

# GRANTS'S

## INTEREST RATE OBSERVER®

Vol. 43, No. 2b

233 Broadway, New York, New York 10279 • www.grantspub.com

JANUARY 31, 2025

### Bubble, interrupted

China's DeepSeek\* caught Mr. Market napping on Monday morning. Last month, the start-up tutored an open-source large-language model (LLM, the "knowledge" on which AI devices rely) that performed similarly to Meta Platforms, Inc.'s flagship Llama 3.1. However—and this is what set the market to freaking—it used just 9% of the Nvidia Corp. H100 GPU-hours that Llama requires. Due to U.S. trade restrictions, the graphics processing units (GPUs) that DeepSeek deployed had just half of the inter-chip transfer rate of H100s. Last week's update displayed inference capabilities that beat OpenAI, Inc.'s month-old o1 model on some metrics. With its reliance on AI-related spending, the market is rightly concerned about DeepSeek's efficiencies. However, this week's disturbances are only the start. Boom and bust—and the ever-present risk of a surprise technological leapfrog—are the topics at hand.

We offer three propositions:

1. Growth in the supply of big buildings crammed with Nvidia's best chips will significantly outpace growth in demand for the same; it was on track to occur even before Monday's thunderclap—or Tuesday night's shocker that Alibaba Cloud has produced an AI breakthrough that may excel DeepSeek's.

2. Fast-moving advances in know-how and software are casting a shadow of obsolescence on data centers yet unbuilt.

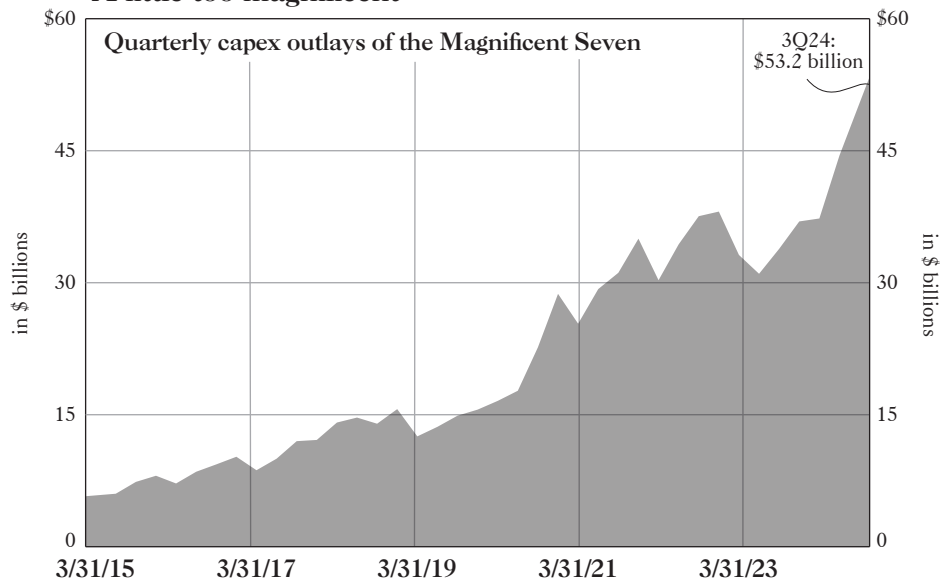
3. The looming multibillion-dollar mismatch, along with the implied future speculative crackup, has taken AI itself unawares. Thanks, AI.

We should have known that there was a banana peel in the path of the great AI capital-spending boom; the government was all in. One of President Joe Biden's final executive orders directs the departments of Defense and Energy to lease federal land for the construction of new data centers with expedited permitting. President Donald Trump, not to be undone, followed up a day after taking office with plans for the Stargate Project, a joint venture in which SoftBank Group Corp., OpenAI and Oracle Corp. are (or were) expected to spend between

\$100 billion and \$500 billion on AI infrastructure, including data centers. The Biden and Trump initiatives come on the heels of more than 100 gigawatts' worth of new data-center projects announced since the beginning of 2023 (think all 94 American nuclear reactors combined). We judge that the coming bust will not only take down data-center companies, but also reverberate across the data-center supply chain and throughout the credit markets. Readers of a certain age will soon recall the millennium fiber-optic and telecom bust. Today, as in 1998–2002, technology businesses are unusually susceptible to overexcitement, overordering and overbuilding.

