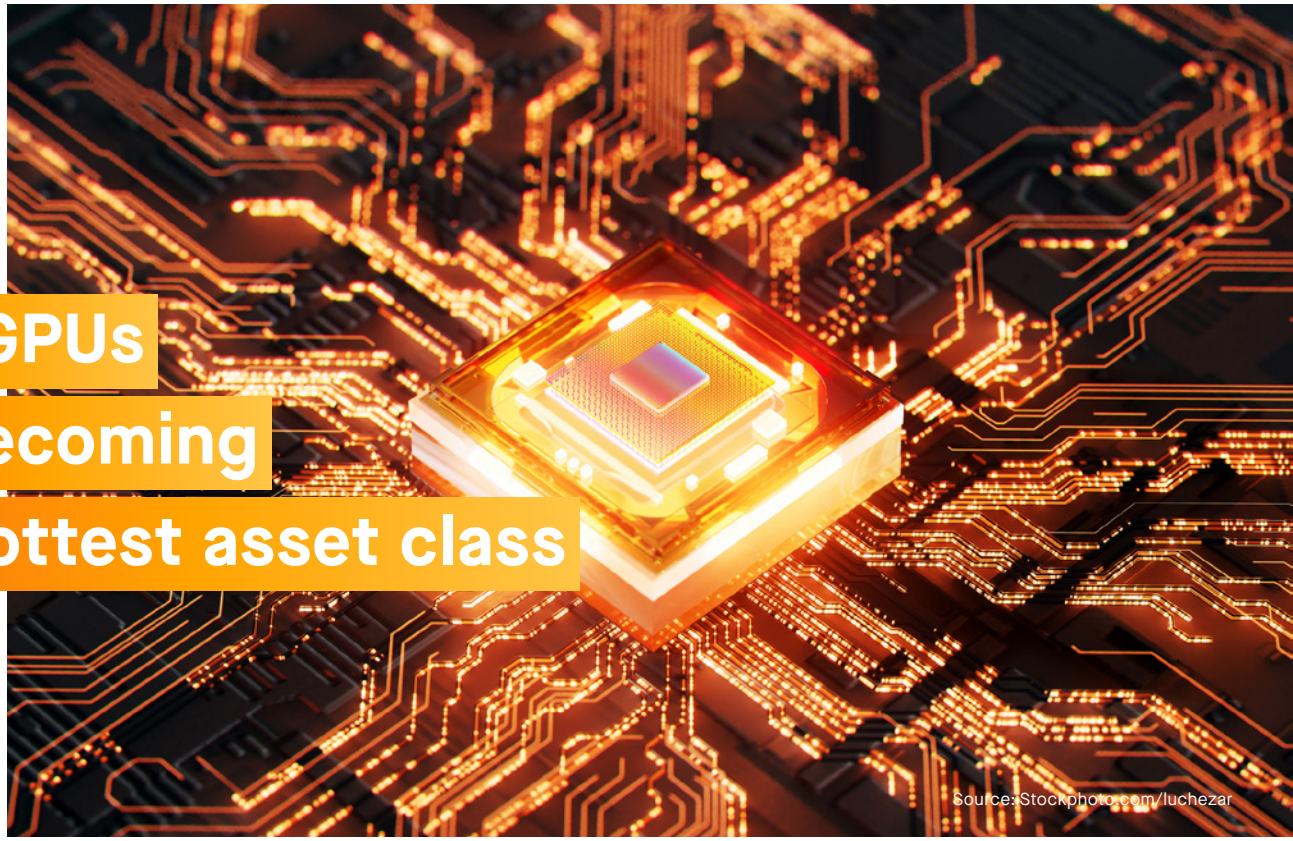
How GPUs are becoming the hottest asset class



Not long ago, GPU chips were a gamer's luxury. Today, they are rapidly evolving into a distinct asset class, one that investors can finance, collateralize, lease, and even treat as yield-generating infrastructure.

