

FROM 5 KW TO 350 KW:

How AI Rack Density Is Rewriting Data Center Design, Cooling, Insurance and Financing

Author: **Tom Harper**



Key takeaways:

- AI is accelerating the shift from traditional rack densities to high-density computing environments that require fundamentally different infrastructure strategies.
- Rising rack density is increasing demands on power, cooling, operational resilience and capital planning.
- Liquid cooling is emerging as a foundational component of many next-generation AI data centers.
- Growing value concentration is reshaping insurance, risk financing and lender considerations across the industry.
- Organizations that align infrastructure, risk management and capital strategy will be best positioned to support future AI-driven growth.

The data center industry is entering a new density era

Artificial intelligence is changing more than computing performance. It is reshaping the economics, engineering requirements and risk profile of modern data centers.

While many enterprise facilities still operate comfortably below 10 kW per rack, AI training and accelerated computing environments are rapidly moving into the 40 kW to 130 kW range. In some of the most advanced deployments, rack densities now exceed 200 kW and can reach as high as 350 kW. Future computing platforms are expected to push those requirements even higher, with some projections indicating that 250 kW could become a standard design assumption for next-generation AI infrastructure.

For data center owners, operators, developers, investors and lenders, rack density influences how facilities are designed, how they are cooled, how projects are financed and how risks are evaluated and insured.

As AI workloads continue to increase demand for compute capacity, the conversation has expanded beyond power and cooling. Organizations are increasingly evaluating value concentration, insurance market capacity, lender expectations and long-term operational resilience alongside traditional infrastructure requirements.

The growth of AI infrastructure is driving significant changes across the digital economy, creating new demands for power generation, transmission infrastructure and facility design.



AI data center rack density trends: Why the market is splitting in two

Average rack density can obscure where new capital is flowing. Much of the installed base remains anchored by enterprise and colocation workloads, but hyperscalers and large AI infrastructure providers are directing substantial investment toward high-density environments. The result is not one uniform market trend but two different planning models.

Industry research shows average server rack densities remain below 8 kW across much of the market, with most facilities operating below 30 kW per rack. More recent findings suggest that average rack power density is increasing gradually, primarily through broader adoption of deployments in the 10 kW to 30 kW range.^{1,2}