Current U.S. data-center capacity

A little too magnificent



source: The Bloomberg

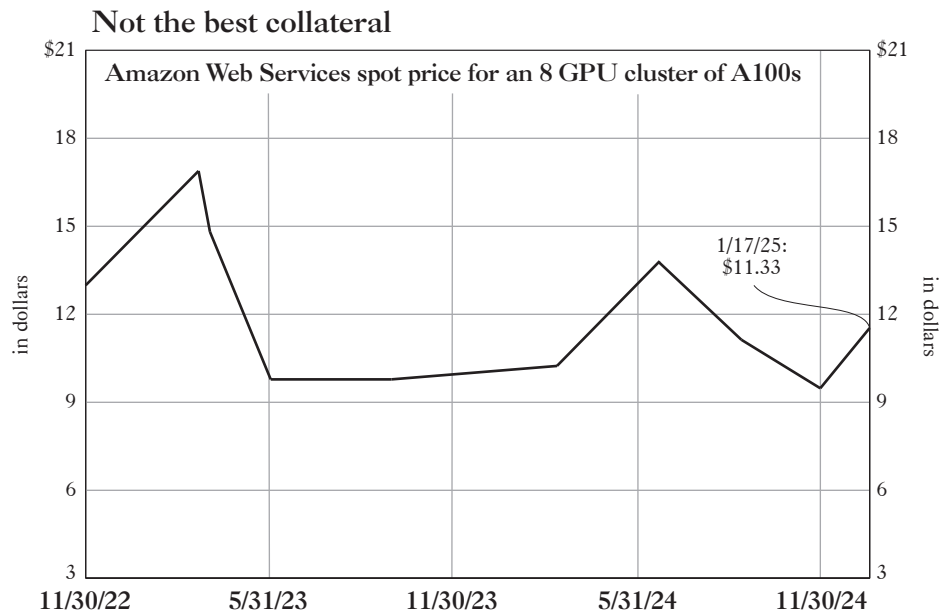
\* Or, call it by its full name, Hangzhou DeepSeek Artificial Intelligence Basic Technology Research Co. Ltd.

(in electrical power terms) stands at 35.5GW, up by 56% since 2023. New-project announcements imply a yearly doubling of capacity, or more, far ahead of the projected growth in demand. Chris Seiple, vice chairman of energy transition and power and renewables at Wood Mackenzie Ltd., tells *Grant's* that 49GW's worth of projects were launched in the latter half of 2024 alone. WoodMac suspects the 100GW announced in 2023–24 is “only a sample of total project development,” and at an average of \$11 million per megawatt, it will cost \$1.1 trillion. Data-center construction has historically taken 1 to 1.5 years, which means U.S. capacity could double before the year is out—and could rise nearly five times in the next two years unless there are delays in the procurement of power-transmission infrastructure.

Even a fourfold rise in current capacity in the next three years is likely to spell trouble. Assume a 27% compound annual growth rate for data-center demand, the top end of the Department of Energy's estimate through 2028 (and just above a recent calculation from Barclays). Demand would then double in three years, but supply would quadruple. Furthermore, project announcements in the second half of last year came in at 8.17GW per month, an astounding pace that smacks of overordering as well as overbuilding. If orders are canceled and projects are abandoned, the bubble may burst before the end of a full construction cycle. It's not out of the realm of possibility that it burst on Monday, Jan. 27.

It's the insatiable demand for computing power from the big data-guzzlers Amazon.com, Inc., Microsoft Corp., Alphabet, Inc., Meta and Apple, Inc. that instigated the data-center build-out. Taking management guidance where it exists and analyst estimates where it doesn't, there could be a collective capex increase from these giants in 2025 of 28.9%, to \$297.2 billion. Such a figure squares with the DOE's and Barclay's estimates of data-center demand growth. But the rate of new data-center projects more than extrapolates the 2023–24 growth in Big Tech capex, which was just shy of 50%.

Notably, even in the context of rising capex, Amazon laid out less year on year for property and equipment it purchased through finance leases in the 12 months to last quarter; one might



source: instances.vantage.sh

infer that the Everything Store is shifting its capex subtly from leasing data centers to building its own. This would jibe with what D.A. Davidson's senior software analyst Gil Luria recently told *Barron's*: namely, that Amazon's increasing operating margins were partly due to spending more on its own chips for its own data centers (and that costs less than buying from Nvidia).