Charles-Henry Monchau, CFA, CAIA, CMT
Chief Investment Officer
charles-henry.monchau@syzgroup.com

Assia Driss
Syz Research Lab Team Coordinator
assia.driss@syzgroup.com

Hugo Morel
Syz Research Lab Team
hugo.morel@syzgroup.com

Sophia Houghton
Syz Research Lab Team
sophia.houghton@syzgroup.com

Syz PRIVATE BANKING

Introduction

In April 2026, KPMG and UK-based alternative investment firm Nuway Capital published their joint research series on GPU investing. In a survey of 120 high-net-worth individuals, family offices, and wealth managers across ten international markets, they found that technology-related assets have overtaken real estate and private equity as the most widely held thematic area within alternative portfolios. SPVs, GPU leasing structures, and the tokenization of physical compute resources are creating alternative pathways for capital deployment, allowing corporations to transfer title to assets and offer investors secured, asset-backed cashflows. This is the beginning of a new investment category.

From gaming chips to financial assets

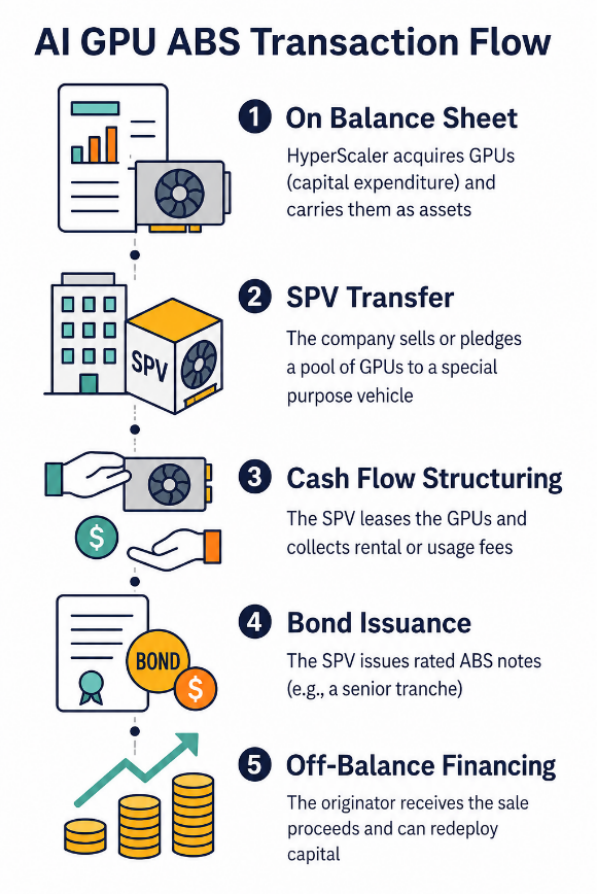
For most of the past decade, GPUs were a gamer's concern, specialized chips designed to render pixels faster than the human eye could track. What set them apart from conventional processors was architecture: where a CPU (Central Processing Unit) executes tasks sequentially, a GPU runs thousands of calculations in parallel. That distinction mattered little outside of gaming and visual effects, until the demand for AI models surged. Coincidently, the same parallel processing power that rendered video game landscapes turned out to be exactly what the technology needed to train vast datasets, leading to a dramatic increase in GPUs' economic value. To some, it is a hard asset, physical hardware that can be financed, leased and lent against, much like an aircraft or a shipping container that has a depreciation schedule and collateral value. To others, it is an income stream, a networked cluster converting electricity into computation, generating recurring cash flow the way a toll road or a power plant does.

Increasingly, GPUs are also being treated as a commodity. This reflects two related developments. First, the computing power generated by GPUs is becoming an interchangeable resource that can be bought, sold, and priced. As companies require growing amounts of computing capacity to train and run artificial intelligence models, many now purchase access to compute rather than the hardware itself. This creates a market where GPU-hours are treated as a standardised product. Second, there is a growing global trade in the physical graphics cards themselves, with demand for advanced chips driving significant cross-border trade and investment, particularly in response to supply shortages and changing market conditions. Together, these trends have shifted the perception of GPUs from specialised pieces of technology to economic resources with commodity-like characteristics. Their value has shifted from the silicon itself to the computation it produces and the financial ecosystem that has grown around it.

In financial markets, a commodity becomes investable once it can be standardised: a unit of measure, an agreed price, and a market deep enough to transfer risk. Oil was a commodity before it was an asset class; what changed was the arrival of benchmarks, futures, financing and risk-transfer tools. The same transition is visible in compute.