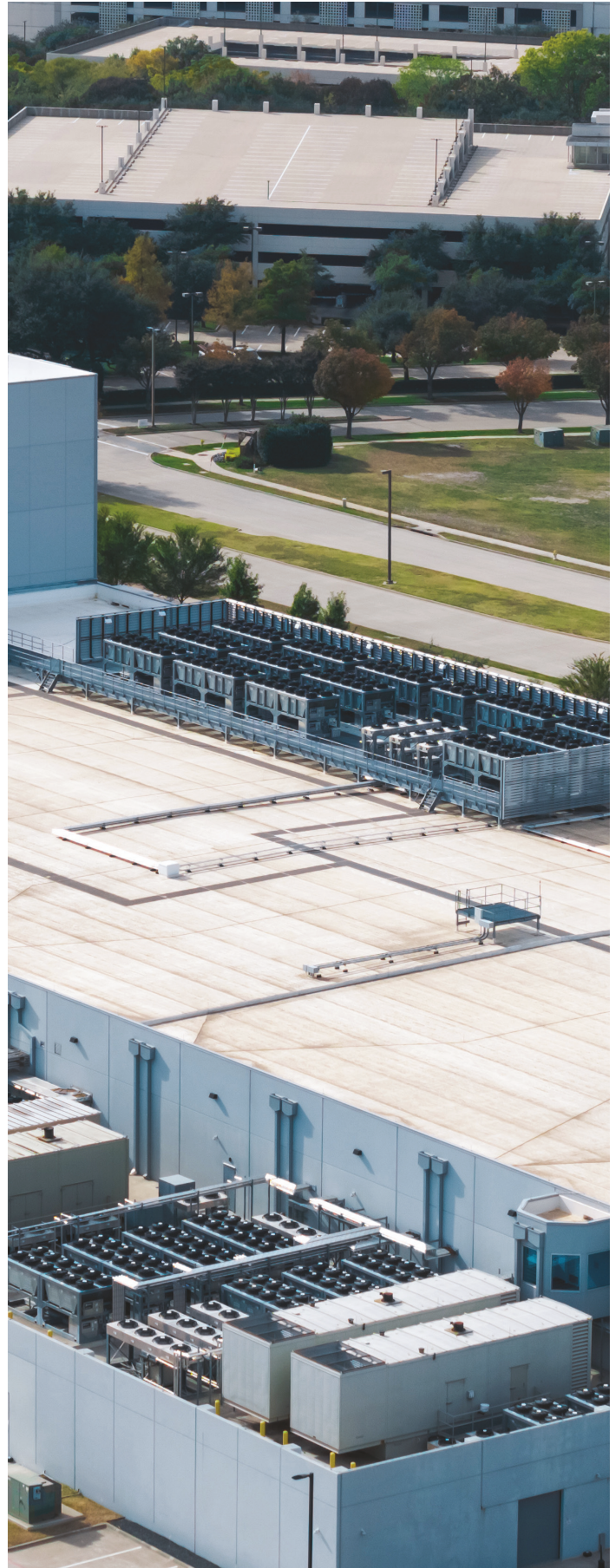
Traditional data centers

Traditional enterprise and colocation facilities generally continue to optimize around moderate density growth, infrastructure efficiency and a mix of customer workloads. Air cooling remains practical for much of this installed base, and upgrades may focus on selective higher-density zones rather than a facilitywide redesign.

AI factories

AI factories are purpose-built around accelerated computing, dense clusters of graphics processing units and large-scale model training or inference. They require more power per rack, more advanced electrical distribution and a thermal strategy designed for much higher heat loads. In these environments, decisions about building design, utilities, cooling, financing and insurance capacity are tightly connected.

Treating both facility types as though they follow the same density trajectory can lead to misaligned investment decisions and inaccurate risk assumptions. Planning should reflect the workload mix, the pace of technology replacement and the amount of capital concentrated in each location.



How AI workloads are driving high-density rack deployments

While average rack density is increasing gradually, the upper end of the market is advancing much faster.

Current industry research places many AI deployments in the 40 kW to 130 kW range, with future chip generations projected to push toward 250 kW per rack. At the same time, deployments operating above 100 kW per rack are becoming increasingly common across AI and high-performance computing environments.³

Perhaps more importantly, the range of rack densities across the market has never been wider. While many enterprise and colocation facilities continue to operate comfortably below 30 kW per rack, leading AI environments are already operating at several times that level. Specialized AI training and high-performance computing (HPC) facilities are running racks in the 200 kW to 350 kW range using advanced liquid-cooling technologies and customized power distribution systems.

This spread highlights an important reality: There is no longer a single definition of a modern data center. A facility designed around traditional enterprise workloads may have vastly different power, cooling, operational and risk-management requirements than a facility designed to support large-scale AI training models.

As density increases, the infrastructure required to support each rack becomes significantly more complex. Higher-density environments require greater power availability, more sophisticated electrical distribution systems, advanced thermal management strategies and increasingly coordinated approaches to operational resilience.

What was once a discussion about server density has evolved into a broader conversation about building design, utility infrastructure, cooling strategy, financing structures and insurance capacity.

AI data center rack density benchmarks

Category	Typical range
Legacy/enterprise average	5 kW-10 kW
Broader market trend	10 kW-30 kW
Standard AI deployment	30 kW-50 kW
Next-gen AI/HPC	40 kW-130 kW
Liquid cooling standard AI racks	100 kW-140 kW
Next-gen liquid systems	140 kW-200 kW
Ultra-high density (cutting edge)	200 kW-350 kW
Projected future ceiling	250 kW



Liquid cooling and thermal management: The new foundation of high-density AI data centers

Where air cooling still fits

Air cooling is not disappearing. It remains practical for many enterprise and colocation facilities and will continue to remove residual heat from components such as power supply units, power distribution units, switchgear and memory. However, extending air cooling to support hotter chips can require trade-offs, including lower rack utilization, higher airflow or lower supply-air temperatures. Those choices can reduce efficiency and may not align with AI architectures designed around densely packed accelerators.

