Goldman Sachs



# Building the AI factories of tomorrow

Hyperscalers and data center developers are fueling an unprecedented infrastructure buildout at a global scale. As AI continues advancing at breakneck speed, the efficient sourcing, deploying, and recycling of capital is vital.

Jason Tofsky and Alexander Tingle, who sit in the Technology, Media, and Telecom group within Investment Banking, discuss how capital markets are innovating to help build the AI factories of tomorrow.

# 1 Investor enthusiasm and record CapEx have spurred comparisons between today's AI boom and the internet bubble. Is this fair?

People love making this comparison, but there's a critical difference: Immensely profitable, free-cash-flowing hyperscalers are driving this AI boom, and they're willing to spend this capital because they believe it's important for their businesses. This setup is in stark contrast to the internet bubble, where many of the high-flying tech stocks were heavily leveraged or generating far less revenue and free cash flow than market leaders are today.

**S&P 500 Forward P/E Multiple**




Source: Compustat, FactSet as of 10/1/2025.

**Hyperscaler corporate returns have remained outsized vs. the broader economy for the past decade**

Cash return on capital invested (CROCI) for eight hyperscalers vs. all sector weighted average (2012–2027E)




Source: Goldman Sachs Research, FactSet. Hyperscalers include Google, Microsoft, Meta, Amazon, Oracle, Baidu, Tencent, Alibaba.

It is generally much simpler to assess the return on investment for CapEx, and one must assume that tech companies' capital allocation strategy is more than the "fear of missing out." It's clearly because of the acceleration in growth. With most CapEx being spent either by hyperscalers directly or via their developers, there's really very little speculative capital being spent on building data centers right now. Importantly, this is all unfolding alongside accelerating growth in the cloud market, with public cloud consumption jumping from ~$90bn in 2019 to $335bn in 2024.*

*Source: "Riding the hyperscaler wave." McKinsey. September 2025.

## 2 AI is evolving at an unprecedented speed, but this pace of change poses new questions and obstacles. How are hyperscalers approaching data center development, chip financing, and power supply?

At the beginning of the cloud era, hyperscalers had largely taken an "insourced" approach to digital infrastructure—finding their own power or land and then either developing the data center themselves or contracting with a third-party developer. Today, hyperscalers are increasingly outsourcing parts of the data center development life cycle, and we have a slightly out-of-consensus view this trend will persist and grow—especially for data centers, chips, and compute.

As a result, third parties (data center companies, neoclouds, etc.) now play an influential role for their ability to source different components and anticipate demand far in advance. Many of today's large-scale data center projects are going behind-the-meter with natural gas generators that require an entirely different set of infrastructure—and just securing the gas turbines takes multiple years. By procuring the gas, turbines, permitting, etc., in advance, however, these third parties suddenly have a product that hyperscalers can buy or contract with directly.

While hyperscalers could theoretically do all of this themselves, outsourcing to third parties can unlock a massive speed-to-market advantage that outweighs the impact on margins. That said, costs for large-scale turnkey data centers are now higher than the $10mn–$12mn per MW seen in recent projects because hyperscalers are asking data center companies to do more on their behalf—increasing the value proposition for digital infrastructure providers in the data center, fiber, and compute markets.

## 3 There seem to be very divergent views on the role that Tier 2 markets will play. How do you see this unfolding?

Of course, there is ongoing debate around residual value for non–Tier 1 data centers. Bulls argue that building a $10bn data center in a Tier 2 market will require spending $20bn–$30bn every five years*—are you just letting that go at the end of a 15-year lease? Probably not. If we experience another DeepSeek moment, however, those Tier 2 locations could be less interesting. That said, the cloud market used to only really have availability in Tier 1 spots like Northern Virginia and Silicon Valley, but leading cloud providers now provide "availability zones" across very diverse geographies globally.

People tend to lump Tier 2 markets into an umbrella category, but there's quite a difference in quasi-remote data centers 50 miles from a data center hub and ones located in the mountains. In fact, hyperscalers are now building a critical mass of compute infrastructure and clustering multiple data centers in markets considered "remote" up until recently.