Today, the growing commoditisation of compute has accelerated the acceptance for GPUs to function as an asset class. Until recently, there was no agreed price for a GPU-hour, with costs varying significantly across providers, regions, and contract terms. This has begun to change through the introduction of benchmark indices, including Ornn AI's compute price index on the Bloomberg Terminal and Silicon Data's rental indices, as well as the launch of compute futures by CME and ICE.

Equally important has been the expansion of financing and securitisation. Companies such as CoreWeave have secured billions of dollars in GPU-collateralised debt Lambda, an AI infrastructure company, completed one of the first GPU-backed asset-backed securities (ABS) transactions. Top-tier lenders such as BlackRock, Blackstone, PIMCO, Carlyle, JP Morgan, Macquarie, are active, and rating agencies have moved from an effective A+ ceiling on data-centre ABS toward higher senior ratings. The GPU financing market was estimated at between USD 20 billion and USD 34 billion by early 2026.



Source: Syz Bank

In addition, risk-transfer tools have started to emerge. To address concerns surrounding the uncertain future value of GPU hardware, specialist insurers now offer residual-value protection, while rating agencies such as Fitch have begun exploring how GPU depreciation should be incorporated into securitisation frameworks.

Finally, increasing standardisation in financing structures has strengthened the market further. The use of special purpose vehicles (SPVs), long-term take-or-pay agreements, and investment-grade counterparties has created financing arrangements that increasingly resemble those found in traditional infrastructure projects. Together, these developments have provided the foundations for GPUs to be valued, financed, and traded in ways that increasingly resemble established asset classes.

The investment case

If the previous section established what the GPU asset class is, this one asks the more consequential question: should you own it?

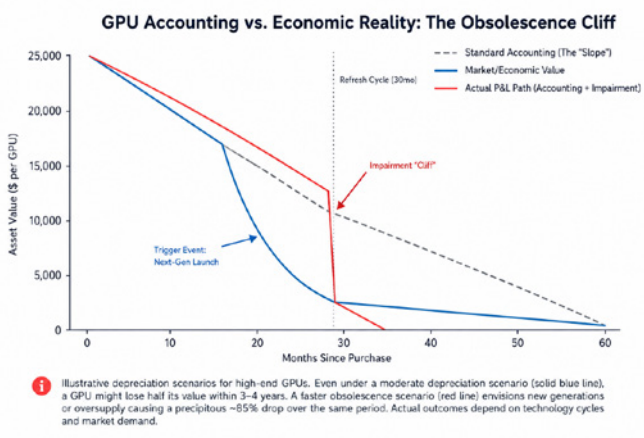
On the positive side, the cash yield is tangible. Unlike most AI exposures, GPU infrastructure can generate contractual income. An H100 bought for roughly $25,000-$40,000 can rent for $2.00-$3.50 per GPU-hour, implying around $17,500-$30,000 of annual revenue at high utilisation, before operating costs. Equipment lenders have extended GPU-backed loans at 8 to 12%; structured neocloud financings and direct cluster ownership have been marketed at net IRRs in the mid-teens to mid-twenties over three-to-five-year horizons, secured against hard assets. The cash-flow profile is closer to energy or transport infrastructure than to venture capital.

The demand side is also compelling. The capital-spending cycle anchoring this theme is the largest in corporate history. Hyperscaler AI commitments are approaching the trillion-dollar mark by 2027. Neoclouds are becoming the new layer of dedicated compute capacity between the chip makers and the largest technology buyers. Even "obsolete" Hopper-class H100 rents did not fall monotonically as newer generations arrived. They rose roughly 40% between late 2025 and early 2026 amid a capacity crunch, with on-demand supply effectively sold out. Demand for compute has, so far, repeatedly outrun the depreciation curve.

For portfolio construction, the diversification argument is coherent as well. Structured GPU credit and infrastructure show relatively low correlation to public equity and bond markets, returns are driven by utilisation and contracted cash flow, not multiple expansion. The theme also fits a nominal-growth regime: it is a real, productive asset whose pricing power is tied to an input in structural shortage. For an alternatives allocation increasingly tilted toward private credit and real assets, GPUs offer a way to participate in AI infrastructure while collecting yield, rather than paying a premium for listed AI equities.