As DeepSeek has already demonstrated, there will be further savings with improved efficiencies in training large-language models and with the open-source release of better digital mousetraps. In another development, Prime Intellect, Inc. released an open-source training method that requires less than a tenth of the usual bandwidth, allowing it to use smaller, geographically dispersed training centers. Reduced training demands of the DeepSeek variety, plus more distributed training, could obviate the need for capital-intensive hyperscale data centers and their engineering headaches (e.g., cooling, networking and power systems). For eager coders, there will be every incentive to make the additional supply unnecessary by the time it arrives. “The hyper-capex AI model will start getting more (justified) questions,” tweeted our friend, the short seller Jim Chanos. We agree.

The mind-boggling expenditures on computing resources have created a growing industry of data-center providers, who sell and lease their “com-

pute” to Big Tech. Many of these companies are former cryptocurrency miners, or mining-suppliers, which desperately needed another use for their GPUs when ethereum-mining became an industry of the past in 2022. These include Applied Digital Corp. (APLD on the Nasdaq), formerly Applied Blockchain; CoreWeave, Inc., formerly Atlantic Crypto Corp.; and Crusoe, Inc. ChatGPT arrived in 2022 just in time to save their bacon, and the ethereum-mining community pivoted right along with their main chip supplier, Nvidia. They have remained thick as thieves: Nvidia invests in all three companies.

...

As cloud-related capex expanded, data-center companies bought more Nvidia chips and facilities, often financing the expansion with loans secured against those same chips. Crusoe was one of the first to receive a GPU-backed loan, in the sum of \$200 million, in 2023. CoreWeave followed with \$2.3 billion, also in 2023, and then again with \$7.5 billion in 2024. These debt issues were led by Magnetar Capital, LLC and Blackstone, Inc. and were secured against CoreWeave's Nvidia H100 GPUs and their cash flows, both of which will decline in value with burgeoning data-center supply. CoreWeave says it took the funds with the goal of doubling its data-center capac-

### The field at a glance

	Share price	Market cap (in \$ bns)	EPS (TTM)	P/E (TTM)	P/B	Debt to equity
Applied Digital	\$6.52	\$1.45	-\$1.73	N/A	3.4	1.43
Equinix, Inc.	\$906.81	\$87.50	\$11.10	81.7	6.4	1.12
Digital Realty Trust, Inc.	\$164.14	\$54.85	\$1.26	128.7	2.5	0.78

sources: company reports, the Bloomberg

ity to 300MW. “Gone are the days of distributed computing,” its website boldly proclaims.

The *Financial Times* sounded the alarm in November over the emergence of this “new asset class” of debt facilities (about \$11 billion total) collateralized by Nvidia GPUs. Certainly, the GPUs are lousy collateral. Based on Applied Digital’s fiscal 2023 annual report, H100s were selling at an average price of \$40,039 each in the middle of 2023. They currently change hands for \$27,000–\$30,000, suggesting a 17% annual rate of depreciation. The chips probably do not depreciate at a steady rate but step down in value with subsequent technological developments. H200s are already on the market, and the world awaits the 2025 release of Nvidia’s data-center-grade Blackwell GPUs.

“You cannot lend against the chips,” Nathan Koppikar, cofounder and portfolio manager of Orso Partners, L.P., tells our U.K. correspondent, Joel Wallenberg. “It’s idiotic.” Orso Partners, which looks after \$386.9 million, is a doubly rare short seller of both AI and private equity and private credit. “You’re lending against Microsoft’s credit risk,” Koppikar continues. CoreWeave expects Microsoft to spend \$10 billion with them through 2030, and the special-purpose vehicle earmarks cash flows from the collateral GPUs to service the debt. The GPUs thus don’t necessarily need to hold their value, but the loans are good PR for Nvidia, Koppikar suggests, and good relations with Nvidia are a sine qua non for the would-be builder of data centers.

