Special Report

HPC-AI Spending in the Cloud is Projected to Grow to Over $23B in 2029

Mark Nossokoff, Jaclyn Ludema, and Earl Joseph
September 2025

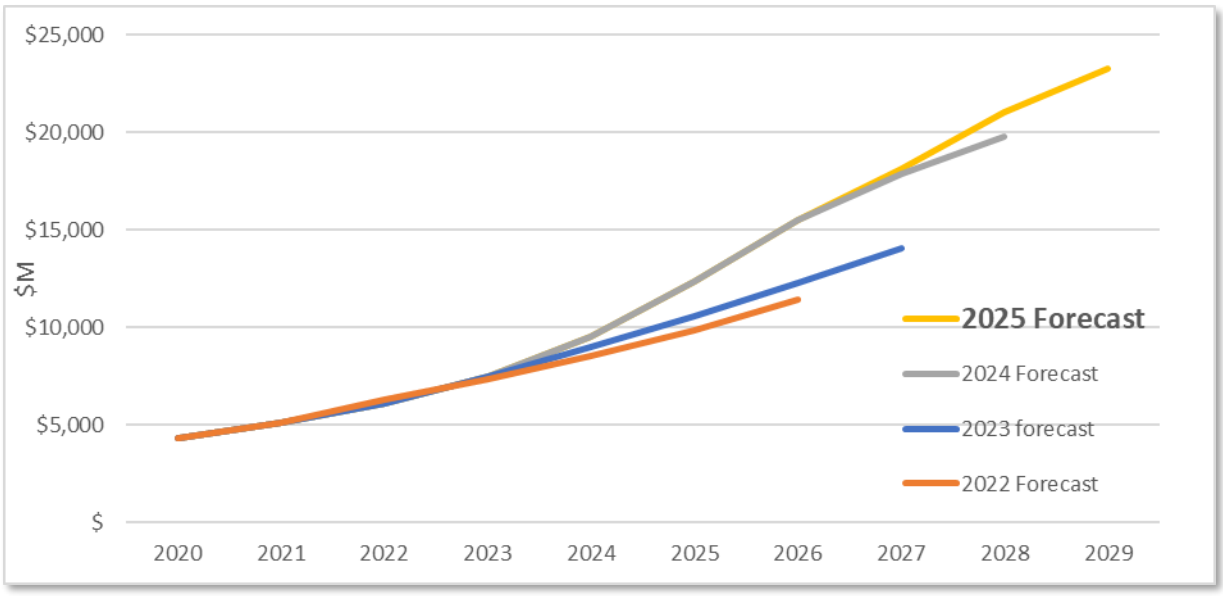
HYPERION RESEARCH OPINION

Driven by continued demands for dynamic and flexible access to the most advanced AI and HPC computational resources, and increased adoption of AI into scientific and engineering workflows, spending for HPC-AI resources in the cloud is projected to grow almost 20% annually over the 2025-2029 planning horizon. Accounting for approximately 15% of overall spending on HPC-AI resources (inclusive of both on-premises and cloud) in 2024, cloud share of spending of the total HPC-AI market is expected to make up over 20% of the market in 2029.

Figure 1 compares the current HPC-AI cloud spending forecast with prior forecasts, while Figure 2 shows the total spending on HPC-AI resources through 2029.

