

Hyperscaler AI Capex 2026: Amazon, Microsoft, Google, Meta

Published July 18, 2026 40 min read



Amazon, Microsoft, Alphabet (Google's parent) and Meta Platforms have collectively guided toward roughly \$700 billion to \$750 billion in capital expenditure for calendar 2026, a figure that continues to be revised upward with nearly every quarterly earnings call. Amazon set the pace in February 2026 by guiding to **\$200 billion** in 2026 capex, more than 50 percent above the **\$131 billion** it spent in 2025 (Source: [cnbc.com](https://www.cnbc.com/2026/02/05/why-amazons-ceo-is-confident-with-200-billion-spending-plan.html#:~:text=spending%20roughly%20%24131%20billion%20on%20purchases%20of%20property%20and%20equipment%20in%202025%2C%20up%20from%20about%20%2483%20billion%20in%20the%20year%20pr)). Alphabet followed a day earlier with guidance of **\$175 billion to \$185 billion**, later raised twice, first to **\$180 billion to \$190 billion** at its first quarter 2026 earnings call in April and again reaffirmed in its June 2026 investor presentation (Source: [sec.gov](https://www.sec.gov/Archives/edgar/data/1652044/000165204426000012/googexhibit991q42025.htm#:~:text=our%202026%20CapEx%20investments%20are%20anticipated%20to%20be%20in%20the%20range%20of%20\$190%20billion%20to%20\$200%20billion%20in%202026%2C%20and%20anticipate%20strong%20long-term%20return%20on%20invested%20capital)). Amazon's Q4 2025 earnings call in October 2025 projected 2026 capex at **\$200 billion**, with a range of **\$180 billion to \$220 billion** (Source: [amazon.com](https://ir.aboutamazon.com/news-release/news-release-details/2026/Amazon-com-Announces-Fourth-Quarter-Results/)). Microsoft's Q4 2025 earnings call in February 2026 projected 2026 capex at **\$200 billion**, with a range of **\$180 billion to \$220 billion** (Source: [microsoft.com](https://www.microsoft.com/en-us/news/microsoft-brings-1gw-of-capacity-online-in-latest-quarter-again#:~:text=Microsoft%20has%20said%20that%20it%20is%20expecting%20Q4%20to%20have%20a%20capex%20of%20\$244%20billion%20and%20full%20year%20spend%20to%20reach%20\$249%20billion)). Meta's Q4 2025 earnings call in February 2026 projected 2026 capex at **\$200 billion**, with a range of **\$180 billion to \$220 billion** (Source: [meta.com](https://www.meta.com/amp/2026/04/29/meta-q1-earnings-report-2026.html#:~:text=Wall%20Street's%20\$27.57%20billion%20estimate%20of%20\$242.75%20billion%20average%20estimate)). Google's Q4 2025 earnings call in February 2026 projected 2026 capex at **\$200 billion**, with a range of **\$180 billion to \$220 billion** (Source: [google.com](https://www.google.com/pressroom/newsroom/2026/02/05/google-q4-2025-earnings-report)). The scale of this buildout is now large enough that independent analysts disagree meaningfully about how to measure it. Goldman Sachs Research put the Wall Street consensus for the hyperscaler group's 2026 capital spending at **\$527 billion** as of December 2025, up from \$465 billion just months earlier (Source: [goldmansachs.com](https://www.goldmansachs.com/insights/articles/ai-infrastructure-spending-2026)). Morgan Stanley's own chief executive told investors on the bank's second quarter 2026 earnings call that data center capital expenditure, initially projected at \$575 billion for 2026, is "coming in at about \$850 billion," with 2027 projected at \$1.3 trillion and 2028 potentially reaching \$1.5 trillion (Source: [seekingalpha.com](https://seekingalpha.com/news/4588888-morgan-stanley-raises-2026-ai-capex-forecast-to-850-billion)). The International Energy Agency, using its own satellite-based tracking of data centers, reported that the largest technology companies' capital expenditure exceeded **\$400 billion** in 2025 and is expected to jump by another **75 percent** in 2026 (Source: [iea.org](https://www.iea.org/articles/ai-data-centers-energy-consumption-2026)). These figures differ because they use different company sets and different definitions of what counts as AI infrastructure, but they converge on the same directional conclusion: hyperscaler capex roughly doubled in two years and shows no sign of plateauing in 2026.

The spending is being driven by acute compute scarcity rather than speculative overbuilding, according to the companies themselves. Amazon Web Services (AWS) chief executive Matt Garman's own operating team said it is "monetizing capacity as fast as we can install it" (Source: [iodive.com](https://www.iodive.com/news/aws-ceo-matt-garman-ai-capex-2026/)). Google Cloud reported it remains "compute constrained in the near term" even after tripling capex guidance since 2022 (Source: [cnbc.com](https://www.cnbc.com/2026/02/05/google-cloud-ai-capex-2026/)). Microsoft chief financial officer Amy Hood said the company expects "to remain constrained at least through 2026" (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/en/news/microsoft-brings-1gw-of-capacity-online-in-latest-quarter-again)). Named infrastructure projects illustrate the physical scale involved: Meta's Hyperion campus in Richland Parish, Louisiana, is now expected to cost more than **\$50 billion** to reach 5 gigawatts of capacity, nearly double the \$27 billion figure disclosed only nine months earlier (Source: [cnbc.com](https://www.cnbc.com/2026/02/05/meta-hyperion-campus-2026/)); the Stargate campus in Abilene, Texas, run by Oracle for OpenAI, is projected to house up to **450,000 Nvidia GB200** graphics processing units (GPUs) on 1.2 gigawatts of power (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/en/news/oracle-openai-stargate-campus-2026/)); and Amazon's Project Rainier in Indiana, built for Anthropic, deployed nearly 500,000 Trainium2 chips at a disclosed cost of **\$11 billion** (Source: [cnbc.com](https://www.cnbc.com/2026/02/05/amazon-project-rainier-2026/)).

This report examines how much each of the four largest hyperscalers is spending in 2026, how those figures have moved since the start of the year, what is constraining and enabling that spending ([memory chip shortages](#), [power availability](#), depreciation accounting choices and vendor financing arrangements among Nvidia, OpenAI and Oracle), and what the implications are for investors, enterprise customers and the broader technology supply chain. It draws on earnings releases, U.S. Securities and Exchange Commission (SEC) filings, investor presentations and research from Goldman Sachs, Morgan Stanley and the IEA, verified directly against each source. As of July 2026, the consensus among company disclosures and independent analysts is that combined big four hyperscaler capex will land somewhere between \$650 billion and \$770 billion for calendar 2026, with Oracle's roughly \$55.7 billion to \$70 billion in related cloud infrastructure spending pushing the wider "hyperscaler" definition used by banks like Morgan Stanley closer to \$850 billion.