One point is worth clarifying, because it is often lost in the headline debate. The chip itself is not the real collateral. The core asset is often a five-year take-or-pay contract with a strong counterparty. The debt amortises over the contract life, the SPV ring-fences the assets and cash flows, and residual GPU value becomes a secondary protection. That makes the asset class more credible than the "lending against depreciating chips" critique implies.

However, the GPU case has limits. The biggest issue lies in how long a GPU holds its value. Even the largest technology companies disagree. Amazon shortened server useful life from six years to five in early 2025, Meta extended it to five and a half years. Microsoft, Alphabet, and CoreWeave still use six. When the best-informed buyers in the world disagree by roughly 20% on depreciation assumptions, collateral value is uncertain. A chip worth $25,000 in 2026 may be worth a fraction of that three years later as future generations arrive.



Source: Syz Bank

Price volatility compounds the picture. Rental economics in this market are cyclical. H100 rental rates peaked near $8 per hour, collapsed to $1–2 on oversupply, recovered roughly 40% into early 2026, then softened again by May 2026. The pricing pattern for Nvidia's B200 points to the same fragility: despite being one of the newest high-end GPUs, marketplace rates have already moved sharply within a few months. The same GPU can look like a high-yielding infrastructure asset or a loss-making machine depending on where in the cycle it is financed.

Source: Theory Ventures, ORNN data

There is also ecosystem concentration. NVIDIA sells to neoclouds, which borrow against GPUs to buy more GPUs, often backed by a small number of hyperscaler or AI-lab customers. CoreWeave illustrates both the opportunity and the fragility: rapid growth, but also customer concentration, leverage, periods of negative operating cash flow, and refinancing pressure. The 2008 analogy is probably too extreme, because today's structures are usually backed by amortising debt and real contracts and not layers of synthetic leverage. But the concentration and leverage risks are still very real.

The chips are also at risk of becoming the wrong chips. Custom silicon from hyperscalers, Google TPUs, Amazon Trainium, AMD's roadmap, export controls, and the move from training to inference all threaten the resale value and pricing power of today's GPU fleets. Export controls introduce a wildcard for secondary-market values that is difficult to price. And operational execution remains underweighted in most financial models: utilisation, power density and cooling are where returns are actually made or lost. Several "AI-ready" facilities marketed to investors have lacked the infrastructure to deliver on the label.

Conclusion

GPU exposure runs from a listed share to a warehouse full of chips, and the closer a position sits to contracted cash flows, the more it behaves like a genuine asset class. Listed equity in companies offers high liquidity in the AI infrastructure buildout, but with full equity volatility attached. Private credit sits closer to the ideal. Senior secured loans tied to long-term contracts with investment-grade counterparties, generating yields of 8-12%, with debt amortizing fully over the contract term and low liquidity. Direct cluster ownership can push returns higher, yet concentrates the hardest risks: depreciation, utilization, operational complexity. Compute futures offer directional exposure, though the market remains thin. Tokenized and decentralized compute networks sit furthest out on the risk spectrum, with variable liquidity and returns still closer to venture than infrastructure.

The debate around GPUs as a financial asset tends to generate more heat than light, as the two sides are rarely describing the same thing. Enthusiasts point to a financial infrastructure that materialized at remarkable speed: benchmarks, debt markets, insurance and futures. Skeptics point to contested residual values, heavy leverage and violent price swings. Both are right. They are simply looking at different parts of the same market. A GPU is a legitimate financial asset when what you own is contracted, amortizing cash flow secured against hard assets, and a speculative bet on a depreciating chip when it is not.

Welcome to Syzerland®

For further information

Banque Syz SA
Quai des Bergues 1
CH-1201 Geneva
T. +41 58 799 10 00
syzgroup.com

Charles-Henry Monchau, CFA, CAIA, CMT
Chief Investment Officer
charles-henry.monchau@syzgroup.com

Assia Driss
Syz Research Lab Team Coordinator
assia.driss@syzgroup.com

Sophia Houghton
Syz Research Lab Team
sophia.houghton@syzgroup.com

Hugo Morel
Syz Research Lab Team
hugo.morel@syzgroup.com