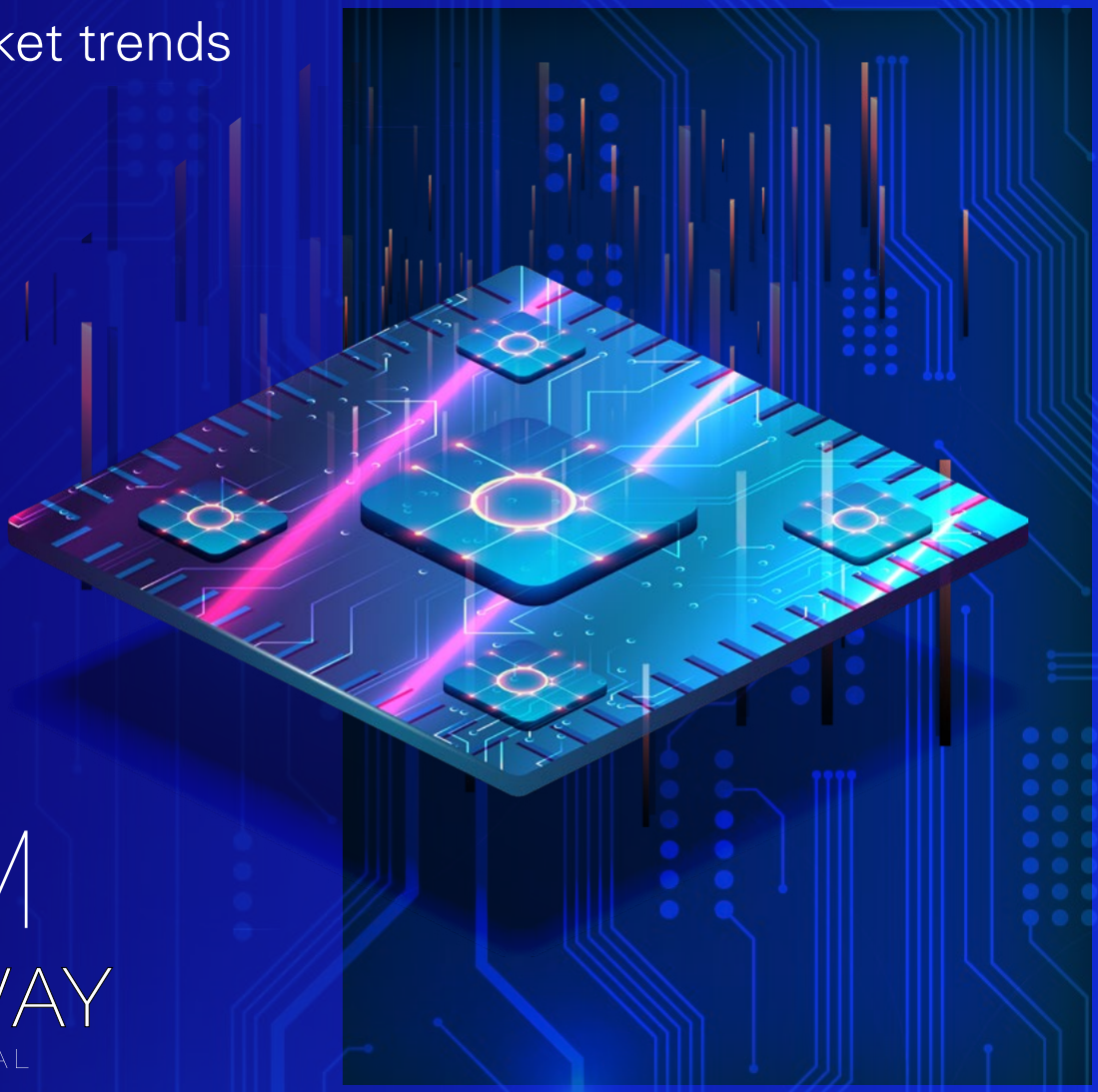
KPMG

Investing in GPUs

Key demand drivers and market trends



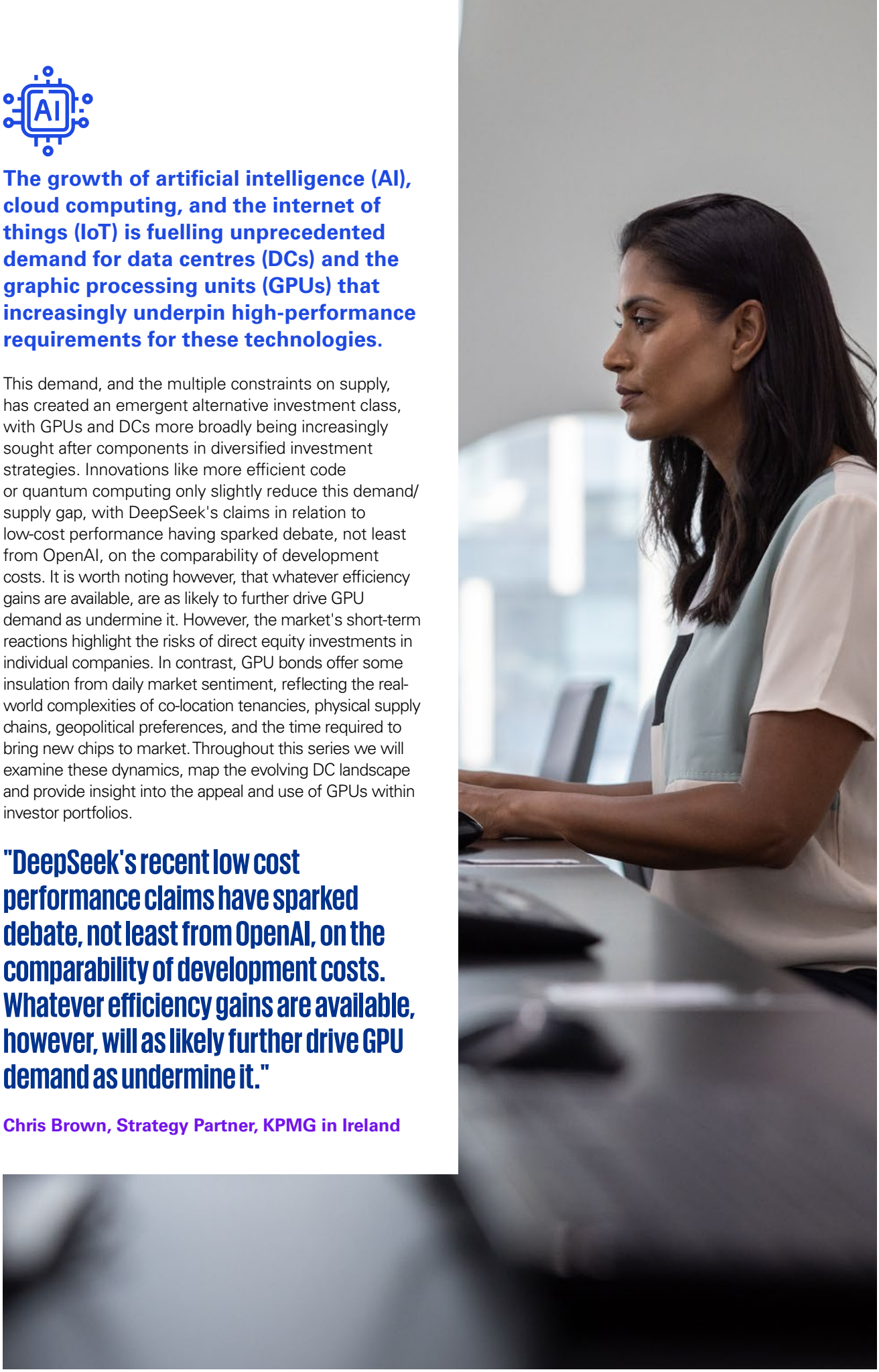
NUWAY CAPITAL

The growth of artificial intelligence (AI), cloud computing, and the internet of things (IoT) is fuelling unprecedented demand for data centres (DCs) and the graphic processing units (GPUs) that increasingly underpin high-performance requirements for these technologies.