For more on this topic, look no further than Abilene, Texas, which many investors and developers considered too remote a few years ago. Flash forward to today, those same people would rather put money in an Abilene project than in other locations because they have conviction in how much development activity and clustering is already underway.

There's no crystal ball for which nontraditional markets will become the next Tier 1 hub, but many will make that transition by virtue of clustering in places like Abilene. These views will certainly evolve as we gain greater appreciation for which workloads are truly latency-sensitive, but we are clearly past the era of Silicon Valley and Northern Virginia representing the only markets that matter. Just as Dallas, Phoenix, and Atlanta have become de facto Tier 1 markets in their own right, we're now seeing the next generation of Tier 1 hubs emerge.

*Source: SemiAnalysis, August 2025. Supermicro Data Sheet, February 2025.

"We are clearly past the era of Silicon Valley and Northern Virginia representing the only markets that matter. Just as Dallas, Phoenix, and Atlanta have become de facto Tier 1 markets in their own right, we're now seeing the next generation of Tier 1 hubs emerge."



**Alexander Tingle**
Managing Director, Technology, Media, and Telecom, Investment Banking
Goldman Sachs

# 4 We have recently seen a convergence in asset manager and digital infrastructure M&A. What's driving this activity?

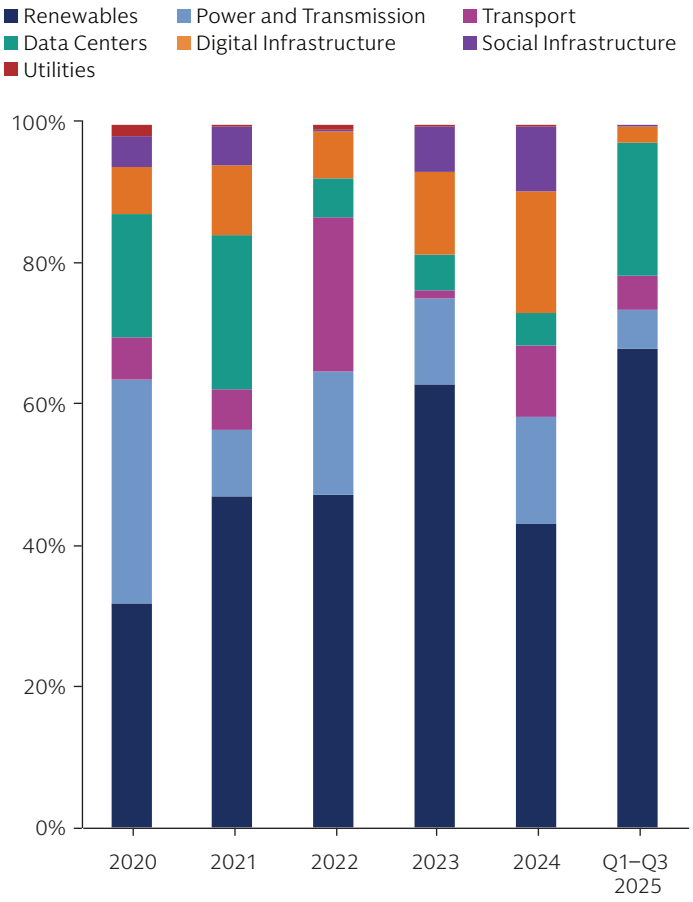
AI data centers are just the next generation of digital infrastructure, supported by an ecosystem of companies originating data center leases, powered land acquisitions, and compute (e.g., GPUs and chips). What's ***new*** are the specialist companies that fit all of those pieces together for customers—acting more like asset managers than platforms as they scale.

This trend is attracting large asset managers that have traditional infrastructure and/or private credit funds but lack digital infrastructure capabilities as they build out their investment and capital formation teams in this space. For these firms, data center operators and development platforms can provide a proprietary origination engine to generate recurring, high-quality investment opportunities for their funds, balance sheet, and other mandates.

Earlier this year, we advised Stream Data Centers on a majority investment from Apollo that will help Stream execute on a multi-gigawatt development pipeline and provide Apollo with an "origination engine" for deploying capital. In May, we similarly advised alternatives manager TPG on its acquisition of specialized digital infrastructure investor Peppertree Capital—providing TPG with immediate scale in wireless communications tower and network infrastructure development.

