



Investing in GPUs

Paper 2: supply side constraints



MM
NUWAY
CAPITAL



GPU investing: supply-side constraints and new business models

The recent explosion of Artificial Intelligence (AI), cloud computing, and the Internet of Things (IoT) has created significant demand for GPUs and data centres (DCs), rapidly transforming them into a key alternative investment class. As discussed in the first paper of this series, the recent growth in AI, in particular, has fuelled vast data processing demands and driven substantial gains for sector champions, whose valuations have surged as they consolidate their position at the heart of the AI infrastructure boom.

At the same time, supply constraints – including limited semiconductor manufacturing capacity, power grid shortages, real estate scarcity, and disrupted supply chains – have intensified competition and boosted potential returns for investors. In this paper, we examine the nature and impact of these constraints, as our series continues to explore key investment trends within the sector.





GPUs: a new investor favourite

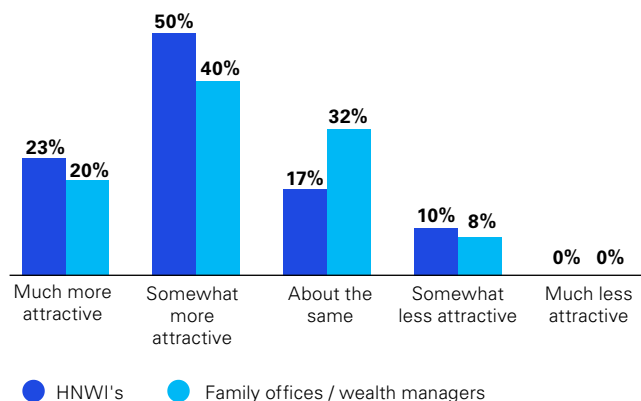
Investors seeking exposure to the digital megatrends shaping the 21st century have many sectors to choose from, including renewable energy, cloud computing, gaming, social media, artificial intelligence, blockchain technology, and quantum computing amongst others. These industries have different underlying drivers, challenges and levels of maturity with capital deployment also varying across each one. However, a common thread connecting these sectors is their reliance on GPUs and DCs as the underlying physical infrastructure that enables their development. Hence, GPUs have offered investors broad exposure to digital transformation, which goes some way to explaining the popularity of this relatively new asset class.

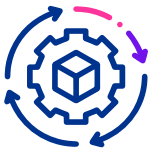
Major producers have demonstrated strong financial performance in recent years, as sustained GPU scarcity has boosted margins and stock valuations across the sector. NVIDIA, for example, with a market capitalisation in the trillions has become a driver of broader stock market performance alongside other technology related mega cap players. The competitive landscape is rapidly evolving, with market rivalry poised to intensify as Google's AI-native Tensor Processing Units (TPUs) become available to third-party developers. In particular, speculation that Meta may lease TPUs from Google has already triggered negative stock price reactions for incumbent players. According to our survey data, this combination of indispensable real-world utility, limited supply, and robust returns has made GPUs an attractive emerging technology asset class in comparison to others such as blockchain and quantum computing.



GPU investor appeal Compared to other emerging technologies (e.g. blockchain, renewable energy, quantum computing), how attractive do you view GPUs as an investment opportunity?

% of respondents, N=120





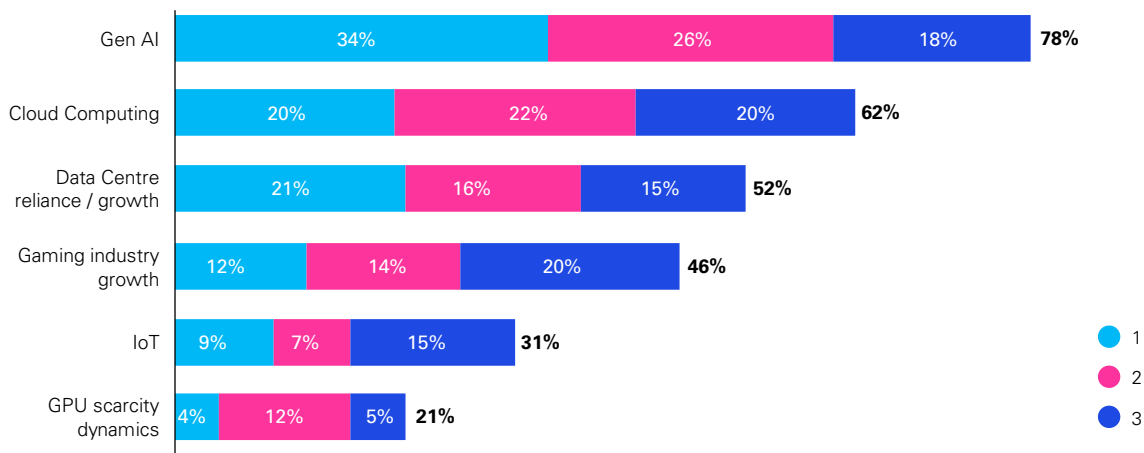
Supply and demand

Investor enthusiasm for GPUs is driven largely by explosive demand from AI, machine learning, and cloud computing. With governments and industries continuing to unearth new AI use-cases the perception of virtually limitless global demand has taken hold, with AI models exponentially increasing in quantity and size over the past few years. This dynamic has made GPUs one of the most attractive asset classes in today's tech-driven landscape.