As a result, many operators are adopting transitional strategies that combine both air- and liquid-cooling technologies within existing facilities. These hybrid approaches can help organizations support higher-density workloads while extending the useful life of existing infrastructure and managing capital expenditures.

Why liquid cooling changes facility design

At the highest density levels, liquid cooling is becoming less of an optimization and more of a design requirement. Rising GPU heat loads and increasing efficiency expectations are accelerating adoption of direct-to-chip cooling, rear-door heat exchangers and, in some cases, immersion cooling technologies. As rack densities continue to climb, thermal management is becoming a defining factor in facility design and operational resilience.

Importantly, liquid cooling introduces new operational, maintenance and risk management considerations. Cooling systems increasingly require supporting infrastructure that may include coolant distribution units (CDUs), manifolds, piping networks, water treatment processes, monitoring systems and specialized maintenance procedures. For many organizations, implementing liquid cooling means designing an entirely new thermal ecosystem around high-density computing environments.

Failure scenarios require early planning

As operators evaluate high-density deployments, several strategic considerations become increasingly important. Organizations must assess when air cooling may cease to be economically viable, whether existing electrical and structural infrastructure can support liquid-cooling retrofits and how cooling choices could influence future expansion plans.

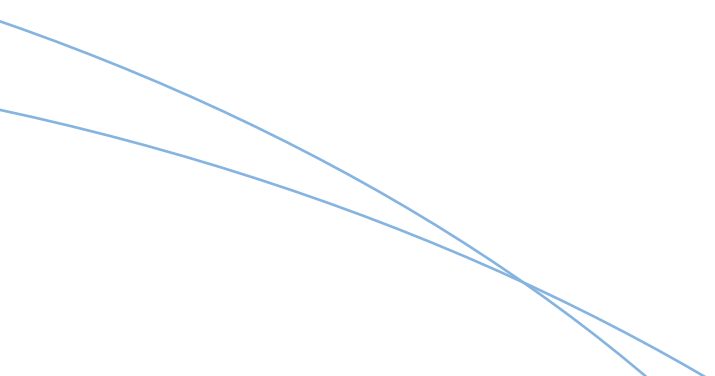
Water management is becoming equally important. Questions surrounding water availability, water quality, leak detection, corrosion control, and long-term sustainability increasingly influence site selection, facility design, and operational planning.

Perhaps most importantly, organizations need to understand how cooling system failures could affect business continuity. In high-density environments, a thermal management failure can quickly become more than a maintenance event. It can trigger operational disruptions, uptime concerns, customer service impacts, contractual obligations and potentially significant business interruption losses.

Josh Claman, Executive Chairman at Accelsius, states, "Insurers are starting to ask the right question, which is not whether liquid cooling works but what a failure costs." He adds,

"With water-based systems, the answer includes the hardware itself. With a two-phase dielectric refrigerant, a leak doesn't conduct, doesn't corrode and doesn't take out the GPUs. That changes both the probable maximum loss and the business interruption exposure."

These decisions increasingly intersect with broader considerations around power availability, site selection, utility infrastructure and long-term resilience planning. As AI infrastructure evolves, the cooling strategy is becoming one of the most important links between facility operations, risk management and capital planning.