This demand, and the multiple constraints on supply, has created an emergent alternative investment class, with GPUs and DCs more broadly being increasingly sought after components in diversified investment strategies. Innovations like more efficient code or quantum computing only slightly reduce this demand/supply gap, with DeepSeek's claims in relation to low-cost performance having sparked debate, not least from OpenAI, on the comparability of development costs. It is worth noting however, that whatever efficiency gains are available, are as likely to further drive GPU demand as undermine it. However, the market's short-term reactions highlight the risks of direct equity investments in individual companies. In contrast, GPU bonds offer some insulation from daily market sentiment, reflecting the real-world complexities of co-location tenancies, physical supply chains, geopolitical preferences, and the time required to bring new chips to market. Throughout this series we will examine these dynamics, map the evolving DC landscape and provide insight into the appeal and use of GPUs within investor portfolios.

"DeepSeek's recent low cost performance claims have sparked debate, not least from OpenAI, on the comparability of development costs. Whatever efficiency gains are available, however, will as likely further drive GPU demand as undermine it."

Chris Brown, Strategy Partner, KPMG in Ireland



Global demand for GPU capacity

Across geographies, public and private sector entities are working at pace to build the infrastructural foundations to enable the digital revolution. The transformation of AI from peripheral to business-critical technology has prompted a mass adoption of the advanced GPUs necessary to handle its huge workloads, leading to a level of demand that could outstrip supply for years to come. Furthermore, the ever-expanding IoT landscape that connects millions of devices and generates significant data volumes, all of which must be stored and processed in real time, requires millions of square feet of new DC capacity even as computing density continually increases in those DC halls. We expect these powerful trends to persist, as the premium on new innovation grows, and governments and businesses rush to realise the vast promise of these new technologies.

Backbone of the digital revolution

GPUs and DCs play indispensable roles in the national technology stack. GPUs are engineered to support the considerable processing demands of modern AI applications, and provide an integral support structure for applications including machine learning, graphics rendering, crypto mining, scientific modelling, and other cutting edge modern technologies. As such, they are increasingly key components in DCs, providing the necessary computational resources to store, process, and distribute the vast amounts of digital information such applications generate. DCs are the physical infrastructure where such computational power, storage, and networking reside, and provide the foundation for almost all private and public online services.

These critical functions have seen both GPU makers and DC builders enjoy tremendous growth in recent years. For example, NVIDIA has become a market dominant player on the strength of its most advanced chips. Today, the global GPU market is estimated at around $65bn, and expected to hit $274bn by 2029, at a CAGR of over 33%[1]. The DC market, meanwhile, is estimated at around $256bn today, and expected to hit $450bn by 2029,[2] creating a joint market size of well over half a trillion dollars.

These headline figures obscure significant regional differences. The US remains the dominant GPU market, accounting for higher volumes and revenues than anywhere else in the world. However, Asia-Pacific is not far behind and is growing fastest, buoyed by manufacturing dominance and increasing demand from tech hubs such as China, Japan, and South Korea. As in other sectors, growth in Europe is slower. However, the region's regulatory features are set to grow demand in the sector, with mandates requiring proximity of data storage and processing to end users increasingly fostering the development of country-specific 'AI clouds'. This creates a necessity for duplicate operations, amplifying demand for GPUs and other critical hardware. Simultaneously, the increasing sophistication of cyber-attacks are creating a need for separate or on-premises DCs and AI-driven security measures, all of which further inflate GPU demand.

Source:

(1). https://www.mordorintelligence.com/industry-reports/graphics-processing-unit-market

(2). https://www.globenewswire.com/news-release/2024/09/18/2948393/0/en/Data-Center-Market-Size-Expected-to-Reach-USD-775-73-Billion-by-2034.html.