Introduction and Background

The term **hyperscaler** refers to a small group of technology companies, chiefly Amazon, Microsoft, Alphabet and Meta, that operate cloud computing and data center infrastructure at a scale large enough to serve millions of customers and, increasingly, to train and run frontier artificial intelligence (AI) models. **Capital expenditure**, commonly abbreviated **capex**, is the accounting term for money spent acquiring or building long lived physical assets, in this case data centers, servers, custom silicon, networking equipment and, increasingly, direct investments in electricity generation. Unlike operating expenses, capex is capitalized on the balance sheet and depreciated over several years rather than expensed immediately, a distinction that matters considerably for how AI spending affects reported profit, as later sections of this report explain.

The scale of 2026 hyperscaler capex has no precedent in the history of the technology industry. Amazon's guided **\$200 billion** for calendar 2026 alone exceeds the entire annual capital budget of most national governments' infrastructure programs, and it followed a 2025 in which Amazon had already spent **\$131 billion**, itself up from **\$83 billion** the year before (Source: [cnbc.com](https://www.cnbc.com/2026/02/05/why-amazons-ceo-is-confident-with-200-billion-spending-plan.html#:~:text=spending%20roughly%20%24131%20billion%20on%20purchases%20of%20property%20and%20equipment%20in%202025%2C%20up%20from%20about%20%2483%20billion%20in%20the%20year%20pr)). Trade publication CIO Dive summarized the pattern bluntly: Amazon, Google and Microsoft combined expect to invest more than **\$500 billion** in capital expenditures in 2026 as the companies race to build out AI

infrastructure, and that figure excludes Meta and Oracle entirely (Source: [ciodive.com](https://www.ciodive.com/news/amazon-adds-200b-ai-buildout-rush/811633/#:~:text=Amazon%2C%20Google%20and%20Microsoft%20combined%20expect%20to%20invest%20more%20than%20%24500%20billion%20in%20capital%20expenditures%20in%202026)). Reuters separately reported that Microsoft, Alphabet, Meta and Amazon together are expected to spend more than **\$500 billion** on AI in 2026 based on guidance available as of late January 2026, before Meta's April upward revision and Alphabet's own increase were factored in (Source: [reuters.com](https://www.reuters.com/business/retail-consumer/microsoft-edges-past-cloud-growth-expectations-2026-01-28/#:~:text=Microsoft%2C%20Alphabet%20and%20Meta%20and%20Amazon%20are%20expected%20to%20spend%20more%20than%20%24500%20billion%20on%20AI%20this%20year)). What is new in the 2026 cycle is not simply the dollar figure but the reasoning behind it. Prior technology investment waves, including the fiber optic buildout of the late 1990s, were driven substantially by anticipated future demand. The current cycle, by contrast, is explicitly described by all four companies as a response to demand that already exceeds available supply. Alphabet chief executive Sundar Pichai told analysts in February 2026 that "we've been supply-constrained, even as we've been ramping up our capacity" (Source: [reuters.com](https://www.reuters.com/business/retail-consumer/microsoft-edges-past-cloud-growth-expectations-2026-01-28/#:~:text=Microsoft%2C%20Alphabet%20and%20Meta%20and%20Amazon%20are%20expected%20to%20spend%20more%20than%20%24500%20billion%20on%20AI%20this%20year)). Google Cloud vice president Amin Vahdat told an internal all-hands meeting that the company needs to double its serving capacity for AI every six months just to keep pace with demand (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/en/news/google-cloud-vice-president-amin-vahdat-told-an-internal-all-hands-meeting-that-the-company-needs-to-double-its-serving-capacity-for-ai-every-six-months-just-to-keep-pace-with-demand)). That framing matters for how investors should read the headline figures in this report: the companies are not describing a discretionary bet so much as a constrained scramble for chips, power and physical building capacity, a dynamic examined in detail in the sections that follow.

This report focuses on calendar year 2026 guidance and actual spending as disclosed through mid July 2026, drawing on the most recent quarterly results available for each company: Amazon's and Alphabet's and Meta's first quarter 2026 results (all reported April 29, 2026, since each operates on a calendar fiscal year), and Microsoft's fiscal third quarter 2026 results (also reported April 29, 2026, covering the quarter ended March 31, 2026, since Microsoft's fiscal year runs from July through June). Microsoft's fiscal fourth quarter, and Amazon's, Alphabet's and Meta's calendar second quarter results, were not yet public as of this report's July 18, 2026 reference date.

Methodology and Scope

This analysis synthesizes disclosures from four primary categories of sources. First, official company disclosures, including SEC filings, investor relations press releases and investor day presentations, which are treated as authoritative for each company's own capex figures and guidance. Second, earnings call commentary from chief executives and chief financial officers, quoted directly where it explains the reasoning behind spending decisions. Third, independent research from banks and research firms, specifically Goldman Sachs Research and Goldman Sachs Global Institute, and comments from Morgan Stanley's chief executive on the bank's own quarterly earnings call, used to contextualize company-specific guidance against the wider market. Fourth, the IEA's 2026 energy and AI research, which offers a rare third party, satellite-verified cross check on hyperscaler capital spending and its associated electricity demand.

Two definitional issues recur throughout this report and are worth stating up front. First, "capex" as reported in company cash flow statements typically captures purchases of property and equipment net of proceeds from sales, which for Amazon in the trailing twelve months ended March 31, 2026 totaled **\$147.3 billion**, an increase of 67 percent year over year (Source: ir.aboutamazon.com). This figure differs from, and is sometimes lower than, forward guidance figures because guidance often includes finance leases and other off balance sheet commitments that do not appear identically in the cash flow statement in the same period. Second, "hyperscaler" is used inconsistently across sources: some analyst notes include only Amazon, Microsoft, Alphabet and Meta, while Morgan Stanley's and the IEA's broader figures fold in Oracle, whose fiscal 2026 capex of **\$55.7 billion** is smaller in absolute terms than any of the big four but is growing faster in percentage terms than all of them (Source: [theregister.com](https://www.theregister.com/)). This report is explicit at each point about which company set and which definition of capex is being used.

