

Financing the AI Arms Race: Inside the Hyperscaler Capex Boom

Introduction

Artificial intelligence is often described as an intangible revolution. Its near-term economic footprint, however, is strikingly physical. Training and deploying frontier models requires large quantities of specialised chips, electricity, cooling, network equipment, and purpose-built data centre capacity. The AI boom is therefore not only a software story but an infrastructure investment cycle of unprecedented scale: the five largest hyperscalers are expected to spend roughly \$750bn in 2026 alone, equivalent to around 38% of their revenues. Yet these companies are also among the world's most profitable and cash-generative businesses. Why, then, are they increasingly turning to debt, leases and private capital to fund the build-out? As AI transforms the economics of Big Tech, the race for computing capacity is increasingly becoming a race for capital.

The Economics of Hyperscale

A hyperscaler operates computing infrastructure across a large network of data centres and can expand capacity rapidly as demand changes. Amazon [NASDAQ: AMZN] Web Services, Microsoft [NASDAQ: MSFT] Azure, Google [NASDAQ: GOOGL] Cloud and Oracle [NYSE: ORCL] Cloud Infrastructure sell this capacity through public-cloud services. Meta [NASDAQ: META] operates infrastructure at a similar scale but uses most of its infrastructure internally for recommendation systems and consumer applications. Its returns therefore come mainly from advertising and user engagement rather than cloud service revenue.

Historically, this was an attractive financial model. As usage increased, providers could spread their large infrastructure costs across more customers. Provided that capacity remained well utilised, this supported high margins and operating cash flow, while recurring contracts made revenue more predictable. Traditional cloud workloads were also relatively diversified, with software hosting and general-purpose computing distributed across customers and applications, reducing dependence on any single workload. Processors and servers still required replacement, but the infrastructure supported a broad range of uses and could often be redeployed as demand changed. Cloud providers could therefore invest ahead of demand, fill the resulting capacity over time and recover the expenditure through recurring revenue.

AI workloads change that balance. Training and serving large models require clusters of accelerators, high-bandwidth memory, storage and cooling systems. These assets are more expensive than conventional servers and are often installed in larger blocks before the associated revenue is fully visible. Capacity is also less interchangeable in the short term. A data centre designed around large GPU clusters and high power density cannot always be redirected towards ordinary enterprise workloads without affecting its economic return.

The revenue opportunity remains substantial. Demand for cloud infrastructure has accelerated, and several providers continue to report that available AI capacity is being absorbed as quickly as it comes online. Yet the timing of the economics has changed. Cash is committed when land and all the equipment is acquired, while revenue arrives gradually as the capacity is deployed and used. The provider must therefore finance the infrastructure before it knows the ultimate level, duration and profitability of demand. This reverses part of the advantage associated with traditional cloud growth, where strong operating cash flow and a more predictable capacity cycle allowed expansion to remain largely self-funded.

The distinction is also visible across the main firms like Amazon and Microsoft, which can recover capex directly through AWS and Azure consumption, while also using the same capacity for their own products. Google

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

combines external cloud demand with internal requirements from Search, YouTube and Gemini. Oracle has become a larger infrastructure provider through major AI contracts, but its smaller cash-flow base makes the associated investment more demanding relative to the group. Meta has no comparable public-cloud revenue stream, although Meta One and the new Meta Enterprise Platform provide more direct ways to monetise AI. For now, however, returns on its infrastructure still depend mainly on improvements in advertising and engagement. AI therefore challenges one of Big Tech's main financial strengths: its ability to fund growth almost entirely through internally generated cash. As infrastructure spending rises before the associated revenue is realised, maintaining this self-funded model becomes more difficult.

From Asset-Light to Asset-Heavy

Hyperscalers historically combined substantial infrastructure investment with the asset-light economics of their broader technology businesses, as revenue and cash flow grew much faster than the capital required to support them. AI is shifting that balance, with growth now requiring much greater upfront investment in computing capacity. FactSet estimates that aggregate capex across Alphabet, Amazon, Meta, Microsoft and Oracle increased from approximately \$95bn in FY2020 to around \$490bn over the last twelve months. It expects spending to exceed \$690bn across the companies' respective FY2026 periods. S&P Global Ratings uses a calendar-year estimate of approximately \$750bn for the same five companies, equivalent to 38% of their projected revenue. Despite differences in methodology, both estimates illustrate how infrastructure investment is consuming a much larger share of hyperscaler revenue and cash generation than it did before the AI cycle.