Exploding demand, limited supply

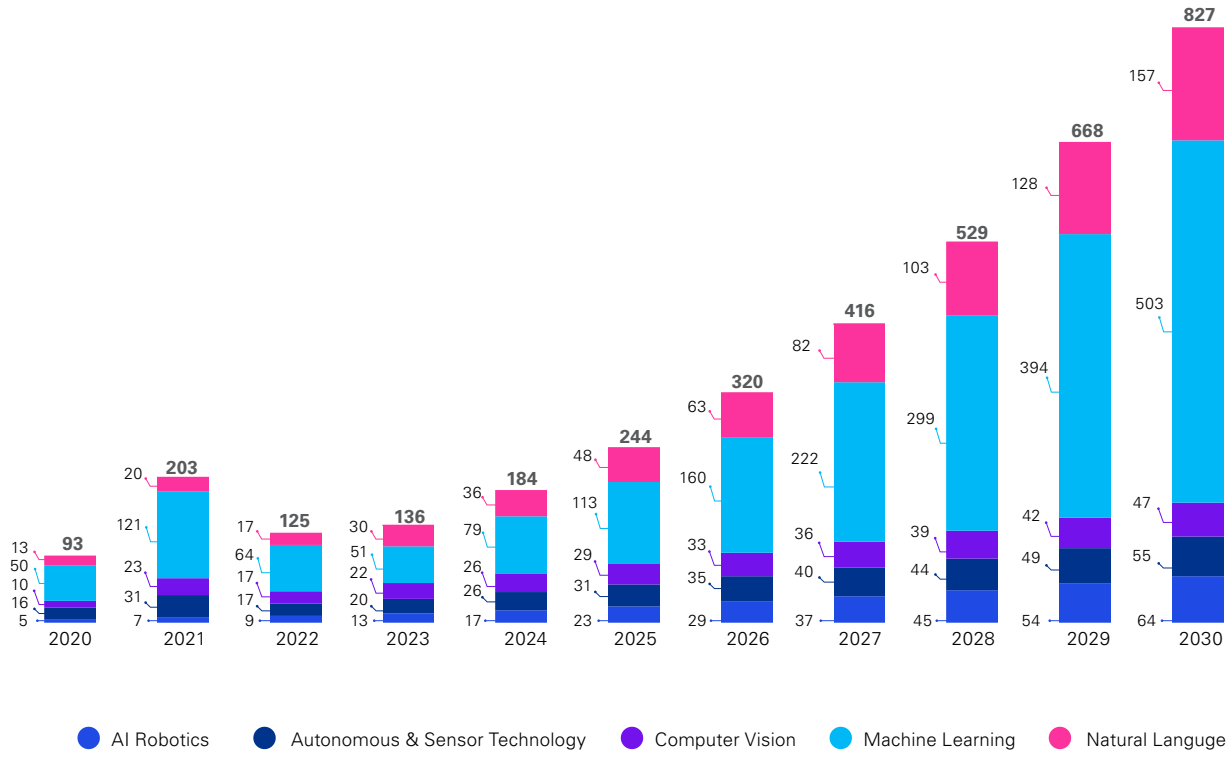
Demand for both GPUs and DCs is closely aligned to other markers of AI and IoT adoption, both of which are spawning applications across almost every sector, from agriculture to energy management and supply chain optimisation, driving significant investor activity. While much hype surrounds AI, and the promise is often several years ahead of the reality, many forecasts predict that AI will be a trillion-dollar market by the end of the decade.

These constraints have seen demand outstrip supply for both asset types, leading to intense competition. It is now commonplace for customers to pre-lease DC capacity well in advance of operation, while smaller players often struggle to source GPUs as larger ones bulk-buy all available stock. Whilst this backdrop of competition and constrained supply presents operational problems for many players in the value chain, it makes both GPUs and DCs compelling investment opportunities within the AI infrastructure landscape.

Whilst demand for GPUs and DCs is substantial, supply is currently constrained by a wide range of factors, including:

- Lack of suitable real estate
- Lack of available grid capacity
- Disrupted supply chains
- Skilled labour shortages
- High technical, environmental, and regulatory operational standards
- Large capital requirements
- Manufacturing bottlenecks
- Long production development cycles

Fig 3: AI global market size forecast, USD bn



Source:
(3). https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide#market-size

Investor channels

Recent returns on investments in both GPUs and DCs have often outpaced those in traditional asset classes. NVIDIA's share price climbed almost 200% in 2024, reflecting high returns across the sector, whilst some DC providers have also recorded double digit growth in their stock prices. However, investors have options beyond direct equity with multiple direct and indirect channels available:

Direct routes include:

- Equity in GPU manufacturers and distributors, such as shares in GPU-focused companies like NVIDIA or AMD, or through GPU-focused ETFs.
- Direct investment through physical ownership of GPU servers, via an alternative investment vehicle.
- Specialised funds targeting tech infrastructure and semiconductor sectors.
- Equity in third-party cloud providers (e.g., AWS) leasing DC and GPU server processing capacity through a "GPU as a Service" (GPUaaS) business model.
- Equity in those involved in the build, fit-out and / or retrofitting of DCs, from design and engineering specialists through to manufacturers in the supply chain. However, the majority of fast growing players in the space are not yet listed with investment via Private Equity (e.g. Blackstone with Winthrop, Exponent with Ethos), and those that are listed tend to be part of wider multi-sector conglomerates (e.g. Siemens, ABB, Schneider).