Amazon: AWS and the \$200 Billion Build-Out

Amazon set the tone for the 2026 hyperscaler capex cycle when it disclosed its full year 2025 results on February 5, 2026. The company reported **\$131 billion** in purchases of property and equipment for 2025, itself up sharply from **\$83 billion** in 2024, and then guided to **\$200 billion** for 2026, a jump of more than 50 percent and roughly \$50 billion above what analysts had modeled (Source: [cnbc.com](https://www.cnbc.com/)). Amazon's own fourth quarter earnings release quoted chief executive Andy Jassy explaining that "with such strong demand for our existing offerings and seminal opportunities like AI, chips, robotics, and low earth orbit satellites, we expect to invest about \$200 billion in capital expenditures across Amazon in 2026, and anticipate strong long-term return on invested capital" (Source: ir.aboutamazon.com). Amazon's stock fell **11 percent** in after hours trading on the news, the sharpest single day reaction among the four hyperscalers' 2026 guidance disclosures (Source: [cnbc.com](https://www.cnbc.com/)), even though on the subsequent earnings call Jassy separately told analysts he was "confident" in the plan and added that "this isn't some sort of quixotic, top-line grab" (Source: [cnbc.com](https://www.cnbc.com/)).

The spending is concentrated in **AWS**, Amazon's cloud computing division, which the company said would receive "a primary focus" of the \$200 billion in 2026 outlays (Source: [ciodive.com](https://www.ciodive.com/)). AWS itself has continued to accelerate: fourth quarter 2025 segment sales grew **24 percent** year over year to **\$35.6 billion**, its fastest growth in 15 quarters (Source: [ciodive.com](https://www.ciodive.com/)), and then accelerated further to **28 percent** growth and **\$37.6 billion** in revenue in the first quarter of 2026, with segment operating income of **\$14.2 billion** (Source: ir.aboutamazon.com). Amazon's own SEC filing shows quarterly purchases of property and equipment of **\$44.2 billion** in the first quarter of 2026, up from \$25.0 billion in the same quarter of 2025 (Source: ir.aboutamazon.com). Amazon did not update its full year 2026 capex guidance alongside the first quarter results, leaving the \$200 billion figure from February as the operative number as of this report's publication.

Amazon's spending mix differs from its rivals in one important respect: a meaningful share is directed at custom silicon and non-traditional infrastructure rather than purchased Nvidia GPUs alone. Jassy said AWS's chips business, which includes the in-house **Trainium** training chips and **Inferentia** inference chips, topped a **\$20 billion** revenue run rate in the first quarter of 2026 and was growing at triple digit percentage rates (Source: ir.aboutamazon.com). The clearest illustration of that strategy is **Project Rainier**, a compute cluster built for Anthropic in New Carlisle, Indiana, which came online in mid 2025 at a disclosed cost of **\$11 billion** and now runs nearly half a million Trainium2 chips, delivering more than five times the compute power Anthropic used to train its earlier models (Source: [aboutamazon.com](https://www.aboutamazon.com/)) (Source: [cnbc.com](https://www.cnbc.com/)). AWS chief executive Matt Garman described the pace of construction as "cornfields to data centers, almost overnight" (Source: [cnbc.com](https://www.cnbc.com/)), and AWS separately disclosed that the cluster represents a **70 percent** increase in its AI computing infrastructure compared with prior deployments (Source: [datacentermagazine.com](https://www.datacentermagazine.com/)). Anthropic itself was expected to scale to more than one million Trainium2 chips for combined training and inference workloads by the end of 2025 (Source: [aboutamazon.com](https://www.aboutamazon.com/)).

Microsoft: Azure, OpenAI, and the \$190 Billion Fiscal Year

Microsoft's fiscal year runs from July through June, so its 2026 spending is best measured across four fiscal quarters rather than a single calendar year guidance figure, and the company has been unusually candid about the quarter to quarter mechanics of that spending. Fiscal second quarter 2026 (the October through December 2025 period) produced a then record **\$37.5 billion** in capital expenditure, a jump of nearly **66 percent** from a year earlier, with roughly two thirds directed toward "short-lived assets" such as GPUs and central processing units (CPUs) and **\$6.7 billion** allocated to data center leases (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)) (Source: [reuters.com](https://www.reuters.com/)). That quarter's Azure cloud revenue grew **39 percent**, narrowly beating a consensus estimate of 38.8 percent (Source: [reuters.com](https://www.reuters.com/)), but Microsoft shares still fell **6.5 percent** in after hours trading, reflecting investor unease that capex was outpacing Azure growth (Source: [reuters.com](https://www.reuters.com/)).

Fiscal third quarter 2026 (January through March 2026) capex came in at **\$31.9 billion**, lower than the prior quarter as chief financial officer Amy Hood had signaled it would be, with the same roughly two thirds share going to short-lived compute assets and \$4.7 billion to data center leases (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)). That same quarter, Microsoft's own earnings release quoted chief executive Satya Nadella noting that "our AI business surpassed an annual revenue run rate of \$37 billion, up 123% year-over-year," on total company revenue of \$82.9 billion, up 18 percent (Source: [microsoft.com](https://www.microsoft.com/)). Azure growth accelerated to **40 percent** on a constant currency basis that quarter, beating expectations of roughly 38 percent, even as overall capex growth outpaced revenue growth by enough to send the stock down about **3.9 percent** the following session (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)). Looking ahead, Microsoft told investors to expect fiscal fourth quarter capex of about **\$40 billion**, which would bring full fiscal year 2026 spending to roughly **\$190 billion**, a figure the company attributed partly to **\$25 billion** of higher than expected component pricing, chiefly memory chips (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)). That would represent well over double the roughly **\$80 billion** Microsoft spent in fiscal 2025 (Source: [ciodive.com](https://www.ciodive.com/)). Chief executive Satya Nadella told investors the company remains "on track to double our overall footprint in just two years" (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)), while Hood cautioned that "even with these additional investments...we expect to remain constrained at least through 2026" (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)).

Much of that spending is tied contractually to **OpenAI**, which Reuters reported plans to spend at least **\$281 billion** with Microsoft under their restructured partnership (Source: [reuters.com](https://www.reuters.com/)), including a specific commitment for OpenAI to buy **\$250 billion** of Azure cloud services (Source: [reuters.com](https://www.reuters.com/)). Microsoft's total cloud backlog, known as remaining performance obligation, more than doubled to **\$625 billion**, of which roughly **45 percent** was attributable to OpenAI alone, a concentration that Reuters flagged as underscoring Microsoft's continued reliance on a single customer even after diversifying with a **\$30 billion** cloud deal with Anthropic (Source: [reuters.com](https://www.reuters.com/)) (Source: [reuters.com](https://www.reuters.com/)).

Physically, the spending is materializing as **Fairwater**, a network of purpose-built AI datacenters that Microsoft describes as its "AI superfactory." The Wisconsin and Atlanta Fairwater sites represent, in Microsoft's own words, "tens of billions of dollars of investments and hundreds of thousands of cutting-edge AI chips" (Source: blogs.microsoft.com), connected across sites so that training jobs that once took months can be completed in weeks (Source: news.microsoft.com). Nadella said in April 2026 that the Wisconsin site was "going live, ahead of schedule," bringing together "hundreds of thousands of GB200s into a single seamless cluster" (Source: [datacenterdynamics.com](https://www.datacenterdynamics.com/)), a project Data Center Frontier had previously reported was part of plans to invest over **\$80 billion** in 2025 alone (Source: [datacenterfrontier.com](https://www.datacenterfrontier.com/)).

