

The Data Center Balancing Act: Powering Sustainable AI Growth

By Victoria Schlotterback, CPA, Equity Research Analyst; Joe Pasqualichio, Equity Research Analyst; and John Bond, CFA, Equity Research Analyst



Background: The Resource Tipping Point

The rise of generative artificial intelligence (GenAI)¹ is driving rapid growth in data center development and electricity consumption. As GenAI workloads increase, the International Energy Agency (IEA) predicts that global electricity demand fueled by AI growth will double by 2026, threatening grid capacity.² This raises critical questions for investors: How will hyperscale data centers overcome resource constraints? How will sustainability play a role? What investment opportunities arise from this power dynamic?

The United States (U.S.) hyperscalers are among the largest buyers of compute and installers of data center capacity.³ Hyperscalers provide cloud computing and data management services to organizations that require vast infrastructure for large-scale data processing and storage. Amazon, Microsoft, and Alphabet (Google), in aggregate, account for over 60% of global hyperscale data center capacity, with Meta closely following in ranking.⁴ These companies have a distinct competitive advantage in accessing limited resources, and we believe, are uniquely positioned to catalyze a sustainable transformation of our power systems, which is underappreciated by the broader market.

Our focus in this piece is on data centers, the buildings that house the compute, storage, and networking equipment that underpins our digital world. Data centers are key to realizing the promise of artificial intelligence, but their growth is pushing against limits on key inputs — power, water and land. We will examine data center market dynamics and the steps hyperscalers are taking to increase the likelihood that these limits do not gate the exciting opportunities at the forefront of GenAI. We will highlight investment implications across the data center supply chain, from energy-efficient chips to renewable energy to advanced cooling technologies, that could revolutionize the power balancing act and perhaps catalyze a sustainable grid transformation.

The 2024 electricity report from the IEA states that data centers

consumed 460TWh in 2022, representing 2% of all global electricity usage. This figure could rise to more than 1,000TWh by 2026 in a “worst-case scenario.”⁵ This increase is equivalent to adding the power consumption of a country such as Sweden at the lowest end of the scale or Germany at the highest.

In the U.S., home to 33% of the world’s data centers, consumption is expected to rise from 200TWh in 2022 to 260TWh in 2026, which would equate to 6% of all power use across the country.⁶ After more than 30 years of falling or flat demand for on-grid electricity, electric utilities forecast that the U.S. will need 38 gigawatts (GW) of new power production, or the equivalent of about 34 new nuclear plants, over the next five years to supply power for data centers, electrification, and new industry.⁷

Historically, rapid improvements in energy efficiency and off-grid supply, such as domestic solar, have helped moderate growth in energy consumption from data centers and data transmission networks globally despite global internet traffic expanding 25-fold since 2010.⁸ Specifically, there have been efficiency improvements in IT hardware and cooling, as well as a shift away from small, inefficient enterprise data centers towards more efficient colocation and hyperscale data centers.⁹ Yet, the forward-looking scenario presents new and unique challenges. AI-specialized data centers require more power than traditional data centers, as AI-specific compute consumes significantly more power per rack than traditional compute.

The increasing electricity consumption of data centers prompts a range of environmental impacts. In addition to the global focus on greenhouse gas emissions, data centers present challenges for water conservation, e-waste, and environmental concerns around natural resources, such as the mining of rare earth elements required to manufacture hardware components. As the growth of data centers pressures existing power grids, this leads to demand for new power plants and/or grid expansion, which impacts local communities, particularly land and water use.

Discovering The Data Center Market

Northern Virginia, home to the greatest concentration of data centers in the world, is experiencing constraints sooner than most.¹⁰ Dominion Energy, a company that supplies electricity in the region, stated in its 2023 Annual Report that data centers, representing 24% of Virginia Power's electricity sales for the year ended December 31, 2023, have been a source of significant increase in demand, which is expected to continue over the next decade.