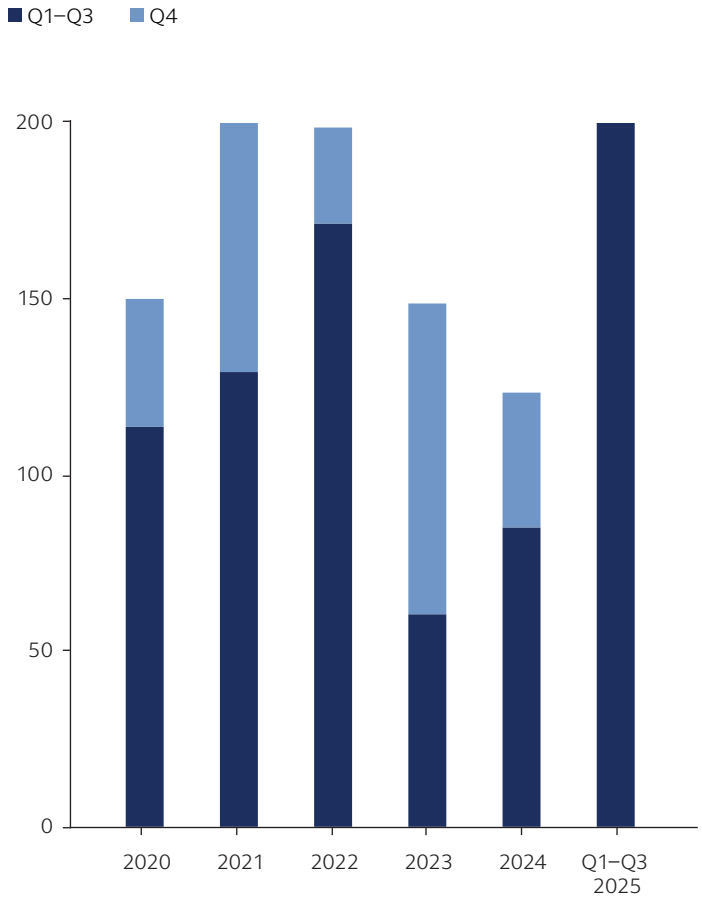
For firms that already have digital infrastructure offerings, they are strengthening their capital formation capabilities in retail and insurance channels—not only for the cost of capital, but to access longer-term financing. The 7–10-year life cycle of private closed-end funds with LPs wanting liquidity is misaligned with digital infrastructure timelines. This duration mismatch has pushed firms toward more evergreen capital with longer time horizons that better suit this type of development—allowing investors to own A+ assets for longer and continue to compound returns.

**Proportion of Capital Raised by Sector-Specific Funds**




Source: Infrastructure Investor Q3 Fundraising Report.