Google Cloud and Alphabet's Doubling Bet

Alphabet's February 4, 2026 earnings call produced what Reuters called a new bar for AI infrastructure spending among the hyperscalers (Source: [cnbc.com](#)). Alphabet's own fourth quarter 2025 earnings release quoted chief executive Sundar Pichai saying that "to meet customer demand and capitalize on the growing opportunities we have ahead of us, our 2026 CapEx investments are anticipated to be in the range of \$175 to \$185 billion," up from **\$91.45 billion** actually spent in 2025, a figure that itself exceeded analysts' average estimate of about \$115.26 billion by a wide margin (Source: [sec.gov](#)) (Source: [reuters.com](#)). Chief executive Sundar Pichai told analysts the company had "been supply-constrained, even as we've been ramping up our capacity," framing the spend as "an eye towards the future" rather than discretionary (Source: [reuters.com](#)). The reasoning was reinforced by results: Google Cloud's fourth quarter 2025 revenue surged **48 percent to \$17.7 billion**, its fastest growth pace in more than four years and, notably, faster than Microsoft Azure's growth rate for the first time in several years, according to D.A. Davidson analyst Gil Luria (Source: [reuters.com](#)).

Alphabet raised its guidance again at its first quarter 2026 earnings call on April 29, 2026, moving the range up to **\$180 billion to \$190 billion** and telling investors it expects 2027 capex to "significantly increase" further (Source: [cnbc.com](#)). First quarter 2026 capex itself came in at **\$35.7 billion**, alongside Google Cloud revenue that had accelerated further to **63 percent** growth and topped **\$20 billion** for the first time in a single quarter, with segment operating income of \$6.6 billion (Source: [cnbc.com](#)). Pichai again described the company as "compute constrained in the near term," adding that cloud revenue "would have been higher if we were able to meet the demand" (Source: [cnbc.com](#)). Alphabet's June 2026 investor presentation reiterated the \$180 billion to \$190 billion range and put the increase in historical context: the company spent approximately **\$31 billion** in capex in 2022, meaning 2026 spending is expected to be roughly six times larger in only four years (Source: [storage.googleapis.com](#)). To help fund that spending without relying solely on debt, Alphabet disclosed an equity distribution agreement on June 1, 2026 supporting an equity raise of up to **\$84.75 billion**, including underwritten common and preferred stock offerings, a **\$40 billion** at the market program, and a **\$10 billion** private placement with Berkshire Hathaway (Source: [stockitan.net](#)).

Google's AI compute strategy leans more heavily than its rivals on custom **Tensor Processing Units (TPUs)**, its purpose-built AI accelerator chips, rather than exclusively on purchased Nvidia GPUs. The seventh generation TPU, **Ironwood**, became generally available at Google Cloud Next in April 2026, delivering 4.6 petaFLOPS of peak compute per chip and 42.5 exaFLOPS when 9,216 chips are linked into a single superpod (Source: [thenextweb.com](#)). Anthropic has become the anchor customer for both the current and next TPU generation, expanding its commitment with Google to **3.5 gigawatts** of compute capacity by 2027 (Source: [thenextweb.com](#)), an expansion on top of a previously announced **\$50 billion** Anthropic investment in domestic computing infrastructure (Source: [datacenterknowledge.com](#)). Even so, Google Cloud vice president Amin Vahdat told employees the company must double its serving capacity for AI every six months, calling the infrastructure race "the most critical and also the most expensive part of the AI race" while adding that Google is not trying to simply "outspend the competition" (Source: [datacenterdynamics.com](#)).

Meta: Superintelligence Labs and the Infrastructure Pivot

Meta's guidance history in 2026 has been the most volatile among the four hyperscalers. On its fourth quarter 2025 earnings call, held January 28, 2026, chief financial officer Susan Li guided to full year 2026 capex of **\$115 billion to \$135 billion**, up sharply from **\$72.22 billion** spent in 2025. Meta's own earnings release stated that the company anticipated "2026 capital expenditures, including principal payments on finance leases, to be in the range of \$115-135 billion, with year-over-year growth driven by increased investment to support our Meta Superintelligence Labs efforts and core business" (Source: [investor.atmeta.com](#)), with the same release disclosing that capital expenditures, including finance lease principal payments, were \$22.14 billion for the fourth quarter and \$72.22 billion for the full year 2025 (Source: [investor.atmeta.com](#)). Total 2026 expense guidance was set at **\$162 billion to \$169 billion**, with Li explaining that "the majority of expense growth will be driven by infrastructure costs, which includes third-party cloud spend, higher depreciation, and higher infrastructure operating expenses" (Source: [datacenterdynamics.com](#)).

Three months later, on its April 29, 2026 first quarter call, Meta raised the range again to **\$125 billion to \$145 billion**, even though the quarter's own capex of **\$19.84 billion** landed well below the **\$27.57 billion** Wall Street had modeled, based on StreetAccount estimates (Source: [cnbc.com](#)) (Source: [sec.gov](#)). Meta shares still fell about **7 percent** in extended trading, reflecting a mixed reaction: the higher forecast alarmed some investors even as the quarter's actual spending had been lower than expected, an unusual combination that analyst Matt Britzman of Hargreaves Lansdown attributed largely to rising memory component prices rather than a change in the underlying build plan (Source: [reuters.com](#)). Meta's own earnings release attributed the raised range directly to component costs: "we anticipate 2026 capital expenditures...to be in the range of \$125-145 billion, increased from our prior range of \$115-135 billion. This reflects our expectations for higher component pricing this year and, to a lesser extent, additional data center costs to support future year capacity" (Source: [sec.gov](#)). First quarter 2026 revenue of **\$56.31 billion** beat the LSEG-compiled analyst estimate of \$55.45 billion (Source: [reuters.com](#)), a 33 percent year over year increase.

Meta has paired its higher capex guidance with workforce reductions, telling Reuters it would begin layoffs affecting roughly **10 percent** of its workforce, or about **8,000 employees**, starting in May 2026 (Source: [cnbc.com](#)). Chief executive Mark Zuckerberg framed the reductions as building "the next evolution of our company" around top performers rather than as a response to capex pressure (Source: [reuters.com](#)), and the company reported it kept its total expense forecast for 2026 unchanged despite the higher capital expenditure outlook, implying anticipated savings elsewhere (Source: [reuters.com](#)). Meta's workforce stood at **77,986** people at the end of March 2026, up 1 percent year over year but down from 78,865 at the end of December 2025 (Source: [reuters.com](#)).