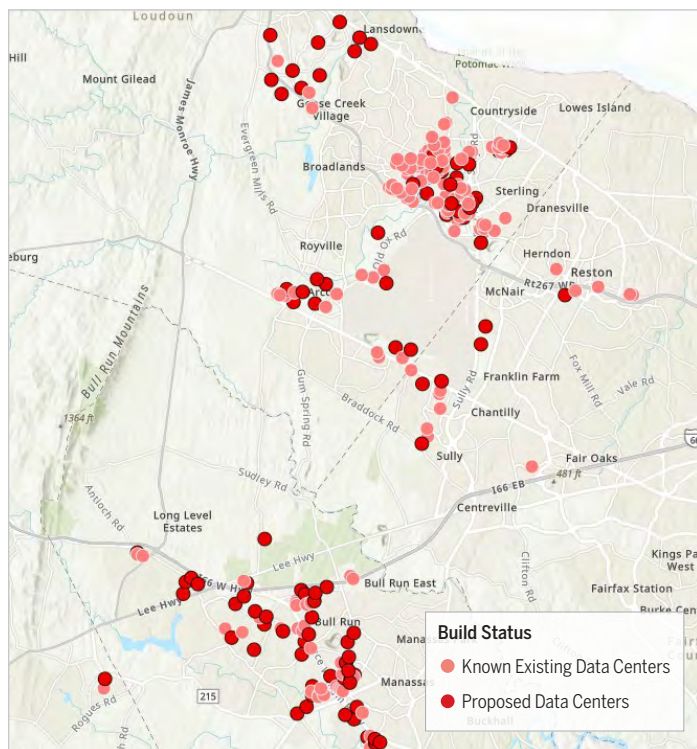
The concentration of data centers primarily in Loudoun County, Virginia, represents a complex challenge. Data center power constraints often arise due to transmission and distribution limitations rather than a lack of power generation capabilities. In other words, the problem isn't insufficient power generation but rather the inability to distribute it effectively. This is leading to significant investments in electric transmission facilities to meet the growing demand.¹¹ As we mentioned previously, local communities can be negatively impacted by this expansion. The potential use of eminent domain – the power of local and state governments to take private property for public use, in exchange for compensation – to complete transmission line projects is a looming concern for citizen groups.

Dominion Energy says it has received customer orders that could double the amount of data center capacity in Virginia by 2028, with a projected market size of 10GW by 2035.¹² We have heard from industry experts that electric transmission constraints are forcing some data centers to wait up to five years or more to secure grid connections. At the same time, counties across Northern Virginia are considering bills that may curb data center development in the region. This mirrors moratoriums around the world such as Ireland's energy grid halting the development of new data centers in Dublin until 2028.¹³ Data center development is now expanding beyond Northern Virginia due to the region's constraints. Power availability, reliability, connectivity, and local regulations are key factors in site selection. The industry is adapting as state regulators change zoning rules, demand environment and resource impact assessments, and impose guidelines on water usage.

As it relates to water usage, cooling presents a multifaceted challenge. While research continues on the topic and each data centers' water intensity varies, the most common range that we hear is the average data center consumes 3-5 million gallons of water per day. Most is local freshwater, putting pressure on drinking water supplies, especially in drought-prone areas. The World Resources Institute forecasts that about one-third of data centers globally are now located in areas with high or extremely high levels of water stress.¹⁴ Accordingly, leading data center operators are prioritizing water conservation and exploring alternative cooling methods, such as immersion cooling and direct-to-chip cooling. Remarkably, Microsoft's massive new data center in Wisconsin, for example, will be cooled with a closed-loop, water recycling system that uses new water only in rare circumstances.¹⁵ Governments are also looking for ways to incentivize efficient cooling. In 2022, the U.S. Department of Energy (DOE) announced up to \$42 billion in funding to support the development of energy-efficient, high-performance cooling solutions for data centers.¹⁶

Companies worldwide closely monitor and influence the quickly evolving regulatory landscape for energy-efficient data centers. Virginia's Clean Economy Act mandates that Dominion Energy and American Electric Power transition to 100% renewable energy sources by 2045 and 2050, respectively. Similarly, the European Energy Efficiency Directive promotes energy efficiency across sectors, and the first step for data centers is mandatory reporting of energy use and emissions from data centers which are larger than 500kW. While some jurisdictions prioritize approvals for data centers with renewable energy commitments, others grapple with spiking energy demand by keeping coal power plants operational.