S&P Global Market Intelligence estimates that the five hyperscalers have collectively spent \$1.1tn on capex over the past five years, with Visible Alpha consensus implying a further \$5.3tn through 2030. The composition of that spending has also changed. FactSet estimates that processors and servers will account for approximately \$380bn in 2026, close to 60% of aggregate hyperscaler capex, compared with around 43% in 2022. The increase reflects larger volumes of AI accelerators and higher costs from data centre construction and core infrastructure. While buildings and power infrastructure may support operations for more than a decade, servers and accelerators are depreciated over much shorter periods and require more frequent replacement. As compute becomes a larger share of total expenditure, a larger proportion of capex becomes recurring as hardware becomes outdated. Capex growth may therefore moderate without absolute spending returning to the levels of the earlier cloud cycle.

Microsoft's disclosures show how quickly the mix has moved. In the fourth quarter of FY2026, the company recorded \$41bn of capex, with roughly two-thirds directed towards short-lived assets, principally CPUs and GPUs. Microsoft Cloud still generated a gross margin of 65%, while the growing infrastructure base increasingly feeds through to costs via depreciation and operating expenses.

The impact of capex on accounting numbers unfolds slowly. Capital expenditure creates an immediate cash outflow, while its effect on the income statement is spread over the asset's useful life through depreciation. Reported earnings may therefore remain strong even as free cash flow (FCF) comes under pressure during the early years of a build-out. This creates a second-stage pressure on operating margins after the initial cash requirement has already been absorbed. Alphabet has repeatedly identified higher depreciation from technical infrastructure as a cost headwind, even while Google Cloud's operating margin improved through higher revenue and efficiency.

When Cash Flow Is No Longer Enough

The financing gap emerges as infrastructure expenditure grows faster than operating cash flow, while hyperscalers simultaneously seek to preserve capital flexibility. Since FCF broadly represents operating cash flow after capital

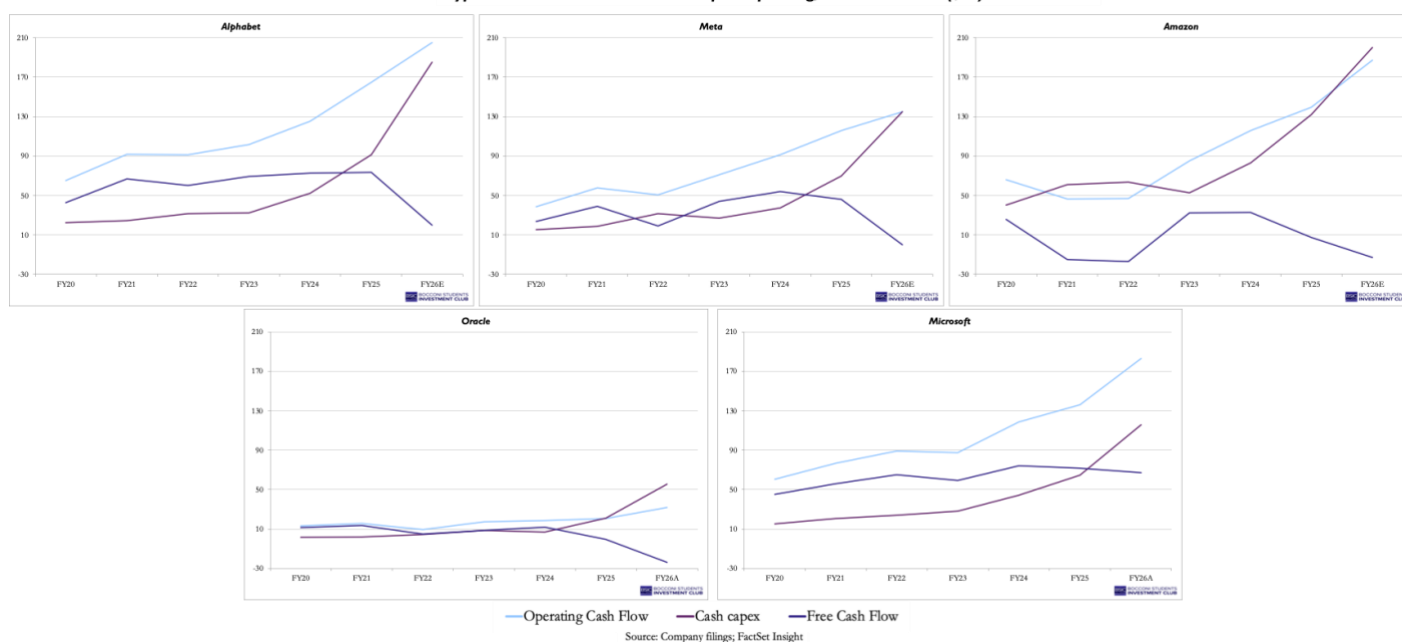
All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

expenditure, rising capex can sharply compress FCF even while revenue, operating income and reported earnings continue to grow. Amazon provides a direct example. For the TTM ended March 2026 operating cash flow rose by 30% to \$148.5bn. Purchases of property and equipment increased to \$147.3bn, driven by the company's rapidly expanding investment programme, including AI and AWS infrastructure. FCF consequently fell to \$1.2bn from \$25.9bn one year earlier. The company was producing more operating cash but almost all of it was being reinvested.