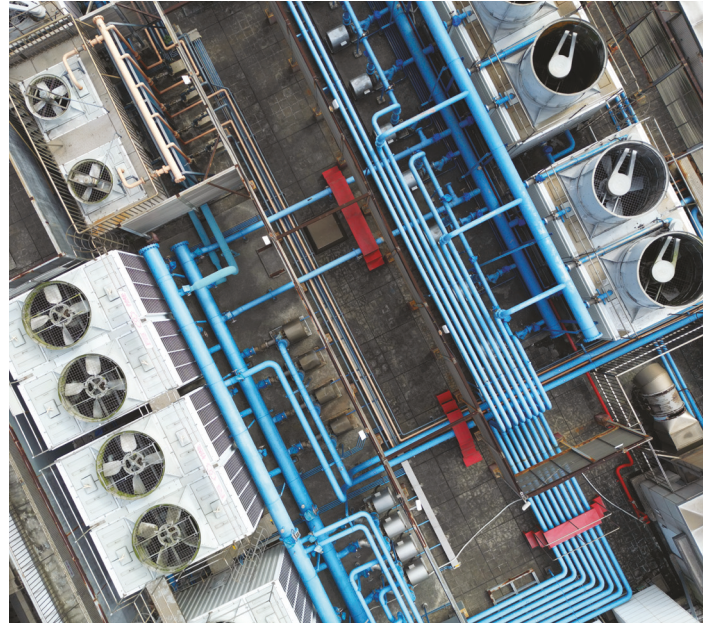


Value concentration is reshaping insurance decisions

As rack density increases, more capital and technology can be concentrated within a smaller footprint. Construction costs for an individual AI-focused campus can reach tens of billions of dollars before servers, networking infrastructure and tenant-owned equipment are included. A site that may once have been viewed primarily as real estate can instead represent highly concentrated, mission-critical infrastructure.

This shift builds on broader concerns surrounding data center construction risk, including the interaction among project values, technology assets, physical hazards and business interruption exposure.

Some campuses can challenge the amount of property and construction insurance capacity available for a single location. Owners and investors therefore need to determine how much value can realistically be transferred to the insurance market and how much risk remains with owners, lenders and investors.



Financing and lender expectations are evolving



When replacement values exceed available insurance capacity, lenders may place greater emphasis on catastrophe modeling, business continuity, loss-scenario analysis and risk improvement plans. Insurance becomes one component of a broader risk-financing framework that can include capital reserves, contractual risk transfer, alternative financing structures and operational investments.

Probable maximum loss (PML) and maximum foreseeable loss (MFL) frameworks can help organizations move beyond a full replacement-value assumption and evaluate credible loss scenarios. These discussions may influence financing structures, lender requirements and project economics well before a facility becomes operational.

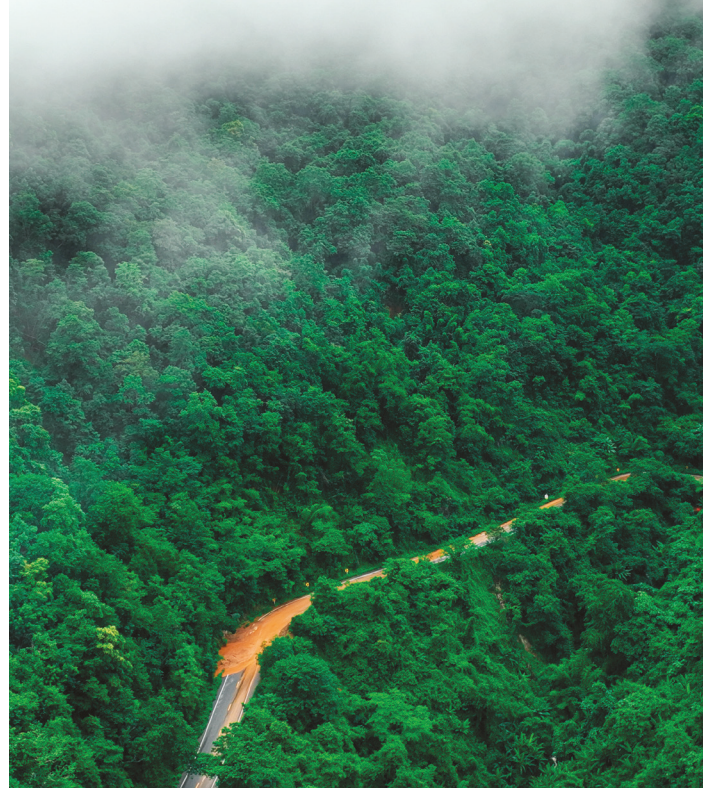
Financing markets are also adapting to the scale of AI campuses. Larger projects may require a broader mix of capital sources and more tailored arrangements. Infrastructure design, insurability and financing should therefore be evaluated together rather than in sequence.