FIGURE 1: EXISTING AND PROPOSED DATA CENTERS IN VIRGINIA



Source: Piedmont Environmental Council, as of 07/12/2024; <https://www.pecva.org/work/energy-work/data-centers/existing-and-proposed-data-centers-a-web-map/>



Victoria Schlotterback, Elisabeth Hiss, and Tiffany Ernest, Brown Advisory's Director of Infrastructure, visiting Quality Technology Services (QTS), a Blackstone portfolio company that delivers secure data center solutions with over two gigawatts of power capacity in North America and Europe. QTS owns four data centers in Loudoun County, Virginia, including Ashburn-Broderick, Ashburn-Lockridge, Ashburn-Moran, and Ashburn-Shellhorn.

FIGURE 2: WORLDWIDE DATA CENTER MARKET



Source: Data Center Map, as of 07/12/2024; <https://www.datacentermap.com/datacenters/>

Rising Hyperscaler Capex: A Race for Resources

Hyperscalers are racing to capture share in the AI market with a sizable wave of investment. As shown in Figure 3, capital expenditure (CAPEX) by Microsoft, Google and Meta increased to an aggregate of over \$100 billion in 2023 primarily driven by data center infrastructure.

Hyperscalers have worked to reduce costs and resource consumption through increasing efficiency, with a strong lever being the semiconductor industry. NVIDIA’s new Blackwell platform enables organizations to build and run real-time GenAI on trillion-parameter large language models at up to 25x less cost and energy consumption than its predecessor.¹⁷ However, while semiconductor efficiency will continue to play a critical role, as transistors approach the size of individual atoms, it’s becoming more difficult for the industry to benefit from Moore’s Law. Jevons paradox points to efficiency resulting in an increase of energy consumption in the long run. In essence, there is less juice to squeeze from the era of miniaturization to counterbalance the rate of growth in compute that semiconductor chips enable. This calls for a broader solution set, including software and system-level efficiency gains.

Hyperscalers: Catalyzing Sustainable Energy

Hyperscalers seemingly remain committed to their ambitious sustainability goals despite the challenges deterring progress as GenAI scales. Microsoft’s indirect emissions (Scope 3) increased by 30.9% from their 2020 baseline, which “primarily comes from the construction of more datacenters”¹⁸ and Google’s total GHG emissions increased 48% compared to their 2019 target base year “primarily due to increases in data center energy consumption.”¹⁹ Still, both companies persist in their bold set of environmental commitments. Microsoft aims to be a carbon negative, water positive, zero waste company that protects ecosystems — all by

2030. Microsoft’s 2024 Environmental Sustainability Report outlines the development of more than 80 discrete and significant measures that should help the company reduce these emissions — including a new requirement for select scale, high volume suppliers to use 100% carbon-free electricity for Microsoft delivered goods and services by 2030. Likewise, Google signed contracts to purchase approximately 4 gigawatts of clean energy generation capacity in locations such as Texas, Belgium, and Australia — more than in any prior year — and celebrated a first-of-a-kind enhanced geothermal project now delivering carbon-free energy to the grid.²⁰

We believe viewing hyperscalers purely as power-hungry giants undermines the bigger picture. Many of their investments to meet their energy needs may be laying the foundation for a more sustainable, resilient grid.