Indirect routes include:

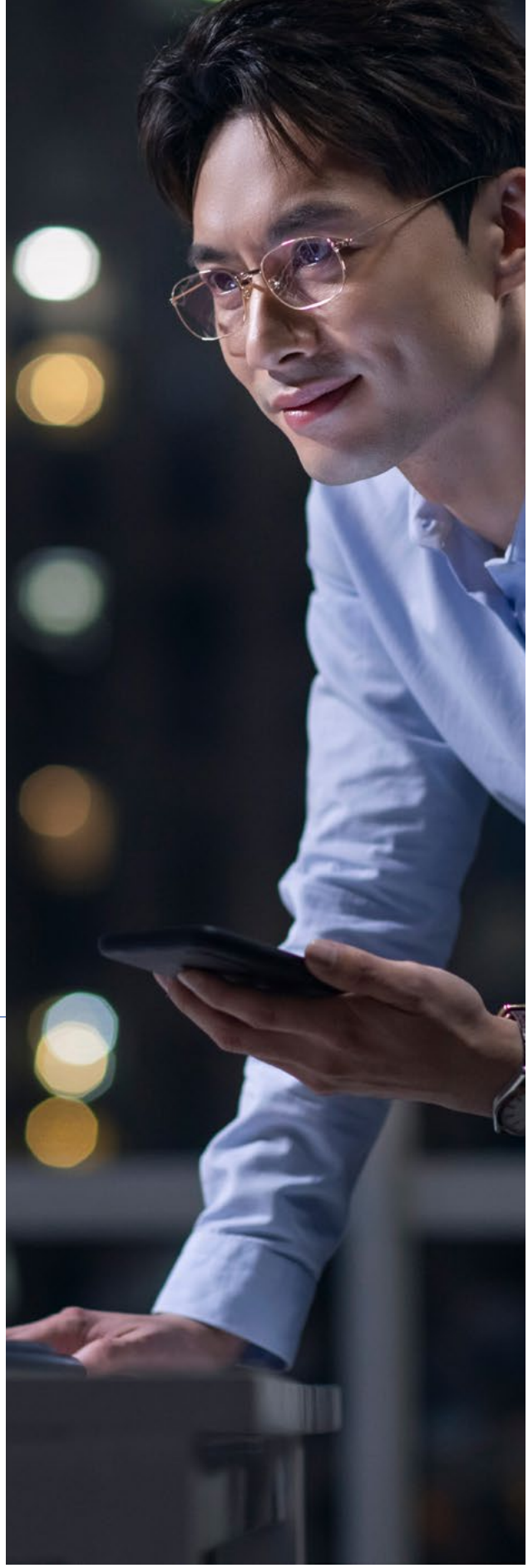
- GPU infrastructure funds, which invest in data centres (DCs), cloud computing, AI and machine learning companies reliant on GPU technology, and futures or options tied to GPU-related markets, technologies, or commodities.
- Blended unit trusts or mutual funds that include GPU-related assets.
- Technology-focused venture capital or private equity funds with GPU exposure.

Each approach has it pros and cons and will appeal to the varying risk appetites of different investors. For example, direct investment into GPU manufacturers - as with any single-stock investment - is susceptible to market volatility and company-specific challenges such as intensifying competition, regulatory changes, and internal cost pressures. For those seeking a lower risk approach, GPU bond structures, such as the one offered by Nuway Capital offer an alternative. These types of models aim to combine the security and elevated returns of physical GPU ownership with the advantages of pooling resources across multiple assets to reduce risk. Later in this series of papers, we will take a deeper-dive into how these different types of structures compare to other alternative assets and their role within broader investment portfolios.