The physical anchor for Meta's spending is **Hyperion**, a data center campus under construction in Richland Parish, Louisiana, originally described in 2025 as a facility that would deliver 5 gigawatts of compute capacity, with a footprint large enough to cover most of Manhattan (Source: [techcrunch.com](#)) (Source: [datacenters.atmeta.com](#)). By July 2026, Meta disclosed the site's total cost had grown to more than **\$50 billion**, up from the **\$27 billion** figure revealed the previous October when Meta and Blue Owl Capital formed a joint venture to help finance and manage the buildout, at the time planned as a 2 gigawatt facility (Source: [cnbc.com](#)). The first 2 gigawatt phase remains on track for 2030, with the full 5 gigawatt buildout expected around 2032 (Source: [datacenterknowledge.com](#)). Meta also operates a separate 1 gigawatt supercluster, **Prometheus**, in New Albany, Ohio, intended to come online in 2026 (Source: [techcrunch.com](#)).

Analysis of Key Segments: GPUs, Memory, and Financing

Behind the headline capex figures sit three structural constraints that increasingly determine how fast hyperscalers can actually deploy the capital they are guiding to: chip supply, memory pricing, and financing structure.

On the compute side, Nvidia's own results are the clearest proxy for how much of hyperscaler capex is flowing into GPUs. Nvidia's data center segment, which sells almost exclusively to hyperscalers, AI labs and specialized cloud providers, generated record revenue of **\$62.3 billion** in its fiscal fourth quarter 2026 (ended January 25, 2026), up **75 percent** year over year, contributing to full fiscal year revenue of **\$215.9 billion**, up **65 percent** (Source: [investor.nvidia.com](#)). That growth accelerated further into fiscal 2027: first quarter data center revenue reached **\$75.2 billion**, up **92 percent** year over year, on total company revenue of \$81.6 billion, up 85 percent (Source: [investor.nvidia.com](#)). Chief executive Jensen Huang described the pattern as "the buildout of AI factories, the largest infrastructure expansion in human history" (Source: [investor.nvidia.com](#)).

Memory chips have emerged as the tightest bottleneck constraining how fast that GPU spending can be converted into usable capacity. **High-bandwidth memory (HBM)**, the specialized memory stacked directly next to AI accelerators, and conventional **dynamic random-access memory (DRAM)** compete for the same fabrication capacity, and manufacturers have prioritized the higher-margin HBM business, squeezing conventional DRAM supply. Data centers are expected to consume more than **70 percent** of the high-end memory chips manufacturers produce in 2026, according to TrendForce (Source: [iq.com](#)). The resulting price shock has been severe: TrendForce recorded conventional DRAM contract prices rising **93 to 98 percent** quarter on quarter in the first quarter of 2026, followed by a further **58 to 63 percent** increase in the second quarter, with NAND flash contract prices climbing a comparable 85 to 90 percent and then 55 to 60 percent over the same two quarters (Source: [iq.com](#)). Samsung Electronics said in July 2026 that it expected second quarter operating profit of approximately 89.4 trillion won, or roughly **\$59 billion**, about **19 times** higher than a year earlier (Source: [iq.com](#)), and SK Hynix's chairman has warned that global memory supply is likely to remain roughly **20 percent** below demand through 2030 (Source: [iq.com](#)). Microsoft's own disclosure that \$25 billion of its fiscal 2026 capex increase was driven by higher component pricing, and Meta's parallel attribution of its raised guidance to higher component pricing, are both direct references to this memory shortage rather than to expanded physical build plans.

Financing structure is the third constraint, and the one attracting the most scrutiny from independent analysts. A substantial share of the compute capacity being built for OpenAI in particular rests on financing arrangements in which chipmakers and cloud providers fund the very customers who then buy their products. Nvidia has agreed to invest up to \$100 billion in OpenAI, Oracle booked a \$300 billion compute order from OpenAI and then raised debt to build the data centers required to fill it, and AMD extended OpenAI a warrant for a stake of roughly one tenth of the chipmaker, a structure that has drawn direct comparisons to the vendor financing arrangements that preceded the 2000 telecommunications equipment bust (Source: [finance.yahoo.com](#)). Analysts at Oxford Economics warned in a research note that "the scale of recent investment increases by tech firms already indicates that they are taking significant risks," adding that if AI productivity gains and the resulting return on investment prove limited or delayed, "a sharp correction in tech stocks, with negative knock-ons for the real economy, would be very likely" (Source: [finance.yahoo.com](#)) (Source: [finance.yahoo.com](#)).

Data Analysis and Evidence

Table 1 below summarizes each hyperscaler's disclosed 2025 actual capex, its initial and most recently revised 2026 guidance, and its most recently reported quarterly capex figure as of this report's July 2026 reference date.

COMPANY	2025 ACTUAL CAPEX	INITIAL 2026 GUIDANCE	LATEST 2026 GUIDANCE	MOST RECENT QUARTERLY CAPEX
Amazon	\$131 billion (Source: cnbc.com)	\$200 billion, Feb 5, 2026 (Source: ir.aboutamazon.com)	\$200 billion (unchanged as of Q1 2026)	\$44.2 billion, Q1 2026 (Source: ir.aboutamazon.com)
Microsoft (fiscal year)	~\$80 billion, FY2025 (Source: ciotive.com)	Ramping through fiscal quarters, no single figure issued	~\$190 billion, FY2026 total, guided April 2026 (Source: datacenterdynamics.com)	\$31.9 billion, fiscal Q3 2026 (Source: datacenterdynamics.com)
Alphabet	\$91.45 billion (Source: reuters.com)	\$175 to \$185 billion, Feb 4, 2026 (Source: sec.gov)	\$180 to \$190 billion, reaffirmed June 2026 (Source: storage.googleapis.com)	\$35.7 billion, Q1 2026 (Source: cnbc.com)
Meta	\$72.22 billion (Source: investor.atmeta.com)	\$115 to \$135 billion, Jan 28, 2026 (Source: investor.atmeta.com)	\$125 to \$145 billion, Apr 29, 2026 (Source: sec.gov)	\$19.84 billion, Q1 2026 (Source: sec.gov)
Oracle (fiscal year, context)	\$21.2 billion, FY2025 (Source: theregister.com)	~\$50 billion, prior FY2026 guidance	\$55.7 billion actual, FY2026; ~\$70 billion guided for FY2027 (Source: thenextweb.com)	\$55.7 billion, FY2026 total (Source: theregister.com)

Summing the midpoints of the four hyperscalers' latest 2026 guidance (Amazon \$200 billion, Microsoft roughly \$190 billion, Alphabet's midpoint of \$185 billion, and Meta's midpoint of \$135 billion) produces a combined total of approximately **\$710 billion** for the big four alone, before Oracle's roughly \$55.7 billion to \$70 billion is added. That arithmetic helps explain why Morgan Stanley's five-company figure of **\$850 billion** for 2026, cited earlier in this report, sits noticeably above Goldman Sachs's four or five company Wall Street consensus of \$527 billion from December 2025: the gap reflects both a several month gap in when each estimate was made, during which guidance rose repeatedly, and differences in which companies and which cost categories (leases, prepayments, and power infrastructure among them) are included.