FIGURE 1

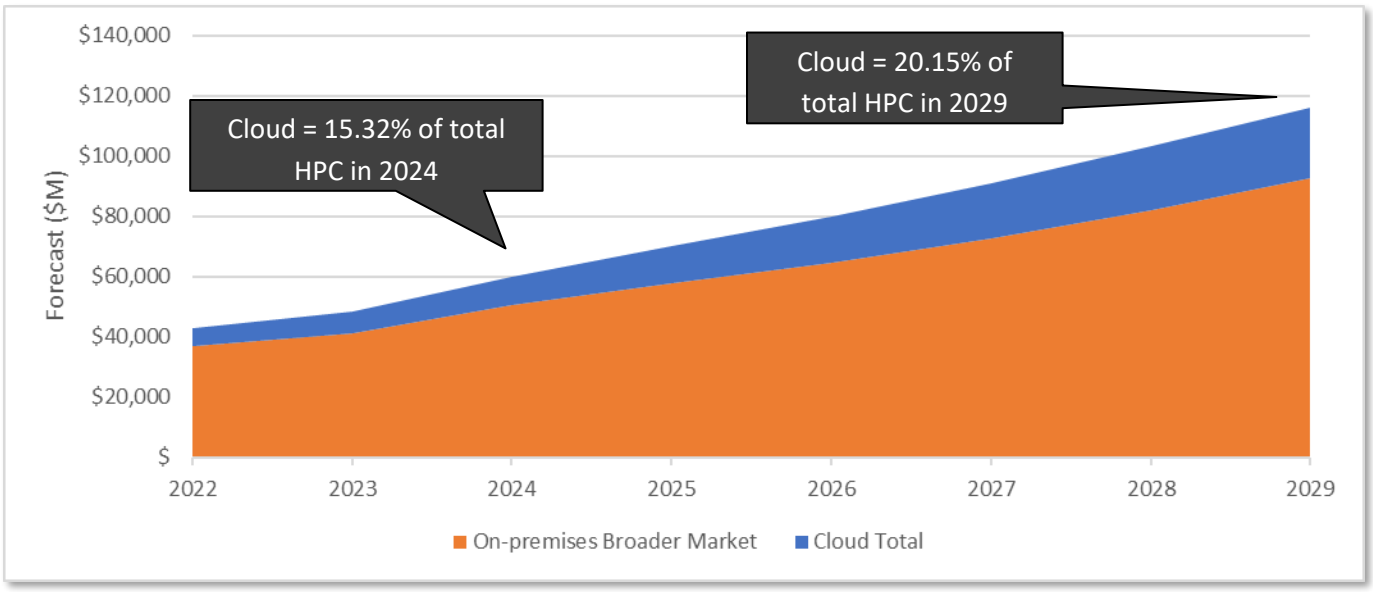
HPC-AI Cloud Spending Forecast Comparison



Source: Hyperion Research, 2025

FIGURE 2

Total HPC-AI Market Forecast



Source: Hyperion Research, 2025

FORECAST OBSERVATIONS

Methodology

Hyperion Research defines the HPC-AI Cloud Service Provider Market as end-users spending across all facets (e.g., compute, storage, networking, software) to run scientific and engineering workloads in the cloud. This contrasts with what CSPs are spending to provision HPC-AI resources to users. The latter is difficult to project as users may run their HPC-AI workloads on what might be considered non-HPC-AI infrastructure, and HPC-AI infrastructure may be used to run non-HPC-AI user applications.

Growth Drivers

Several factors are driving the continued growth of HPC-AI spending in the cloud:

- Access to latest technologies and software. With vendors accelerating the roadmap cadences to a 12-year cycle, on-premises data centers are becoming less able to provide their users with the most advanced computational capabilities. Data centers are adapting to integrating cloud access capabilities into the HPC-AI workflow processes they support. New AI software, models and tools are key contributors to the growth.
- Resource costs. The increasing price of the most advanced computational capabilities, particularly GPUs, requires data centers to provide an increasing portion of the HPC-AI CAPEX budget for them. If those expensive resources aren't fully utilized, they may sit idle. There is a utilization threshold that differs across institutions where the cost of the expensive resources may be less expensive to access in the cloud on-demand, as opposed to expending a large CAPEX investment for similar on-premises resources.

- **Bursting capabilities**. As organizations generally plan on-premises resource capabilities for "typical" utilization expectations, there are times when user applications and workloads exceed what on-premises resources can deliver. Cloud-based resources are required to meet this cyclical demand.
- **Running large-scale jobs**. In addition to bursting needs, there are times when the on-premises resources cannot meet the demands of the workload (e.g., large-scale AI training jobs). The scale of resources afforded by the cloud can meet these oftentimes short-term needs.
- **Evolution of cloud-based tools.** The cloud service providers (CSPs), ranging from the predominant full-service public CSPs (e.g., AWS, Google, Microsoft) to the "neoclouds" focusing on AIaaS and GPUaaS capabilities have continued to evolve their tools to both ease migrating on-premises workloads to the cloud and developing cloud-native applications.

ADDITIONAL HPC-AI CLOUD FORECAST PERSPECTIVES

It is also instructive to examine the forecast from the perspective of industry verticals, geographic region, and sector.

HPC-AI Forecast by Industry Vertical

The Biosciences and CAE verticals are projected to each surpass $4B in 2029 while the Economics/Financial, Government Labs and Academia verticals are projected to grow at a 5-year CAGR exceeding 23%.

Table 1 provides the HPC-AI industry vertical cloud spending forecast projections.