Wallenberg asked Joshua Stevens, associate director for Macquarie Group Ltd.’s commodities and global-markets business, about GPU depreciation with regard to Macquarie’s \$500 million loan to Lambda, Inc., another data-center company. That loan is se-

cured against H100 GPUs that are *not* contracted by an end-customer. Stevens indicated that such a step up in risk is now typical in this hot market: “Spreads have tightened significantly for transactions supported by off-takes to acceptable credits,” such as Microsoft. The GPUs truly have to function as collateral here, but Macquarie feels protected by the estimate of a seven-year useful life of a GPU, coupled with typically “3–4 year debt terms.” However, at the 17% rate of depreciation noted above, the GPUs would lose roughly another half of their value in 3.5 years. Nevertheless, Stevens focuses on the expected cash flows from the GPUs: “The operating cost to run a GPU... is less than 20% of the revenue it is likely to generate.”

One might well question whether the price of compute for H100 GPUs will hold indefinitely. The graph above shows a 32% tumble in the spot price of compute using A100s from its peak, as the newer H100s were being adopted.

Macquarie moved into data-center equity on January 14, with a \$900 million preferred and common-equity deal to fund Applied Digital’s new 400MW data center, including an option for a further \$4.1 billion (seemingly the entire cost, at \$12.5 million per megawatt). The preferreds, which are perpetual and optionally payable in kind, yield 12.75%. Macquarie mitigates its risk by requiring an executed lease for each megawatt of the new data center’s capacity in order to release an increment of \$2.25 million.

However, the stricter terms given with one hand are taken by another: A stated purpose of the deal is to allow APLD to repay outstanding bridge loans, \$150 million of which is a Macquarie 18-month 0.25% promissory note. The combination is thus something of a synthetic PIK-toggle arrangement for APLD. The reliance on payment-in-kind is not surprising, giv-

en that Applied Digital, whose debt the company itself estimates as CCC-rated, consistently runs an operating loss (\$99 million for fiscal 2024) and has high capex requirements. While Koppikar declined to tell us which shorts he currently has on, he did say “the worst [data-center provider] is Applied Digital.”

If some of the above sounds eerily familiar, you may be noticing a rhyme with the fiber-optics build-out of the late 1990s through 2002 (and its subsequent bust). The fiber/telecom bubble outlasted the dot-com bubble, as some argued that telecommunications infrastructure, with its intrinsic value, was safer than dot-com companies. Then as now, it’s not that the underlying technology wasn’t useful, but rather that the capital response to early high demand led to eventual excess capacity: By 2002, only 10% of long-haul fibers carried signal, and then just at a tenth of their capacity. Equity investors weren’t the only ones to lose: High-yield defaults climbed to 12.8%, more than 3 points of which was WorldCom’s \$30 billion default. The pre-leasing of planned data centers offers more protection than the fiber build-out did, but only where those deals have been cut.

Our bearish guide Koppikar continues: “The market is pricing in sustainability in many, many companies, and this is wrong.” As the data-center glut arrives, the price of compute will drop and Big Tech will pare back its capital expenditure. GPU depreciation will accelerate, meaning that neither cash flows nor hard collateral will make data-center creditors whole. “Once the capex drops off, everyone will be hit up and down the supply chain. Tech bubbles are so potent because they sell to each other.”

The point is easily illustrated. In its quarterly report, filed on Jan. 14, Applied’s top three customers contributed 51%, 28% and 16% of revenues (they were probably tech giants). APLD names Super Micro Computer, Inc. as a critical supplier of servers. Super Micro, in its latest quarterly report, disclosed that it relied on two customers for 38% of sales, three customers for 61.4% of accounts receivable and one supplier for 69.8% of its equipment outlays (surely Nvidia). As of last quarter, Nvidia’s top three customers accounted for 36% of revenues,

and its four largest customers for 53% of receivables.

With connections this tight and brittle, when trouble comes to AI land, it will probably come all at once.

Not that AI is losing sleep over the risks. GPT-4o, when queried, acknowledges that “some analysts” warn of a bubble (it knows us!), though it doesn’t “experience emotions or worry.” The

chatbot concludes that “the overall industry trend points toward a need for continued expansion.” Nowadays even bots talk their book.

•

*Grant's® and Grant's Interest Rate Observer® are registered trademarks of Grant's Financial Publishing, Inc.  
PLEASE do not post this on any website, forward it to anyone else, or make copies (print or electronic) for anyone else.  
Copyright ©2025 Grant's Financial Publishing Inc. All rights reserved.*