Amazon, Microsoft, Meta, and Google are the four largest purchasers of corporate renewable energy power purchase agreements (PPAs), having contracted almost 50GW to date, equal to the generation capacity of Sweden.²¹ Microsoft announced a monumental \$10 billion renewable energy deal with Brookfield Asset Management. Commencing in 2026, Brookfield will spearhead the deployment of more than 10.5 gigawatts of renewable energy capacity, equivalent to the output of 10 nuclear power plants.²²

Nuclear power is gaining renewed attention as a reliable, low-carbon source of always-on electricity that could help balance intermittent renewables, although hurdles exist around cost, timing, and safety concerns. Data center operators are exploring the potential to site facilities near existing nuclear plants. Amazon recently purchased a nuclear-powered data center campus in Pennsylvania for \$650 million to supplement its renewables.²³ Microsoft entered into the world’s first fusion energy purchase agreement with energy developer Helion Energy for the provision

of electricity from its first fusion power plant expected to be online by 2028.²⁴ We're also seeing companies such as Microsoft showing early signs of investing in small modular reactors (SMRs).²⁵ As nuclear technology continues to advance, utility regulators could become more comfortable approving new nuclear capacity. In December 2023, Microsoft published a [policy brief](#) that highlights the importance of carbon-free electricity and the role that advanced nuclear and fusion energy may play in a decarbonized energy future.²⁶ Listen to the Brown Advisory NOW Podcast on this subject:



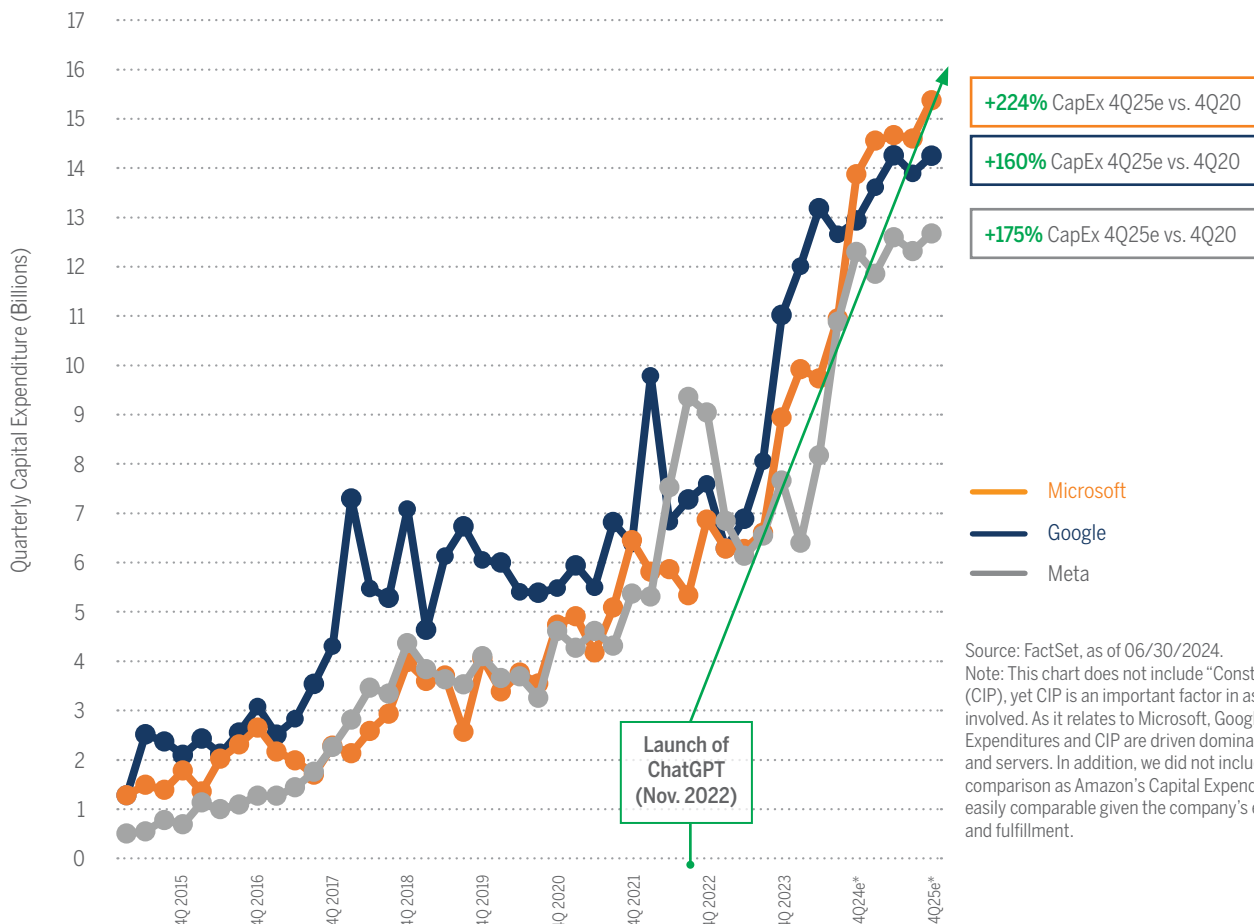
As renewables continue to develop, grid interoperability becomes increasingly important. Companies utilizing heat batteries and other storage methods are already making it easier and cheaper to store excess renewable generation. The key to unlock is for data centers to dynamically interact with power systems, potentially transforming them into giant batteries that absorb excess solar and wind energy. However, achieving this potential depends on several factors, including scalability, reliability, flexible policies, and emerging market structures.