Operational risk extends beyond physical damage

Natural hazards such as hail, tornado, flood, wildfire and convective storms can affect sites containing significant physical assets and concentrated computing capacity. Yet the risk extends beyond property damage. Cooling system failures, utility interruptions, cyber events, delayed start-ups, technology failures, supply chain disruption and business interruption can all affect high-density AI environments.

Liquid cooling adds questions about equipment ownership, contamination, leakage, maintenance responsibilities and downtime. In multitenant environments, contracts should clearly allocate responsibility for tenant-owned computing equipment, cooling infrastructure and associated interruption losses. Operational procedures should also define monitoring, escalation, isolation and recovery responsibilities.

The central issue is dependency: High-density environments place more capital, technology and operational reliance in a smaller space. Organizations are better positioned when facility engineering, insurance, risk financing and continuity planning use the same loss scenarios and design assumptions.



The future of AI data centers: Density, power demand and growth through 2030

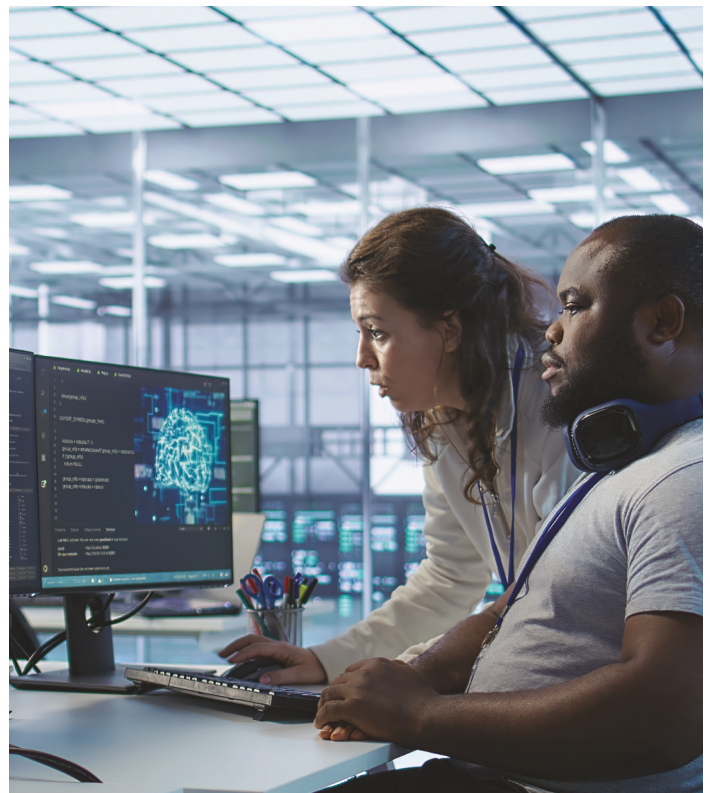
Over the next several years, both rack density and power demand are expected to continue moving higher.

Near-term outlook (2026-2027)

The broader market is likely to continue shifting from traditional 5 kW to 10 kW assumptions toward 10 kW to 30 kW deployments, while AI-focused facilities increasingly require 40 kW to 130 kW capability as a baseline planning requirement.

Industry forecasts suggest the global data center market could expand at approximately 15% annually through 2027, with some projections indicating upside potential approaching 20%, driven by continued growth in AI, cloud computing and other compute-intensive workloads.³

At the same time, demand for capacity remains exceptionally strong. Primary-market vacancy rates fell to near-historic lows despite substantial increases in available supply.⁴ This combination of rapid demand growth and constrained capacity is putting additional pressure on power infrastructure, development timelines and capital planning decisions.