TABLE 1

HPC-AI Cloud Spending by Vertical ($M)

	2024	2025	2026	2027	2028	2029	CAGR 24-29
Bio-Sciences	$2,274	$2,730	$3,210	$3,791	$4,239	$4,510	14.7%
CAE	$2,025	$2,640	$3,321	$3,877	$4,318	$4,608	17.9%
Chemical Engineering	$239	$311	$389	$345	$400	$442	13.1%
DCC & Distribution	$539	$698	$874	$944	$1,095	$1,210	17.6%
Economics/Financial	$662	$877	$1,122	$1,316	$1,643	$1,955	24.2%
EDA	$675	$891	$1,133	$1,344	$1,580	$1,746	20.9%
Geosciences	$656	$879	$1,133	$1,362	$1,590	$1,757	21.8%
Mechanical Design	$39	$48	$57	$58	$67	$74	13.7%
Defense	$736	$958	$1,203	$1,408	$1,633	$1,804	19.6%
Government Lab	$744	$1,027	$1,319	$1,561	$1,864	$2,118	23.3%

TABLE 1

HPC-AI Cloud Spending by Vertical ($M)

	2024	2025	2026	2027	2028	2029	CAGR 24-29
University/Academic	$389	$557	$760	$980	$1,264	$1,490	30.8%
Weather	$305	$402	$512	$608	$716	$838	22.4%
Other	$256	$359	$485	$563	$653	$721	23.1%
Total	$9,540	$12,376	$15,519	$18,157	$21,062	$23,274	19.5%

Source: Hyperion Research, 2025

HPC-AI Cloud Forecast by Region

North America is projected to grow annually at more than 22% to over $10B in 2029. The following table summarizes the HPC-AI geographic region forecast.

TABLE 2

HPC-AI Cloud Spending Forecast by Region ($M)

Region	2024	2025	2026	2027	2028	2029	CAGR 24-29
North America	$3,759	$5,025	$6,487	$7,717	$9,141	$10,333	22.4%
EMEA	$2,534	$3,222	$3,957	$4,537	$5,244	$5,795	18.0%
APAC	$2,849	$3,602	$4,407	$5,113	$5,750	$6,121	16.5%
ROW	$398	$528	$667	$790	$926	$1,024	20.8%
Total	$9,540	$12,376	$15,519	$18,157	$21,062	$23,274	19.5%

Source: Hyperion Research, 2025

HPC-AI Cloud Forecast by Sector

The industry sector is projected to grow to approximately $15.5B in 2029 while the academic sector is forecast to exhibit the highest CAGR over the forecast period at 30.8%.

The following table provides the HPC-AI forecast by sector details.

TABLE 3

HPC-AI Cloud Spending Forecast by Sector ($M)

	2024	2025	2026	2027	2028	2029	CAGR 24-29
Industry	$6,625	$8,526	$10,640	$12,324	$14,165	$15,519	18.6%
Government	$2,526	$3,293	$4,118	$4,852	$5,633	$6,265	19.9%
Academia	$389	$557	$760	$980	$1,264	$1,490	30.8%
Total	$9,540	$12,376	$15,519	$18,157	$21,062	$23,274	19.5%

Source: Hyperion Research, 2025

FUTURE OUTLOOK

Conversations regarding the use of cloud computing resources have decidedly evolved from "if" cloud resources are appropriate for large-scale scientific, engineering, and AI-related computing to "when" they are appropriate and "how" to best leverage them. The notion of "continuum computing" which encompasses edge, cloud, and on-premises computing infrastructure and the methods and tools required by researchers and scientists to conduct their research dynamically and transparently across any advanced technical computing environment is becoming rooted in the HPC-AI resource planning models for many institutions.

Coupled with anticipated advancements by CSPs and the emergence of new applications, LLMs, workloads and increased adoption of AI into a wide range of workflows, HPC-AI spending growth in the cloud is expected to continue at a strong rate and may even reach a tipping point, causing upward growth revisions in future forecasts.

About Hyperion Research, LLC

Hyperion Research provides data-driven research, analysis and recommendations for technologies, applications, and markets in high performance computing and emerging technology areas to help organizations worldwide make effective decisions and seize growth opportunities. Research includes market sizing and forecasting, share tracking, segmentation, technology and related trend analysis, and both user & vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). Hyperion Research provides thought leadership and practical guidance for users, vendors and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue
St. Paul, MN 55102
USA
612.812.5798
www.HyperionResearch.com and www.hpcuserforum.com