Even though Microsoft is still free-cash-flow positive, it is under the same pressure to invest more as revenue grows. For Q4 of FY2026, Microsoft generated \$55.4bn in operating cash flow and spent \$35.8bn in cash for PP&E, leaving \$19.6bn of FCF. Its reported capex was higher at \$41bn because it also included finance leases. Microsoft can still fund its investment internally, but capex is leaving less capital available for acquisitions and shareholder distributions.

Oracle shows the more stretched version of the same model. It generated \$32bn of operating cash flow in FY2026 but reported negative FCF of \$23.7bn as spending on Oracle Cloud Infrastructure expanded. The company raised \$43bn in debt and \$5bn in equity during the year, with management planning a further \$40bn of debt and equity financing for FY2027. Its rapidly growing contracted backlog provides revenue visibility, but the infrastructure required to fulfil those contracts must often be deployed well before the associated revenue is recognised. Oracle therefore faces the same timing mismatch as the larger hyperscalers, but with less internally generated cash flow and greater starting leverage.

Hyperscaler Cash Generation and Capital Spending, FY2020 - FY2026 (\$bn)



External financing is already becoming a larger part of the aggregate funding mix. FactSet estimates that incremental debt as a share of capex increased from approximately 9% in FY2024 to 32% over the twelve months to mid-2026. Aggregate debt across the five companies reached around \$700bn. This is a material change from a model in which new infrastructure could be funded almost entirely through operating cash flow.

Yet, the financing pressure remains uneven. Microsoft, Alphabet, Amazon and Meta entered the cycle with sufficient balance-sheet capacity to manage the current programme, while Oracle's capex and contractual commitments are significantly larger relative to its cash generation. For the largest firms it is primarily a capital-

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

allocation question, while for Oracle it has become more directly connected to credit capacity. Customer prepayments and customer-supplied hardware can reduce the upfront funding requirement; Oracle, for example, reported \$75bn of such arrangements across its large AI contracts. However, these arrangements are project-specific and do not eliminate the broader financing gap. Internally generated cash remains the main source of funding, but it can no longer cover the full infrastructure programme without limiting other uses of capital. Hyperscalers are therefore increasingly turning to corporate debt, leases, and external capital structures to bridge the gap.

Financing the AI Arms Race

As hyperscaler capex absorbs a growing share of internally generated cash, the question becomes who will supply the marginal dollar while companies preserve flexibility for other uses of capital. The answer includes a layered funding structure that moves costs and risk away from companies' balance sheets or at least delays when it appears there. The need for funding is met through a mix of corporate borrowing, leases and private investments: some put debt directly onto a balance sheet and others place debt in separate entities or delay when lease obligations are recognised.

There are three main ways to hold a data centre. A hyperscaler can own it directly, recording the asset on its balance sheet and funding it through cash or corporate borrowing; it can lease it, spreading payments over years and recording the lease liability once the facility becomes available to use; or a joint venture (JV) or special-purpose vehicle (SPV) can own it and borrow to finance it, while the hyperscaler carries a lease or buys its capacity. The most direct source of funding is corporate debt. Hyperscalers issued more than \$100bn of debt in 2025, with maturities longer than five years, locking in funding for multi-year build-outs. For financially strong companies this is a natural first move, as long-term borrowing matches the life of assets, debt costs less than equity and it preserves buybacks and dividends. The market is nonetheless becoming more concerned about credit risk. Credit default swap (CDS) spreads measure the cost of protection against a borrower defaulting on its obligations, and these spreads have risen, specifically for hyperscalers with lower credit ratings, reflecting increased bond supply and uncertainty around returns on AI projects.

The differences between some of these companies are substantial: Oracle's CDS spread rose from below 50 to 227 basis points, whilst Microsoft, Amazon and Meta, as a comparison, maintained significantly tighter spreads alongside their stronger credit ratings. The message, consistent with the earlier financing-gap discussion, is that credit risk is real but uneven, and it reflects a change in funding structure rather than a sign of distress for the industry.