Investor sentiment

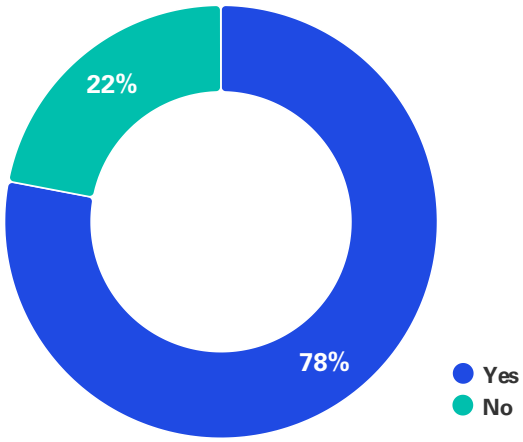
Our research – consisting of a survey of 120 high-net-worth individuals (HNWIs) and investment professionals – indicated that a significant cohort of market participants have already gained exposure to the GPU market (78% of those surveyed). ETFs are the preferred vehicle of choice, with 71% of advisers and 62% of high-net-worth individuals (HNWIs) investing in them in recent years. Top GPU-related ETFs in the market include the VanEck Semiconductor ETF and the iShares Exponential Technologies ETF, the former of which has a material NVIDIA skew, accounting for ~20% of the fund's total exposure. This showcases a strong preference, in particular within HNWIs, for diversified exposure to the GPU market (62% preference versus 38% for direct investment in GPU manufacturers).

GPU bonds were also a prominent choice amongst our survey audience, particularly among investment professionals. Some analysts have suggested that NVIDIA's corporate bonds are a safe-haven investment within the GPU-landscape due to the company's substantial free cash flow, minimal leverage, and strong market position in AI infrastructure, offering bond investors confidence in its ability to meet interest payments and make principal repayments[4].



Have you ever invested in or advised investing in GPU-linked investment(s)?

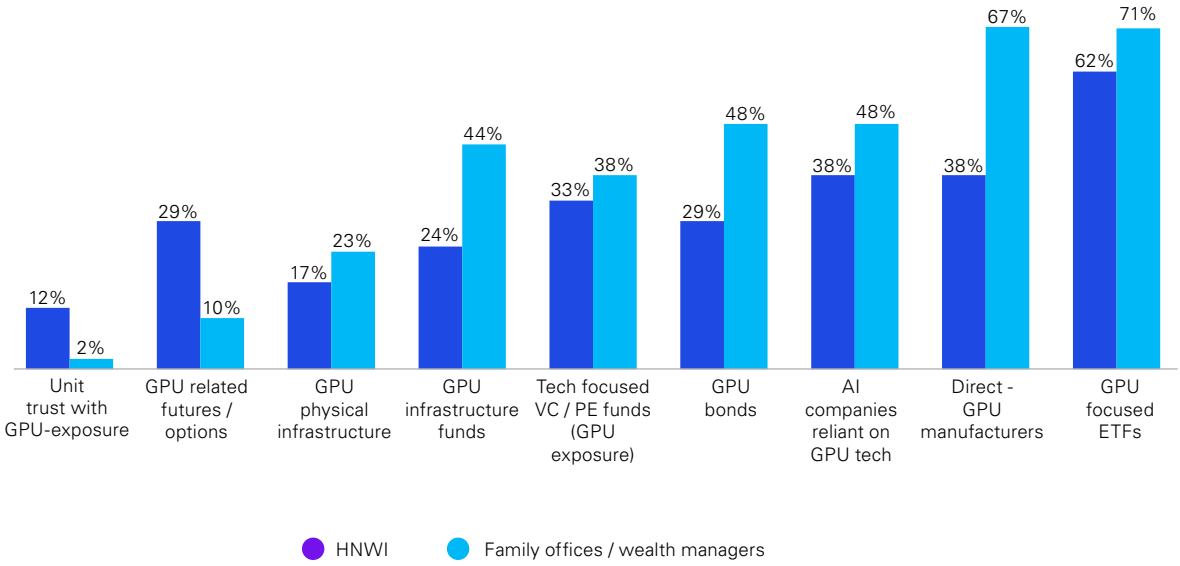
% of respondents, N=120



Source:
(4). https://www.morningstar.com/news/marketwatch/20240622223/why-nvidias-bonds-are-a-safe-haven-alternative-to-its-high-flying-stock

Which type(s) of GPU investment(s) have you invested in? / have you advised investing in?