What is often overlooked is how AI itself is being harnessed to solve the energy and power grid constraints. Prominent energy companies worldwide have swiftly embraced AI, recognizing its potential to optimize planning, facility management, environmental impact, energy storage, and distribution. The technology sector is also prioritizing the use of digital technology, including AI, to gather insights and automate the management of electricity and water consumption. This is not a new phenomenon in the technology sector. In 2016, Google announced that by applying DeepMind's machine learning to their own Google data centers, the company managed to reduce the amount of energy used for cooling by up to 40%.²⁷

The industry is also piloting early examples of supporting grids such as Google's demand response method that allows them to reduce power demand of a data center during times of stress in grids by shifting now-urgent computing tasks to others times and locations. [The U.S. Department of Energy published research](#) suggesting demand response can be useful tool to reduce the need for new power investment on the grid. This minimizes environment at risk in grid expansion as well as detrimental community impact.

All of this points to a future where data centers are not just power consumers, but active partners in decarbonizing the grid. Hyperscalers, with their substantial energy demand, have both the incentive and the influence to drive this transformation. Erika Pagel, Brown Advisory Portfolio Manager, Co-CIO of Private Client, Endowments & Foundations, aptly describes this as an "unlikely relationship, or marriage" of technology and energy.²⁸

FIGURE 3: GROWTH IN QUARTERLY CAPITAL EXPENDITURES – DRIVEN BY DATA CENTERS



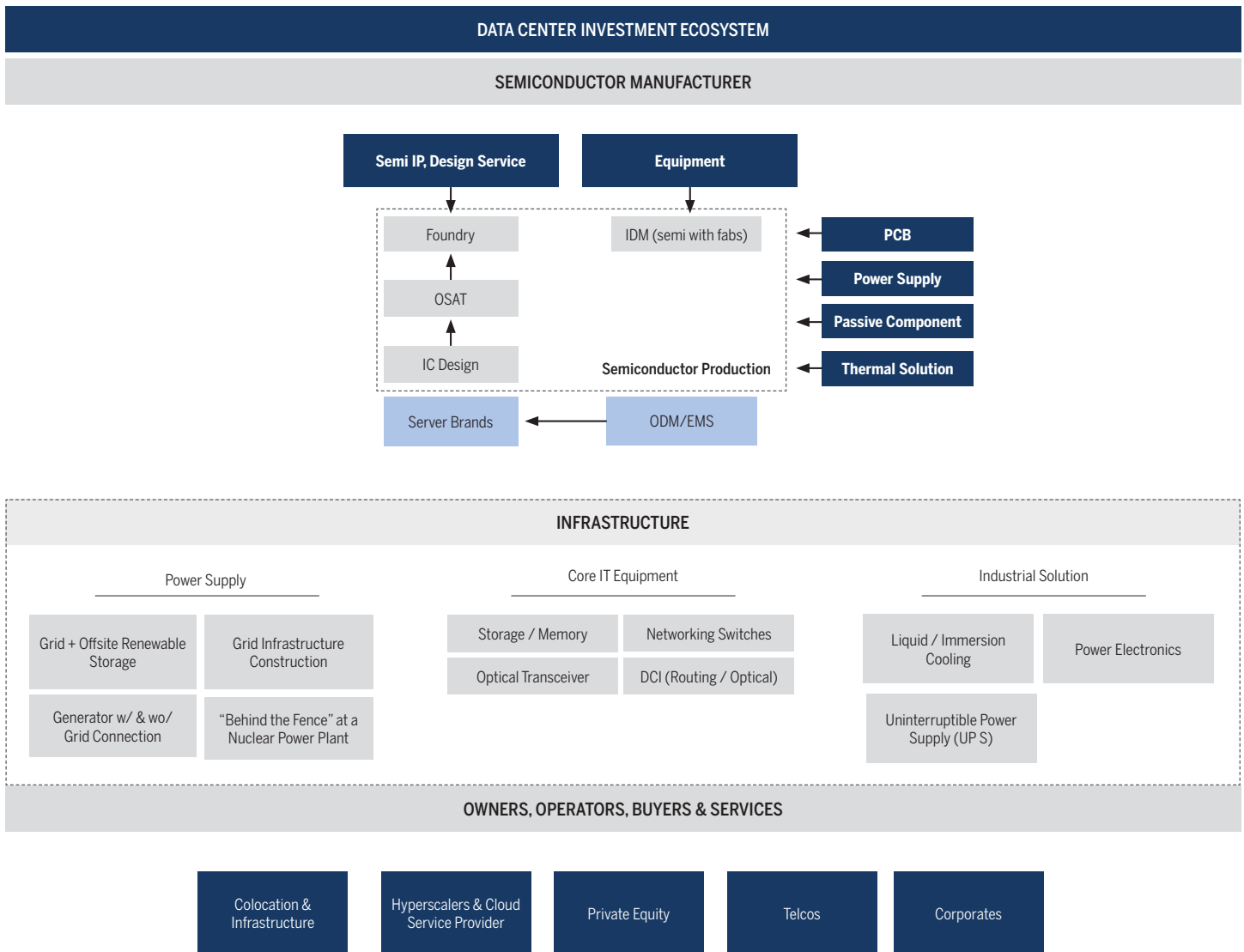
Source: FactSet, as of 06/30/2024.
 Note: This chart does not include "Construction in Progress" (CIP), yet CIP is an important factor in assessing total capital involved. As it relates to Microsoft, Google, and Meta, Capital Expenditures and CIP are driven dominantly by data centers and servers. In addition, we did not include Amazon in this comparison as Amazon's Capital Expenditures are not as easily comparable given the company's exposure to logistics and fulfillment.