Leasing is another major source of funding, with hyperscalers securing data-centre capacity without paying for construction up front. Meta reported about \$103.8bn of commitments under leases that had not yet started at the end of 2025. These mainly covered data centres, colocations and network infrastructure, with terms running from one year to thirty years. Across Microsoft, Meta, Oracle, Amazon and Alphabet, payments under leases that have not yet started reached \$1.09tn, four times the roughly \$285bn of lease liabilities already recorded on their balance sheets. These future obligations are disclosed in financial statement notes and are not hidden, and are undiscounted future payments, whereas recorded lease liabilities reflect present values, so the two amounts cannot simply be added together or treated as equivalent measures of debt. Nevertheless, these contracts commit companies to substantial fixed payments before the liabilities appear on their balance sheets.

The third channel, private capital, is likely to grow as funding needs grow. Goldman Sachs expects hyperscalers' spending on AI and data centres to reach \$5.3tn by 2030 and warns that public debt markets may reach limits on

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

how much lending they can absorb. In addition, Goldman points to infrastructure funds as a large pool of money. As of September 2025, they held more than \$1.7tn in assets and approximately \$400bn of dry powder, having raised a record \$221bn in 2025. AI infrastructure suits this capital because its long duration matches the investment horizons of pension funds and insurers seeking long-term assets. Leases with financially strong technology companies offer a steady income and some may even offer protection against inflation. When deployed through project-level vehicles, private capital can fund specific projects without placing the associated debt directly on the parent company's consolidated balance sheet. These funding sources can also be combined through a JV or SPV. The Bank for International Settlements (BIS) describes arrangements in which a separate entity develops or buys data centres: sponsors contribute equity and the entity borrows from private lenders or institutional investors. The hyperscaler owns a minority stake and signs a long-term lease or agreement to buy capacity. This substitutes upfront capex with multi-year operating expenses, while leaving most project debt at the vehicle level. The hyperscaler's contractual payments then provide the cash used to service that debt. The BIS calls these arrangements "shadow-borrowing" as they resemble debt economically but largely sit outside corporate balance sheets, creating closer financial links between tech companies, private credit funds, insurers and banks.

Hyperion: Inside a \$30bn AI Financing

A blueprint for how hyperscalers can finance increasingly capital-intensive AI infrastructure without relying exclusively on their balance sheets emerged in October 2025, when Meta partnered with Blue Owl Capital [NYSE: OWL] to fund the development of Hyperion, its 2GW, 4m-square-foot data centre campus in Richland Parish, Louisiana. Rather than entirely financing the project itself, Meta retained a 20 percent interest in the joint venture (JV), while Blue Owl-managed funds acquired 80 percent. Morgan Stanley [NYSE: MS] arranged more than \$27bn of debt and about \$2.5bn of equity through a special-purpose vehicle (SPV) to finance the build-out, with the financing entity, rather than Meta itself, raising the debt.

Yet Meta remains at the centre of Hyperion's economics: it will provide construction management and property management services for the project, and the company is set to lease all the facilities of the campus once construction is complete. More importantly, Meta also provided the joint venture with a residual-value guarantee for the first 16 years of operations, where the social media firm would make a cash payment to the JV upon non-renewal or termination of a lease. The structure therefore allows Meta to mobilise tens of billions of dollars of external capital without directly borrowing the equivalent amount itself, while providing investors with contractual protection linked to Meta's creditworthiness.

This creates the central tension behind Hyperion: moving debt outside Meta's balance sheet does not necessarily move the underlying economic exposure with it. This close economic link is reflected in S&P's A+ rating of the project bonds, only one notch below Meta's AA- corporate rating, despite the debt being formally issued by a separate vehicle. Hyperion therefore highlights a broader feature in AI infrastructure financing: while legal ownership, external equity and project-level borrowing can keep substantial debt away from the hyperscaler's consolidated balance sheet, long-term leases and guarantees may leave much of the project's economic risk tied to the hyperscaler itself.

Who Ultimately Bears the Risk?

The mix of funding sources explains where the money comes from. The more important question is what happens if the assets underneath are worth less than expected, since some of the structures reviewed above shift part of that exposure away from the hyperscaler's consolidated balance sheet. To understand who bears the risk, we need to look at how long the assets last, which guarantees the hyperscaler has provided and how the lenders are connected.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

The core problem is that debt and leases run for many years, with bond maturities extending beyond five years after issuance, and Meta's leases that have not yet started run from one year to thirty years. GPUs and servers can become outdated much sooner, meaning that the technology housed within a financed facility may require several replacement cycles over the life of the underlying lease or debt. This matters because data centres contain different types of assets. Building and power infrastructure can remain valuable for decades, making long-term financing a good fit. Chips and servers, however, face much faster replacement cycles, and so the risk is higher when long-term financing supports assets that need replacing.