**Private Infrastructure Capital Raised ($bn)**



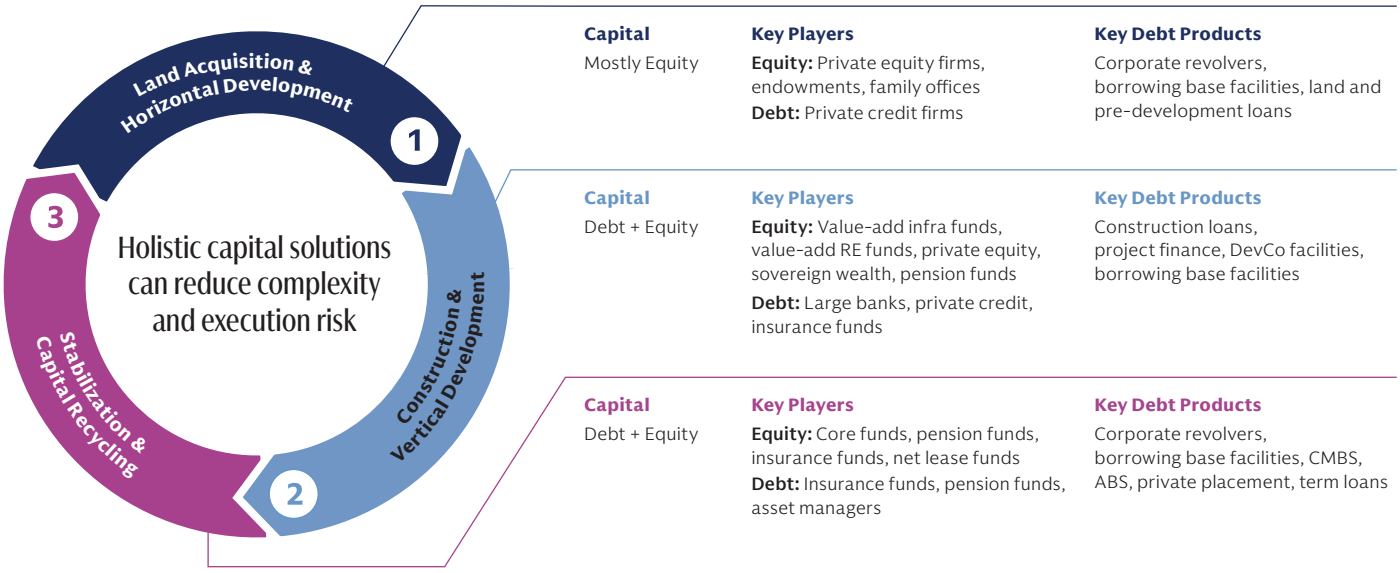

Source: Infrastructure Investor Q3 Fundraising Report.

## 5 Can you dive a little deeper on this concept of "origination engines"?

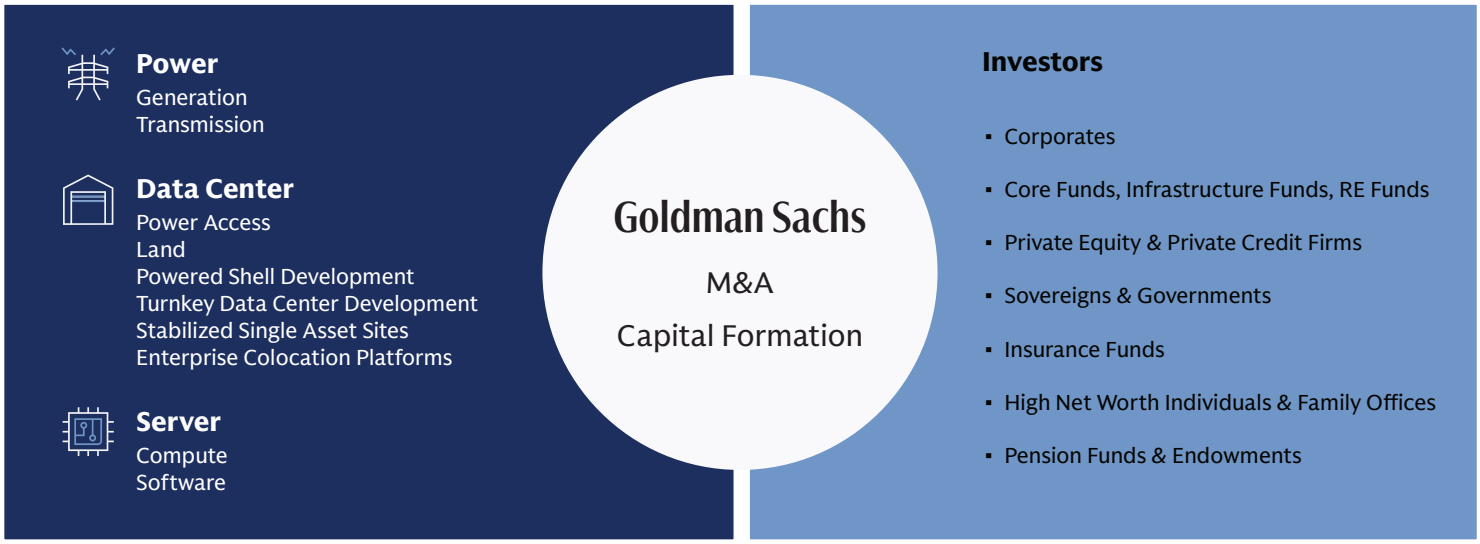
Traditionally, infrastructure and real estate funds would acquire data center platforms by buying a company in its entirety and funding development through an existing fund. However, platforms only focusing on one piece of the AI infrastructure value chain (land, power development, vertical data center development, stabilized data centers) narrow the universe of capital participants to a specific risk and cost profile. For example, land acquisition financing is generally too risky for insurance capital, stabilized data center yields are too low for private credit funds, and so on.