Investment Implications

GenAI sparks an inflection point to stress-test our power infrastructure while unleashing innovation to make data centers more efficient. We believe that hyperscalers have the scale, capital, and engineering prowess to balance compute growth with resource sustainability in the long-run. This will require a multi pronged approach including substantial investments in renewables, storage, and grid balancing capabilities; focus on water conservation and alternative cooling technologies; deployment of onsite generation and dedicated renewable energy where feasible; policy advocacy to modernize power grids and market structures; and innovation in AI infrastructure to be more efficient. As investors, we monitor companies across the data center supply chain, including operators, semiconductors, power electronics, cooling, and software providers, to name a few. Semiconductors are the early beneficiaries in a compute cycle, with the other components of data center infrastructure to follow. Time will tell the extent to which data centers catalyze scaling grid storage and flexibility, creating markets for new technologies and software.

An interesting software investment example is Cadence Design Systems', "Cadence Reality DC," the first campus-wide digital twin simulator for data centers.²⁹ Customers are modeling liquid cooling within Cadence Reality DC to facilitate more efficient cooling for high-density AI equipment.³⁰ The most innovative companies in the supply chain will enable sustainable data center development. One potential solution is Modular Data Centers (MDCs). An MDC is a compact and portable collection of all the key components of a traditional data center, including servers, networking equipment, cooling, and storage.³¹ Dell, for example, has a portfolio of pre-fabricated and assembled MDCs which significantly reduces time to deploy and resources needed onsite.³²