The AI boom

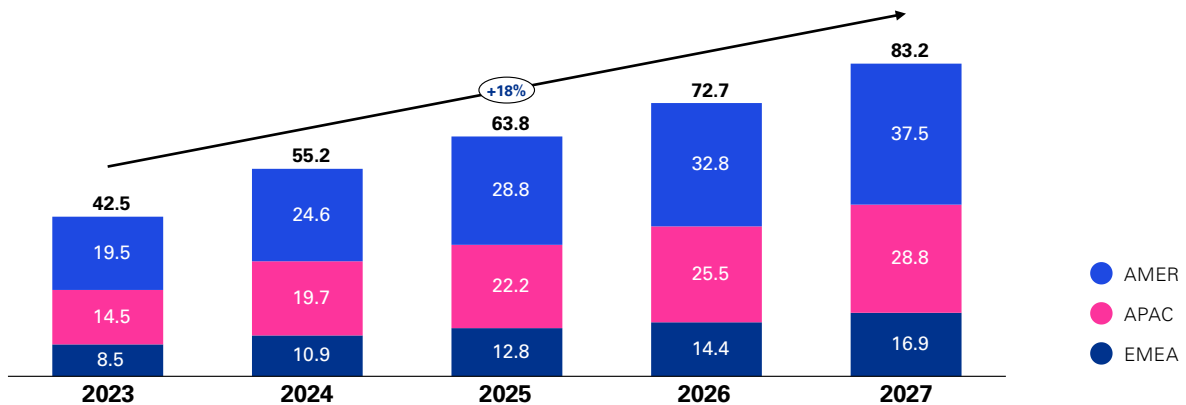
What are the most attractive themes / trends driving your interest in investing in GPUs (rank top 3)?

% of respondents, N=85



However, while surging demand is the primary force driving GPU market growth, supply-side constraints are also central to investor interest with global data centre capacity expected to grow at a CAGR of 18% out to 2027 (in contrast to exponential demand growth). Global supply chains remain disrupted by an uncertain geopolitical outlook and major players like NVIDIA have been hit by production problems caused by earthquakes and design flaws, hampering efforts to meet spiralling demand¹.

Forecast global data centre capacity (gigawatts), 2023 – 2027^{2,3,4}



Source:

- <https://www.reuters.com/technology/nvidias-supply-snags-hurting-deliveries-mask-booming-demand-2024-11-21>
- <https://www.jll.ie/en/trends-and-insights/research/data-center-outlook>
- <https://www.goldmansachs.com/insights/articles/ai-to-drive-165-increase-in-data-center-power-demand-by-2030>
- KPMG Analysis

Limits on new DC capacity have compounded these issues, as cloud providers battle to secure scarce GPUs to maximise existing capacity. Several factors contribute to this bottleneck, including:

- **Real estate shortages:** DCs (especially hyperscalers), must be sited strategically to ensure their power, cooling, connectivity, and security demands can be met. These specific requirements make suitable sites increasingly scarce in and around major technology hubs.
- **Grid limitations:** with DC energy consumption growing exponentially, many locations are struggling particularly with grid capacity, necessitating restrictions and delays to new construction.
- **Sustainability restrictions:** the environmental impacts of DCs has also become a more high profile issue, causing governments to propose stricter regulations on energy and water consumption, particularly in Europe and the USA.
- **Complex operational standards:** issues such as the complexity of cooling systems during the build-out phase, the challenges of retrofitting with water-cooling solutions, and long development cycles all reduce the speed to market for new DC capacity.
- **Skills shortages:** DC construction is a high-skill enterprise, with demand for engineers and technicians consistently outpacing supply.
- **Capital requirements:** DC construction demands significant capital investment, and rising interest rates and inflation have made financing hyperscale projects harder in recent years.

Investors see such constraints as a source of long-term price stability for GPUs, enhancing their appeal and resilience compared to other technology assets. As AI adoption accelerates, GPUs will remain a supply-limited yet essential resource, positioning them as a key asset in high-growth digital portfolios.