With room for all types of capital to engage in AI infrastructure, developing or acquiring the right data center platform maximizes the opportunities for deploying capital across that value chain. By either acquiring a data center platform in an asset-light fashion or developing an integrated solution internally, an "origination engine" uses said platforms as a conduit for raising and deploying different types of capital—insurance, retail, private credit funds, firm balance sheet, etc.—against infrastructure projects that suit their respective cost and risk profiles.

**AI infrastructure is currently funded in parts by disparate investors, but more comprehensive solutions can accelerate growth.**




**Building bespoke capital solutions can open the door for underinvested pools of capital to participate across the AI value chain.**



## 6 "AI factories" are growing in scale, with multiple +1 GW projects announced in 2025. How are developers recycling capital to meet demand?

This question is becoming more important every day. Just in the last few months, we saw the largest data center construction loan for syndication announced to date for Vantage's new 1.5 GW campus in Texas and Wisconsin. At this scale, everything gets more complex—pushing lenders and developers to be more efficient at every stage.

In the typical development life cycle, a large data center developer in need of capital goes to a lender and says, "I've got this contract; can you help me finance it?" Lenders will be comforted by the tenant contract already in place, but this initial stage still carries the highest cost of capital due to the heightened risk and uncertainty. Later, when that data center is fully leased and cash-flowing, or de-risked, the developer can sell it in the stabilized infrastructure market via ABS/CMBS/private placement and recycle that capital into new projects. The last three years have seen $21.6bn in data center ABS/CMBS volumes, and another $19.2bn YTD as activity accelerates.*

But the middle stage is getting especially interesting as developers use creative financing structures to lock in both construction ***and*** long-term financing up front. Earlier this year, Crusoe, the developer for Stargate's first data center, used a fully funded forward takeout agreement with Blue Owl Capital and Primary Digital Infrastructure to fund construction and secure permanent financing once it's built.

Tailoring these joint ventures (JVs) to the needs of different capital sources can drastically expand the universe of participants. Through longer contracts mitigating residual risk, removing early termination rights, and ensuring triple-net structures, these forward sales would become viable for evergreen sources of capital like insurance and retirement plans.

We believe that AI infrastructure will lead the way for capital formation over the coming decade writ large.

*Source: Green Street, Finsight as of September 30, 2025.

## 7 The compute equipment in AI data centers costs 3–4x more than the physical structure. How are AI labs and neoclouds financing chip purchases?

Intensifying competition has led chip suppliers and neoclouds to adopt creative strategies—such as credit support and other incentives—to accelerate market adoption, fundamentally reshaping how end-customers finance purchases.

The largest players, such as the hyperscalers and platform companies, are using balance-sheet support, credit enhancements, and flexible contracts to drive adoption. If a neocloud wants to buy even up to $20bn GPUs, strategic partners are able to provide partially guaranteed payments and structured offtake options, or "credit wrapping"—re-renting, remarketing, or repurchasing GPUs—to improve the neocloud's credit profile, which helps reduce financing costs and GPU obsolescence risk. Critically, credit wrapping with an investment-grade (IG)–rated anchor also expands the universe of participants to include insurance capital, syndicated credit, private credit, and more.

But for those without credit support from an IG anchor (like many AI labs), financing is mostly coming from the private credit market at a higher cost of capital—but innovation is coming.

We are trying to build financing solutions that blend IG cash flows from IG projects with non-IG contracts, which enables IG credit support for non-IG customers and then allows for overall financing size to increase for a wider mix of customers. This is very similar to the existing borrowing base facility technology in data centers and other areas of digital infrastructure.

"This next era of digital infrastructure will involve rewiring how capital flows into the AI ecosystem. Our work centers on collaborating with first-of-their-kind platforms that link bespoke M&A and capital formation solutions across the risk curve, unlocking underinvested pools of capital—including insurance and private wealth—that will define the next decade of asset management growth."