Table 2 below lists the named, physically identifiable AI infrastructure projects referenced throughout this report, illustrating how the capex figures translate into gigawatts of power and specific chip deployments.

PROJECT	OWNER / TENANT	LOCATION	SCALE	DISCLOSED COST
Stargate Abilene	Oracle (developer) / OpenAI (tenant)	Abilene, Texas	1.2 gigawatts; up to 450,000 Nvidia GB200 GPUs (Source: datacenterdynamics.com)	\$15.9 billion current, up to \$31.9 billion projected (Source: enr.com)
Fairwater superfactory	Microsoft	Wisconsin and Atlanta, connected network	Hundreds of thousands of GB200 and GB300 GPUs (Source: news.microsoft.com)	Tens of billions of dollars (Source: blogs.microsoft.com)
Hyperion	Meta	Richland Parish, Louisiana	5 gigawatts targeted; 2 gigawatts by 2030 (Source: datacenterknowledge.com)	More than \$50 billion (Source: cnbc.com)
Project Rainier	AWS (developer) / Anthropic (tenant)	New Carlisle, Indiana	Nearly 500,000 Trainium2 chips, scaling past 1 million (Source: aboutamazon.com)	\$11 billion (Source: cnbc.com)
Ironwood / Anthropic TPU network	Google Cloud (developer) / Anthropic (tenant)	Multiple U.S. sites	Expanding to 3.5 gigawatts by 2027 (Source: thenextweb.com)	Part of a \$50 billion Anthropic infrastructure commitment (Source: datacenterknowledge.com)

The projects in Table 2 make a broader point: Anthropic alone appears as the named tenant behind at least three of the five flagship projects listed (Amazon's Rainier, Google's Ironwood network, and indirectly via its own \$50 billion domestic infrastructure plan), underscoring how a handful of frontier AI labs, OpenAI and Anthropic chief among them, are the ultimate demand source that hyperscalers cite to justify their own capex increases.

Independent forecasts of aggregate hyperscaler or data center capex vary by as much as \$300 billion for 2026 depending on scope and timing, a discrepancy worth stating plainly rather than smoothing over. Goldman Sachs Research's Wall Street consensus figure of **\$527 billion**, compiled in December 2025, reflected analyst models that predate Alphabet's and Meta's subsequent guidance increases and used a narrower company set (Source: goldmansachs.com). Goldman's own Global Institute, publishing in May 2026 using a supply side model anchored to Nvidia's forward data center revenue estimates, arrived at **\$765 billion** for 2026 alone, rising to \$1.6 trillion annually by 2031 and totaling roughly \$7.6 trillion cumulatively across that period (Source: goldmansachs.com). Goldman separately noted that historical technology investment cycles, such as the late 1990s telecommunications buildout, suggest as much as \$200 billion of further upside to current hyperscaler 2026 estimates given the industry's still-strong balance sheets (Source: goldmansachs.com). Morgan Stanley chief executive Ted Pick told the bank's own July 2026 earnings call that data center capex initially projected at \$575 billion for 2026 is "coming in at about \$850 billion," rising to \$1.3 trillion in 2027 and potentially \$1.5 trillion in 2028, and cautioned that the industry is only "around 10% to 15% of the way through the investment cycle" (Source: seekingalpha.com). The IEA's own tracking, based on satellite imagery of physical construction rather than analyst models, found that AI-specific data centers, described in its report as "AI factories," more than tripled in capacity in the 18 months preceding its 2026 report (Source: iea.org), and that global electricity demand from data centers grew 17 percent in 2025 overall, while demand specifically from AI-focused data centers surged **50 percent** (Source: iea.org).

Case Studies and Real-World Examples

Microsoft's Fairwater Superfactory: Wisconsin and Atlanta

Microsoft's Fairwater network is the clearest example of a hyperscaler treating its AI data centers as a single distributed machine rather than a portfolio of independent buildings. The Wisconsin site, in Mount Pleasant, was introduced in September 2025 as, in Microsoft's own description, "the largest and most sophisticated AI factory we've built yet" (Source: blogs.microsoft.com), designed to deliver up to 10 times the performance of prior supercomputers through high bandwidth GPU interconnects and facility scale liquid cooling (Source: datacenterfrontier.com). By November 2025, Microsoft had connected the Wisconsin site to a second Fairwater facility in Atlanta, describing the pair as its "first AI superfactory," a network capable of training new generations of AI models in weeks rather than months (Source: news.microsoft.com). By April 2026, chief executive Satya Nadella said the Wisconsin site was going live ahead of schedule, calling it "the world's most powerful AI data center" and describing plans to bring "hundreds of thousands of GB200s into a single seamless cluster" (Source: datacenterdynamics.com). The project sits within a fiscal 2025 capital plan that had already exceeded \$80 billion (Source: datacenterfrontier.com), and Microsoft's broader cloud footprint of more than 400 data centers across 70 regions globally (Source: blogs.microsoft.com).

Meta's Hyperion: A Single Project Whose Cost Doubled in Nine Months

Meta's Hyperion campus in Richland Parish, Louisiana, is arguably the single clearest illustration of how quickly 2026 AI infrastructure cost estimates have moved. When Meta and investment firm Blue Owl Capital formed a joint venture in October 2025 to develop the site, it was planned as a 2 gigawatt facility with an associated cost of \$27 billion. By July 2026, Meta disclosed the project had been expanded to a 5 gigawatt target with a total cost exceeding **\$50 billion**, according to a company blog post (Source: [cnbc.com](https://www.cnbc.com)). Analysts covering the project noted that at 5 gigawatts, Hyperion "is no longer just a large data center" but a facility large enough to influence regional electricity generation, transmission, and utility planning across northeast Louisiana (Source: datacenterknowledge.com). Louisiana governor Jeff Landry credited the project with revitalizing "one of our state's beautiful rural areas" and creating high paying jobs (Source: datacenters.atmeta.com), and Meta disclosed the site would draw more than 7,500 skilled trade workers at peak construction. The episode illustrates a broader pattern across the industry in 2026: initial cost disclosures for gigawatt scale AI campuses have consistently proven to be interim figures rather than final budgets, typically revised upward as detailed engineering and power procurement proceed.