FIGURE 4: MAPPING OUT THE AI INFRASTRUCTURE VALUE CHAIN



Source: Morgan Stanley Research

Looking at the broader supply chain, hyperscalers' investments in grid solutions could also yield benefits across utilities, storage providers and more. In balancing near-term challenges with long-term upside, we find attractive investment opportunities in technology enablers that are helping data centers overcome constraints and become more grid flexible. From power management, cooling and energy storage, the opportunity set is large and diversifying. Technology evolves quickly, and the tale of GenAI is unfolding at an unprecedented pace. The vast majority of GenAI use cases and implementations are still in the early stages. Yet, what is quite clear is that as GenAI applications and inferencing scales, power demand

is set to increase commensurately. There are endless unknowns of what the future holds, and we do not hold a crystal ball. This is what makes the role of an investment analyst thrilling. Data centers demonstrate the critical intersection of environmental sustainability being both a business imperative and a competitive advantage. If hyperscalers catalyze the transition to a more robust, resilient and sustainable power grid while continuing to achieve step-changes in compute efficiency, we believe immense value will be unlocked.

1. Generative artificial intelligence (GenAI) is a type of AI that can create new content and ideas, including conversations, stories, images, videos, and music. Source: AWS, as of 07/12/2024: <https://aws.amazon.com/what-is/generative-ai/>
2. Source: IEA, as of 07/12/2024; <https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf>
3. Source: ENCONNEX, as of 02/22/2024; <https://blog.enconnex.com/where-are-most-data-centers-located-top-markets>
4. Source: Synergy, as of 07/12/2024; <https://www.srgresearch.com/articles/hyperscale-data-centers-hit-the-thousand-mark-total-capacity-is-doubling-every-four-years>
5. Source: IEA, as of 07/12/2024; <https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf>
6. Source: IEA, as of 07/12/2024; <https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf>
7. According to filings made to the Federal Energy Regulatory Commission and compiled by Grid Strategies. <https://gridstrategiesllc.com/wp-content/uploads/2023/12/National-Load-Growth-Report-2023.pdf>
8. Source: IEA, as of 07/11/2023; <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>
9. Source: IEA, as of 07/12/2024; <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>
10. Source: Northern Virginia Technology Council 2024; <https://www.nvtc.org/communities/data-center-and-cloud/report/>
11. Source: Dominion Energy Annual Report 2023; https://s2.q4cdn.com/510812146/files/doc_downloads/2024/2024/03/20/20/Dominion-Energy-2023-Annual-Report-and-Annual-Report-on-Form-10-K.pdf
12. Source: Data Center Frontier, as of 10/11/2023; <https://www.datacenterfrontier.com/energy/article/33013010/dominion-virginias-data-center-cluster-could-double-in-size>
13. Source: RTE, as of 01/10/2023; <https://www.rte.ie/news/dublin/2022/0110/1272869-eirgrid-datacentres-dublin/>
14. Source: WRI, as of 05/15/ 2015; <https://www.wri.org/data/aqueduct-water-stress-projections-data>
15. Source: Microsoft, as of 05/08/2024 <https://news.microsoft.com/2024/05/08/microsoft-announces-3-3-billion-investment-in-wisconsin-to-spur-artificial-intelligence-innovation-and-economic-growth/#:~:text=The%20new%20datacenter%20will%20use,sustainability%20and%20responsible%20resource%20management>
16. Source: ENERGY.GOV, as of 09/22/2022; <https://www.energy.gov/articles/doe-announces-42-million-develop-high-performance-cooling-systems-data-centers>
17. Source: NVIDIA Sustainability Report 2024. <https://images.nvidia.com/aem-dam/Solutions/documents/FY2024-NVIDIA-Corporate-Sustainability-Report.pdf>
18. Source: Microsoft 2024 Environmental Sustainability Report, reporting on 2023 fiscal year; <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1IMIE>
19. Source: Google Environmental Report 2024; <https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf>
20. Source: Google, as of 07/02/2024; <https://blog.google/outreach-initiatives/sustainability/2024-environmental-report/>
21. Source: IEA, as of 07/12/2024; <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>
22. Source: Energy Digital, as of 05/03/2024; <https://energydigital.com/articles/microsoft-brookfield-sign-worlds-biggest-clean-power-deal>
23. Source: Environment + Energy LEADER, as of 03/11/2024; <https://www.environmentenergyleader.com/2024/03/aws-650m-sustainable-storage-hub-in-pennsylvania/>
24. Source: Helion, as of 05/10/2024; <https://www.helionenergy.com/articles/announcing-helion-fusion-ppa-with-microsoft-constellation/>
25. Small modular reactors (SMRs) are advanced nuclear reactors that have a power capacity of up to 300 MW(e) per unit, which is about one-third of the generating capacity of traditional nuclear power reactors. Source: IAEA, as of 09/13/2023; <https://www.iaea.org/newscenter/news/what-are-small-modular-reactors-smrs>
26. Source: Microsoft, as of 12/2024; <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1fApf>
27. Source: Google DeepMind, as of 07/20/2024; <https://deepmind.google/discover/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40/>
28. Source: Brown Advisory, as of 06/06/2024; <https://www.brownadvisory.com/us/insights/cio-perspectives-podcast-powering-ai-boom-investing-aerospace-and-surge-drug-innovation>
29. A digital twin is a virtual representation of an object or system designed to reflect a physical object accurately. It spans the object's lifecycle, is updated from real-time data and uses simulation, machine learning and reasoning to help make decisions. Source: IBM, as of 07/16/2024; <https://www.ibm.com/topics/what-is-a-digital-twin>
30. Source: Cadence, as of 03/18/2024; https://www.cadence.com/en_US/home/company/newsroom/press-releases/pr/2024/revolutionary-cadence-reality-digital-twin-platform-to-transform.html
31. Source: Data Span, as of 08/24/2023; <https://dataspan.com/blog/what-is-a-modular-data-center/>
32. Source: DELL Technologies, Modular Data Center Brochure; https://www.delltechnologies.com/asset/en-us/solutions/infrastructure-solutions/briefs-summaries/modular-data-center-brochure.pdf?ref=cpcb_thisishowyoua-kehtedge1_cta_primaryeditoriallight_readmore