**Jason Tofsky**
Partner, Technology, Media, and Telecom, Investment Banking
Goldman Sachs

## 8 Looking outside the US, what dynamics are at play in emerging global data center markets?

Even a year ago, there was much discussion around building data centers across the Middle East, South America, and elsewhere. These regions will certainly play a critical role for data center capacity at some point, but these markets also have their own local workloads and data center demand that need to be built first.

In a shifting US political climate and heightened focus on sourcing as much capacity as possible, hyperscalers have broadened the aperture to other fuel types as the US grid continues finding ways to bring new power online—despite all the doomsday talk. Long-term, however, hyperscalers still have net-zero emissions goals and want to do more with clean energy—making countries with strong renewable power bases like Brazil attractive markets.

For now, whether or not these countries are used as overflow markets will depend on how quickly new power keeps coming online in the US and how friendly these international administrations are to enabling workloads that go abroad—and that remains to be seen.

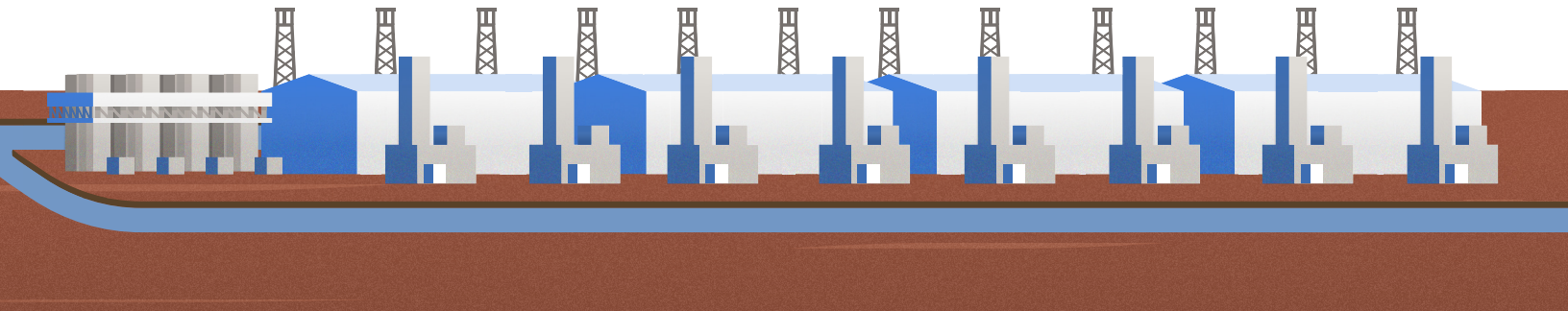
## 9 From a financing perspective, how are providers of capital evolving and finding efficiencies as more data centers get built?

If you look back at almost any transformative technology in history, the complexities and scale of each innovation required new financing methods to address those nuances. Investment banks, corporate bond markets, holding company structures, investment trusts, and a catalogue of other financial innovations all emerged to meet the unprecedented capital demands of their era.

Today is no different. The AI ecosystem has been financed in parts thus far—with different providers and capital solutions for each segment—but the rapidly growing scale of digital infrastructure demands more comprehensive solutions that offer greater certainty in both the availability and pricing of capital.

Structuring these bespoke solutions is much easier when developers are operating under a single contract throughout the data center life cycle. Not only does this approach reduce friction costs, but it also provides capital cost certainty and greater visibility—making it much easier to tailor financing structures to the nuances and complexities of digital infrastructure.

For us, our Capital Solutions Group is integrating all the different pockets and channels of capital at Goldman Sachs to strategically source, deploy, and recycle capital across the digital infrastructure ecosystem.



## Disclaimer

This document has been prepared by Goldman Sachs Investment Banking and is not a product of Goldman Sachs Global Investment Research. This document should not be used as a basis for trading in the securities or loans of the companies named herein or for any other investment decision. This document does not constitute an offer to sell the securities or loans of the companies named herein or a solicitation of proxies or votes and should not be construed as consisting of investment advice.

© 2025 Goldman Sachs.