Oracle and OpenAI's Stargate: The Physical Face of a \$500 Billion Pledge

Announced in January 2025 by OpenAI, SoftBank, Oracle and MGX, the **Stargate** initiative was pitched as a five year, up to \$500 billion infrastructure program to build the world's largest AI supercomputing clusters (Source: [datacenterknowledge.com](#)). Its flagship site in Abilene, Texas, came online in September 2025, described by OpenAI's chief financial officer Sarah Friar as just the beginning: "we're just getting going here in Abilene, Texas, but you'll see this all around the United States and beyond" (Source: [cnbc.com](#)). Oracle founder Larry Ellison said the Abilene campus would eventually house more than 450,000 Nvidia GB200 GPUs and consume 1.2 gigawatts of power, "enough power for one million four-bedroom homes in the United States," which he called "a pretty good-sized city" (Source: [datacenterdynamics.com](#)). Independent tracking firm Epoch AI, using satellite imagery, cooling equipment models and regulatory filings, estimated the site's current capital cost at \$15.9 billion, rising to a projected \$31.9 billion as it scales toward roughly one million H100-equivalent chips of compute (Source: [epoch.ai](#)). By June 2026, industry coverage described Stargate as having evolved "from a \$500 billion promise into a real-world test of whether the AI industry can build infrastructure at a scale never attempted before," with success hinging on whether the initiative can scale faster than the risks it creates around energy, financing, governance and community consent (Source: [datacenterknowledge.com](#)). Oracle's own fiscal 2026 results reflect the strain: the company's capex reached \$55.7 billion for the year, overshooting its own prior guidance, while its remaining performance obligations (contracted revenue not yet recognized) grew from \$553 billion to **\$638 billion** in the fourth fiscal quarter alone (Source: [investor.oracle.com](#)), a backlog that reportedly includes \$300 billion attributable to OpenAI alone (Source: [theregister.com](#)).

Amazon's Project Rainier: Custom Silicon at Gigawatt Scale

Project Rainier, in New Carlisle, Indiana, demonstrates Amazon's distinct strategy of pairing gigawatt-scale infrastructure with proprietary chips rather than relying solely on merchant GPUs. Built exclusively to train and run Anthropic's Claude models, the 1,200 acre site was farmland roughly a year before it opened, and AWS chief executive Matt Garman described the transformation as "cornfields to data centers, almost overnight" (Source: [cnbc.com](#)). Amazon has committed to building around 30 data centers at the site over the coming years (Source: [cnbc.com](#)). The facility, unlike many competing announcements, was already fully operational at the time of its unveiling rather than a future promise, with AWS describing Project Rainier as providing Anthropic "more than five times the compute power" it had previously used for model training (Source: [aboutamazon.com](#)). Data Centre Magazine reported the cluster represented a 70 percent increase in AWS's total AI computing infrastructure in a single project, deployed less than 12 months after it was first announced at AWS's re:Invent conference in December 2024 (Source: [datacentremagazine.com](#)).

Google's Ironwood and the Anthropic TPU Expansion

Google's approach illustrates a fourth distinct model: rather than building single named mega campuses on the scale of Hyperion or Stargate, Google has expanded its custom TPU fleet and signed multi-gigawatt capacity commitments with external AI labs. The seventh generation TPU, Ironwood, became generally available at Google Cloud Next in April 2026 as, in Google's words, the "first Google TPU for the age of inference," delivering 4.6 petaFLOPS per chip and enabling superpods of 9,216 chips that together reach more than 24 times the capacity of El Capitan, currently the world's most powerful conventional supercomputer (Source: [thenextweb.com](#)). Anthropic's expanding relationship with Google, now reaching 3.5 gigawatts of committed compute by 2027, was described by Data Center Knowledge as showing "how surging enterprise demand is forcing AI companies to lock in long-term compute at industrial scale" (Source: [datacenterknowledge.com](#)), a shift the report characterized as resembling "energy procurement rather than traditional cloud infrastructure" (Source: [datacenterknowledge.com](#)).

Implications and Future Directions

The most consequential unresolved question in hyperscaler AI capex is not how much is being spent but how that spending is being accounted for, and depreciation policy sits at the center of the debate. Under U.S. accounting rules, the useful life of a long lived asset such as a server is an estimate, and revising that estimate changes reported operating income prospectively without restating prior periods (Source: [shittysits.substack.com](#)). Microsoft extended the depreciable life of its server and network equipment from four to six years starting fiscal 2023, a change it said would benefit that year's operating income by approximately \$3.7 billion (Source: [shittysits.substack.com](#)), and Alphabet made a similar four to six year extension the same year, expecting a roughly \$3.4 billion reduction in depreciation expense that in practice came in at \$3.9 billion, adding \$3.0 billion to net income (Source: [shittysits.substack.com](#)). Amazon, notably, moved in the opposite direction: after extending server useful life from five to six years in 2024, it reversed course in 2025, shortening a subset back to five years and citing the "increased pace of technology development, particularly in the area of artificial intelligence and machine learning" (Source: [shittysits.substack.com](#)). Investor Michael Burry has publicly argued that this divergence understates industry-wide depreciation expense from 2026 through 2028 by approximately \$176 billion, implying Oracle's 2028 earnings could be overstated by 26.9 percent and Meta's by 20.8 percent, though the analysis cited here does not adopt Burry's specific dollar figures and treats them as one contested estimate among several (Source: [shittysits.substack.com](#)). The practical takeaway for readers of hyperscaler earnings is that reported operating margins during the AI buildout are sensitive to an accounting estimate that companies themselves control and that is not comparable on a like for like basis across the four hyperscalers, since each has chosen a different useful life assumption for substantially similar Nvidia and custom silicon hardware.

A second implication concerns financing structure and systemic risk. Because Amazon, Microsoft, Alphabet and Meta are, unlike many prior infrastructure booms, funding the bulk of their spending from operating cash flow and, increasingly, targeted debt and equity issuance (Alphabet's \$84.75 billion equity program and Oracle's disclosed plan to raise roughly \$40 billion in debt and equity in fiscal 2027 being the clearest examples) (Source: [stocktitan.net](#)) (Source: [theregister.com](#)), the risk of an abrupt capex pullback is comparatively lower for the big four than it is for the AI labs and specialized cloud providers, such as OpenAI and CoreWeave, that depend on vendor financing from Nvidia, Oracle and AMD to make good on multi hundred billion dollar compute commitments. That distinction matters for how investors should interpret any future slowdown: a hyperscaler capex deceleration driven by satisfied demand would look very different from a disorderly unwind triggered by a weak link in the OpenAI centered financing web that Oxford Economics and other analysts have flagged as reminiscent of the vendor financing arrangements that preceded the 2000 telecommunications bust.