The views expressed are those of the author and Brown Advisory as of the date referenced and are subject to change at any time based on market or other conditions. These views are not intended to be and should not be relied upon as investment advice and are not intended to be a forecast of future events or a guarantee of future results. Past performance is not a guarantee of future performance and you may not get back the amount invested.

The information provided in this material is not intended to be and should not be considered to be a recommendation or suggestion to engage in or refrain from a particular course of action or to make or hold a particular investment or pursue a particular investment strategy, including whether or not to buy, sell, or hold any of the securities or issuers mentioned. It should not be assumed that investments in such securities or issuers have been or will be profitable. References to specific securities or issuers are to illustrate views expressed in the commentary and do not represent all of the securities purchased, sold or recommended for advisory clients.

Sustainable investment considerations are one of multiple informational inputs into the investment process, alongside data on traditional financial factors, and so are not the sole driver of decision-making. Sustainable investment analysis may not be performed for every holding in every strategy. Sustainable investment considerations that are material will vary by investment style, sector/industry, market trends and client objectives. Certain strategies seek to identify companies that we believe may be desirable based on our analysis of sustainable investment related risks and opportunities, but investors may differ in their views. As a result, these strategies may invest in companies that do not reflect the beliefs and values of any particular investor. Certain strategies may also invest in companies that would otherwise be excluded from other funds that focus on sustainable investment risks. Security selection will be impacted by the combined focus on sustainable investment research assessments and fundamental research assessments including the return forecasts. These strategies incorporate data from third parties in its research process but do not make investment decisions based on third-party data alone.

Capital Expenditure (CAPEX) are funds used by a company to acquire, upgrade, and maintain physical assets such as property, industrial buildings, or equipment.

Power Purchase Agreements (PPAs) are contracts between two parties, one which generates electricity (the seller) and one which is looking to purchase electricity (the buyer).