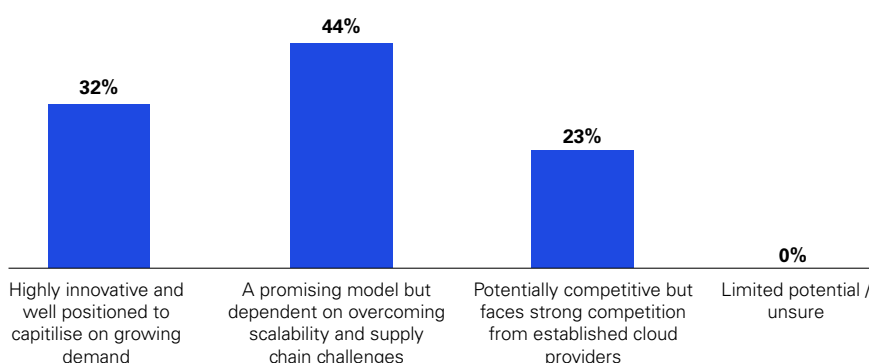


As high infrastructure costs and supply bottlenecks make it harder for businesses to deploy their own GPU hardware, established providers, startups, and investment vehicles are responding with new business models such as GPU-as-a-Service (GPUaaS): a cloud-based model that lets end users rent GPU processing power on demand, rather than grapple with the complexities and expense of owning their own hardware.

As an innovative tool to loosen the supply bottleneck, GPUaaS has emerged as a vital resource for startups, research institutions, and other cash-constrained players with an urgent need for access to high-performance computing. For end users, it offers flexibility, scalability, and cost-efficiency, and has gained rapid traction in recent years as GPU demand outpaces supply. Our survey of professional investors and high-net-worth individuals indicates a view that GPUaaS offers reliable recurring revenue opportunities underpinned by strong fundamentals. However, it must overcome scalability and supply chain challenges to realise its potential.

How do you perceive the 'GPU as a Service' business model in addressing market needs for high-performance computing?

% of respondents, N=120



Such challenges aside, with AI's increasing sophistication continuing at pace, GPUaaS is well positioned to capitalise, offering a multitude of benefits to data rich businesses, for example:

- **Avoid high capital costs:** building data centres with GPU clusters requires massive upfront investment, making it cost-prohibitive for the majority of businesses. Smaller companies, researchers, and developers can access high-performance GPUs without requiring their own infrastructure.
- **Sidestep supply chain constraints:** limited semiconductor supply and bulk-buying by large tech companies have made it harder for smaller firms to directly acquire GPUs.
- **Manage growth:** GPUaaS allows businesses to scale up or down as needed, paying only for the computing power they use, rather than maintaining expensive, underutilised hardware.
- **Work with trusted providers:** companies like AWS, Microsoft Azure, Google Cloud, and NVIDIA's DGX Cloud now offer GPUaaS solutions, enabling customers to run GPU-intensive workloads remotely.

GPU boom appears set to continue for now

As GPU demand grows and supply struggles to keep pace, investors will continue to seek exposure to the AI infrastructure boom. Emerging technologies such as analogue AI and Field-Programmable Gate Arrays (FPGAs) pose potential challenges to GPUs. For now, however, GPUs remain the dominant hardware powering AI. New business models like GPUaaS offer an attractive entry point, democratising access to high-tier computing using, reducing barriers to participation in the GPU economy and yielding dependable cash flow for investors. This KPMG series aims to equip interested readers with the insights needed to navigate the rapidly evolving market landscape. Our next piece will look to contextualise GPU investing within the bigger picture of alternative investments, to help investors understand their relative risks and returns as well as their potential within a diversified portfolio strategy.

Please note: this series does not constitute investment advice.

Contact Us:



Chris Brown
Partner
KPMG Strategy
t: +353 87 050 4453
e: christopher.brown@kpmg.ie



Colin Boshier
Founder, Nuway Capital Ltd.
t: +44 (0) 203 876 2887
e: colinboshier@nuwaycapital.com



Gerard Duffy
Director
KPMG Strategy
t: +353 87 111 6976
e: gerard.duffy@kpmg.ie



William Fielding
Founder, Nuway Capital Ltd.
t: +44 (0) 203 876 2887
e: william@nuwaycapital.com



Claire Mitchell
Manager
KPMG Strategy
t: +353 87 001 0629
e: claire.mitchell@kpmg.ie



Stephane Hélène
Founder, Nuway Capital Ltd.
t: +44 (0) 203 876 2887
e: stephane@nuwaycapital.com



Tony Sheehan
CFO, Nuway Capital Ltd.
t: +353 86 735 2365
e: tony.sheehan@nuwaycapital.com



Visit Nuway Capital
webpage:



kpmg.ie

© 2026 KPMG, an Irish partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation. The KPMG name and logo are registered trademarks of KPMG International Limited ("KPMG International"), a private English company limited by guarantee. If you've received this communication directly from KPMG, it is because we hold your name and company details for the purpose of keeping you informed on a range of business issues and the services we provide. If you would like us to delete this information from our records and would prefer not to receive any further updates from us please contact unsubscribe@kpmg.ie.

Produced by: KPMG's Creative Services. Publication Date: January 2026. (11563)