A problem that arises is that guarantees can bring asset-value risk back to the hyperscaler. Some leases include a residual-value guarantee, meaning that if an asset is worth less than an agreed amount at a specified date, the hyperscaler may be required to cover some or all of the shortfall under the terms of the contract. Meta, as an example, disclosed data centre leases starting in 2029 with an initial commitment of \$12.3bn, alongside a residual-value guarantee with an aggregate threshold of \$28bn. This threshold is not an expected payout, as Meta classified the guarantee to be "not probable" and recorded no liability for it.

These structures can also create additional channels through which financial stress is transmitted. The BIS notes that they connect hyperscalers with private credit funds and banks, which may provide credit lines to the financing vehicles. For example, weaker demand for data centre contracts could reduce a vehicle's expected income and make refinancing harder, potentially generating losses for the funds, insurers or other investors holding that debt. AI infrastructure is increasingly financed across private credit, asset-backed finance, CMBS, ABS and corporate debt, making aggregate exposure harder to trace and compare. Many private loans and structured investments trade infrequently, meaning their values may adjust more slowly than prices in the public markets, so investors may underestimate declining asset values or the concentration of their exposure. These concerns do not reflect a crisis: most hyperscalers have investment-grade ratings, several are rated in the AA/AAA range; contracted lease payments support a multitude of projects; and rating agencies already factor some commitments into adjusted-debt measures. Ultimately, who bears the cost depends on the contractual structure. In project vehicles, equity investors absorb first losses; private-credit funds and insurers bear credit exposure; while hyperscalers can retain economic risk through binding leases and guarantees.

The Return on AI Investment

Everything covered so far only matters if the returns on spending are justified by the profit AI can generate. Evidently, investment and returns operate on different timelines: capex is being spent now, while the revenue that is supposed to justify it follows with a lag. Any assessment of returns today therefore remains necessarily incomplete. Companies make money from AI in different ways: AWS, Microsoft Azure, Google Cloud and Oracle sell computing capacity and related services, allowing part of their investment to be monetised directly through cloud revenue. Microsoft and Google sell AI subscriptions and assistant products. Meta mainly uses AI to improve recommendations, engagement and advertising, so the majority of its returns come from advertising performance rather than a separate AI stream.

Revenue alone cannot tell us whether these investments are worth it: what matters more is whether the additional profit generated by AI investments is enough to compensate for the cost of funding it. In financial terms, incremental investment creates economic value when the return on the additional capital deployed exceeds the company's weighted average cost of capital (WACC). If the return on the next dollar of AI investment falls below WACC, additional growth can destroy value even if revenues continue to increase. Therefore, the question is not if AI revenue exists, but if the marginal dollar spent is earning an adequate return.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

This comparison cannot be made with precision by outside investors. AI spending and revenues are generally embedded within broader cloud, advertising and product businesses rather than separately disclosed:

- Cloud revenue growth indicates demand, but not how profitable that demand is
- Cloud margins provide evidence of profitability but combine AI with other cloud services
- Comparing capital spending growth with revenue shows whether investment is running ahead of current revenues, although today's spending may support future growth
- Free cash flow captures the cash remaining after investment, but does not isolate returns from AI
- Company-wide ROIC measures overall capital efficiency, but does not distinguish between AI and other investments

Taken together, these measures provide useful evidence, but they cannot yet establish conclusively whether incremental AI investment is earning above its cost of capital. The investment cycle can therefore evolve along three broad paths. In the strongest case, monetisation catches up with spending, utilisation remains high and today's capex generates returns comfortably above the cost of capital. In a middle case, investments earn roughly the cost of capital: the business grows but creates little additional economic value. In the weakest case, capacity grows faster than demand, utilisation disappoints and returns fall below the cost of capital.

Conclusion

The AI race is ultimately becoming as much a capital-allocation challenge as a technological one. Hyperscalers are devoting unprecedented amounts of capital even before the full scale and profitability of the future demand are known, increasingly supplementing their internal cash generation with debt, leases and private capital. These structures can expand financing capacity and distribute risk; however, they cannot make that risk completely disappear: leases, guarantees and contractual commitments often leave hyperscalers economically tied to the infrastructure they finance. The key question is therefore no longer simply how much capacity can be built. It becomes whether the profits generated by artificial intelligence can justify the capital required to build it. Ultimately, the winners of the AI race may not be those that spend the most, but those that turn this unprecedented investment into durable returns.

TAGS: AI, Hyperscaler, Capex, Financing, Google, Meta, Microsoft, Oracle, Amazon

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.