A third implication concerns physical bottlenecks that increasingly bind ahead of capital availability: power and memory. The IEA's finding that AI focused data center electricity demand surged 50 percent in 2025 alone, against a broader 17 percent increase in all data center electricity demand (Source: [iea.org](#)), means that even hyperscalers with unconstrained access to capital face multi year queues for grid interconnection, a dynamic evident in Meta's Hyperion timeline, where the first 2 gigawatt phase will not be complete until 2030 despite the project having been announced in 2025. Memory supply is a similarly hard constraint: with data centers set to absorb more than 70 percent of high end memory chip production in 2026 (Source: [iq.com](#)) and manufacturers such as Micron reporting HBM capacity sold out through 2027 (Source: [iq.com](#)), hyperscaler capex guidance for 2027 and beyond may increasingly be constrained less by willingness to spend and more by physical availability of chips, power and skilled construction labor.

Frequently Asked Questions (FAQs)

How much are Amazon, Microsoft, Google and Meta spending on AI in 2026? As of mid 2026, Amazon has guided to \$200 billion in total 2026 capital expenditure (Source: [ir.aboutamazon.com](#)), Microsoft has effectively guided to about \$190 billion for its fiscal 2026 year (Source: [datacenterdynamics.com](#)), Alphabet has guided to \$180 billion to \$190 billion (Source: [storage.googleapis.com](#)), and Meta has guided to \$125 billion to \$145 billion (Source: [sec.gov](#)). Combined, the midpoints of these ranges sum to roughly \$710 billion for the big four alone.

How much is AWS spending on AI infrastructure in 2026? Amazon has not broken out an AWS specific capex figure, but it said the "primary focus" of its \$200 billion 2026 capital plan is AWS (Source: [ciodive.com](#)), and AWS's own chips business, covering Trainium and Inferentia, reached a \$20 billion annualized revenue run rate in the first quarter of 2026 (Source: [ir.aboutamazon.com](#)).

What is Microsoft's Azure AI capex for fiscal 2026? Microsoft's fiscal 2026 (July 2025 through June 2026) total capex is expected to reach approximately \$190 billion, after quarterly capex of \$37.5 billion in fiscal Q2 and \$31.9 billion in fiscal Q3, with fiscal Q4 guided at approximately \$40 billion (Source: [datacenterdynamics.com](#)).

How much is Google Cloud spending on AI in 2026? Alphabet has guided to \$180 billion to \$190 billion in total 2026 capital expenditure (Source: [cnbc.com](#)), directed chiefly at servers, data centers and networking equipment supporting Google Cloud and Google DeepMind.

Why did Meta raise its 2026 capex guidance twice in four months? Meta's own earnings release attributed the April 2026 increase, from \$115 billion to \$135 billion up to \$125 billion to \$145 billion, to "expectations for higher component pricing this year and, to a lesser extent, additional data center costs to support future year capacity" (Source: [sec.gov](#)).

What is driving the GPU and memory chip shortage behind hyperscaler capex? Data centers are expected to consume more than 70 percent of high end memory chip production in 2026 (Source: [iq.com](#)), pushing conventional DRAM contract prices up 93 to 98 percent quarter on quarter in the first quarter of 2026 alone as manufacturers prioritize higher-margin, harder-to-produce HBM (Source: [iq.com](#)).

Are hyperscalers borrowing money to fund AI capex? Increasingly yes: Alphabet launched an equity program of up to \$84.75 billion in June 2026 (Source: [stocktitan.net](#)), and Oracle disclosed plans to raise roughly \$40 billion in debt and equity in fiscal 2027 on top of \$18 billion in debt raised in 2025 (Source: [theregister.com](#)). Amazon, Microsoft and Meta have so far relied more heavily on operating cash flow.

What are the main risks of the 2026 AI capex boom? Analysts point to three: depreciation policy choices that may understate current expenses by tens of billions of dollars industry wide (Source: [shittysits.substack.com](#)), circular vendor financing arrangements among Nvidia, OpenAI, Oracle and AMD that Oxford Economics has compared to pre dot-com bust telecom financing (Source: [finance.yahoo.com](#)), and physical bottlenecks in power and memory that could delay returns on invested capital regardless of how much capital is available.



Conclusion

The 2026 hyperscaler capex cycle is unlike any prior technology investment wave in both scale and structure. Amazon's \$200 billion guidance, Microsoft's roughly \$190 billion fiscal year plan, Alphabet's \$180 billion to \$190 billion range, and Meta's \$125 billion to \$145 billion range together point to combined big four spending in the neighborhood of \$700 billion for calendar 2026, a figure that climbs toward the \$850 billion Morgan Stanley's own chief executive cited on the bank's July 2026 earnings call once Oracle and other cloud infrastructure providers are included. Every one of the four companies has revised its guidance upward at least once since the start of 2026, and none has yet indicated that 2027 spending will be lower, with Alphabet, Microsoft and Oracle all explicitly signaling further increases.

What distinguishes this cycle from earlier infrastructure booms is the consistency with which company executives describe the spending as reactive rather than speculative: each of the four hyperscalers has told investors it remains supply constrained even after historic capex increases, a claim corroborated by independent data on GPU shipment growth, memory chip price inflation and gigawatt scale power procurement queues. At the same time, the depreciation accounting choices each company has made, the concentration of contracted revenue tied to a small number of AI labs such as OpenAI and Anthropic, and the vendor financing arrangements linking Nvidia, Oracle, AMD and OpenAI together mean that the ultimate return on this spending remains genuinely uncertain and will likely only become clear over the multi year period these companies themselves are now building for. Readers tracking this topic should watch three concrete signals in the second half of 2026 and into 2027: whether Microsoft's fiscal Q4 2026 capex lands at or above its guided \$40 billion, whether Amazon and Meta report calendar Q2 2026 results that hold to or exceed their respective guidance ranges, and whether memory chip prices, which both Microsoft and Meta have explicitly blamed for part of their 2026 spending increases, begin to ease as new fabrication capacity from Samsung, SK Hynix and Micron comes online after 2027.

Tags: hyperscaler ai capex 2026, amazon aws capex, microsoft azure capex, google cloud capex, meta ai infrastructure spending, ai data center spending, nvidia data center revenue, big tech capex comparison, ai infrastructure investment, gpu capacity spending

DISCLAIMER

This document is provided for informational purposes only. No representations or warranties are made regarding the accuracy, completeness, or reliability of its contents. Any use of this information is at your own risk. GPUSmith shall not be liable for any damages arising from the use of this document. This content may include material generated with assistance from artificial intelligence tools, which may contain errors or inaccuracies. Readers should verify critical information independently. All product names, trademarks, and registered trademarks mentioned are property of their respective owners and are used for identification purposes only. Use of these names does not imply endorsement. This document does not constitute professional or legal advice. For specific guidance related to your needs, please consult qualified professionals.