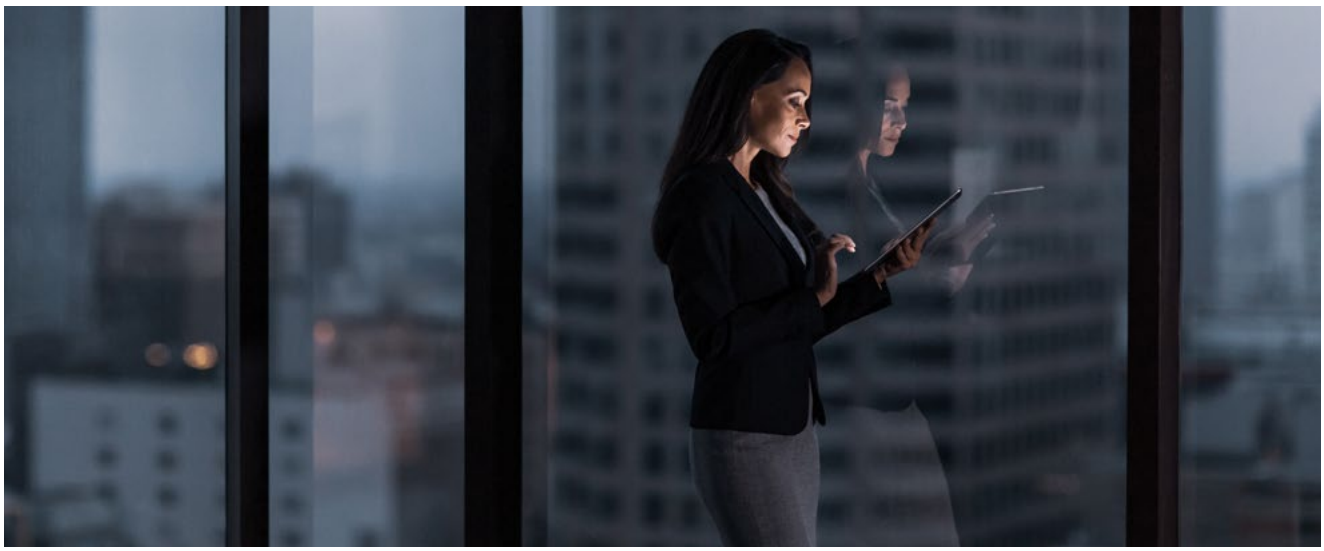
% of respondents, N=94



Another investment concept that has emerged within the broader bond category, are structures that utilise the GPUaaS business model[5]. From the end user perspective, this is a cloud-based business model that provides businesses and developers with on-demand access to powerful GPU resources, enabling them to efficiently run computationally intensive tasks like AI training, rendering, and data processing without owning physical hardware. Via a bond structure, investors can access returns (via a fixed coupon rate) associated with the rental income from the use of these GPUs.

A clear and present opportunity

The scarcity of GPUs and DC capacity is unlikely to be a blip but a sustained feature of the digital revolution, as supply struggles to match demand in these strategically critical hardware components. For investors, this amplifies their appeal as high-growth asset classes. This KPMG series aims to equip readers with the market insights needed to navigate the rapidly evolving GPU investment landscape. Our next piece will examine supply-side constraints on GPU production in greater detail, as well as their likely impact on market dynamics.



Source:
(5). https://www.thearmchairtrader.com/features/nuway-capital-nvidia-gpu-investment/

Please note: this series does not constitute investment advice.

Contact Us:

Chris Brown
Partner, Head of Strategy
KPMG in Ireland
t: +353 87 050 4453
e: christopher.brown@kpmg.ie

Gerard Duffy
Director, FS Strategy Lead
KPMG in Ireland
t: +353 87 111 6976
e: gerard.duffy@kpmg.ie

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

Molly Boyne
Consultant, Strategy
KPMG in Ireland
t: +353 86 191 3324
e: molly.boyne@kpmg.ie

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

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

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

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

Visit Nuway Capital webpage:

kpmg.ie

Produced by: KPMG's Creative Services. **Publication Date:** February 2025. (11127)