Medium-term outlook (2028-2030)

The longer-term outlook suggests continued acceleration across both capacity demand and infrastructure investment.

Current forecasts estimate that approximately \$6.7 trillion of global investment may be required by 2030 to support projected data center demand, with AI workloads representing an increasingly significant share of future capacity requirements.⁵

Power availability is expected to become one of the defining constraints on industry growth. Some research projects suggest that data center power demand could increase by 165% to 175% by 2030 compared with 2023 levels, driven largely by AI workloads and the continued expansion of cloud infrastructure. Additional forecasts suggest US data center grid demand could nearly triple by the end of the decade, placing greater pressure on utilities, transmission infrastructure and long-term energy planning.⁶

Density and power are increasingly two sides of the same equation.

Higher-density deployments can reduce the physical footprint required per unit of compute capacity, but they simultaneously increase demands on electrical infrastructure, cooling systems, redundancy planning and operational resilience. As a result, infrastructure strategy, financing decisions and risk management frameworks must increasingly account for both compute growth and the energy infrastructure required to support it.

What data center owners, investors and operators should consider next

Average rack density is increasing steadily across much of the market, but maximum rack density is rising far more rapidly. That gap is emerging as one of the defining planning challenges for organizations that design, build, finance, own and insure data center infrastructure.

The industry is separating into two distinct operating environments: traditional enterprise facilities and AI-driven, high-density infrastructure. Each requires different assumptions around design, operations, capital allocation and risk management.

Cooling strategy is moving from an operational consideration to a foundational business decision. As density increases, thermal management directly influences reliability, expansion potential and long-term operating costs. Liquid cooling is increasingly becoming a standard component of AI infrastructure rather than a niche HPC solution.

At the same time, value concentration is reshaping how organizations think about insurance, financing and resilience. As facility values continue to grow, the ability to secure appropriate insurance capacity and financing support may become just as important as access to land, power and fiber connectivity.

The next generation of AI data centers will not be defined solely by the number of megawatts they can deliver. They will also be defined by how effectively owners and operators manage heat, power, value concentration, operational resilience and insurability from the earliest stages of development.

Ensuring your organization is prepared

Want to explore how rising AI data center rack density could affect your organization's infrastructure strategy, risk profile or future development plans?

Contact Gallagher's Data Center practice to start the conversation.

Sources:

1. Uptime Institute. "2024 Global Data Center Survey."
2. Uptime Institute. "2025 update on rack power density trends."
3. JLL. "2025 Global Data Center Outlook."
4. CBRE. "North America Data Center Trends report."
5. McKinsey & Company. "Data center investment outlook."
6. Goldman Sachs Research. "Data center power demand forecast."

AJG.com The Gallagher Way. Since 1927.



The information contained herein is offered as insurance industry guidance and provided as an overview of current market risks and available coverages and is intended for discussion purposes only. This publication is not intended to offer financial, tax, legal or client-specific insurance or risk management advice. General insurance descriptions contained herein do not include complete Insurance policy definitions, terms, and/or conditions, and should not be relied on for coverage interpretation. Actual insurance policies must always be consulted for full coverage details and analysis.

Gallagher publications may contain links to non-Gallagher websites that are created and controlled by other organizations. We claim no responsibility for the content of any linked website, or any link contained therein. The inclusion of any link does not imply endorsement by Gallagher, as we have no responsibility for information referenced in material owned and controlled by other parties. Gallagher strongly encourages you to review any separate terms of use and privacy policies governing use of these third-party websites and resources.

Insurance brokerage and related services provided by Arthur J. Gallagher Risk Management Services, LLC License Nos. IL 100292093 / CA 0D69293

© 2026 Arthur J. Gallagher & Co., and affiliates & subsidiaries | PRODUS203419