking ahead

Large scale requirements

Leasing requirements are scaling to unprecedented levels, with AI users now urgently seeking blocks of 100 MW — and in some cases, exceeding 500 MW.

Transformer shortage

The ongoing shortage of transformers in the U.S. has considerably extended lead times for deliveries, affecting data center development and potentially causing delays of two-to-three years.

Pushing design limits

AI workloads are rapidly raising power density demands. Currently, leases are commonly signed at 120 kW per rack, and next-gen GPU racks are expected to require up to 600 kW — forcing major shifts in cooling strategies and facility design.

Outward push

Near-zero vacancy in primary urban markets is pushing development into outlying suburban and rural markets, where land and power are more readily available. As this shift accelerates, logistics will become a significant bottleneck.

Alternative power

As power limitations drive innovation, data center operators are increasingly turning to alternative off-grid and hybrid power solutions, including microgrid solar, wind, natural gas, and nuclear energy.

Pricing will keep rising

Low vacancy will persist as supply chain and utility constraints limit new deliveries, creating upward pressure on rents and sale prices for existing data centers and powered land.

New

OUTLOOK

2026

A new year, a fresh outlook.

As 2026 unfolds, there's a real sense of potential in the U.S. What does it mean for the data center sector?

Our top Market Intelligence experts break it down – sector by sector – so you can start the year with confidence.

Explore now

For more market insights and information visit **avisonyoung.com**

Howard Berry
Principal
National Director of Data Center Solutions
+1 408 375 6853
howard.berry@avisonyoung.com
CA Lic # 01795693

Stephen Schlenker
Associate
National Data Center Solutions
+1 505 977 2958
stephen.schlenker@avisonyoung.com
TX Lic # 695315

Chris Pedersen
Associate
National Data Center Solutions
+1 408 627 5686
chris.pedersen@avisonyoung.com

Zachary Cutler
US Data Center Lead, Market Intelligence
+1 610 557 0515
zachary.cutler@avisonyoung.com

Howard Huang
Associate
National Data Center Solutions
+1 510 500 7387
howard.huang@avisonyoung.com
CA Lic # 02248233

Alexandria Butler-Boitnott
Administrative Coordinator
National Data Center Solutions
+1 916 563 3901
alexandria.butlerboitnott@avisonyoung.com

AVISON YOUNG