

Forecast Document

Worldwide HPC in the Cloud Forecast, 2020-2026

Alex Norton, Mark Nossokoff, and Earl Joseph
June 2022

HYPERION RESEARCH OPINION

The cloud market for HPC has undergone a fundamental shift in the past two years. Historically, user sites treated cloud resources primarily as additional to on-premises HPC systems to address peaks in workload demand. Generally, funds for HPC cloud resources were separate from the HPC system budgets and were not commonly critical components in future system procurement planning. Over the past two years, that paradigm has changed dramatically. Hyperion Research has conducted a number of studies that suggest cloud is becoming a major component in future resource planning for many HPC user sites. Many sites are now weighing cloud next to on-premises procurement strategies and altering the size and timing of future on-premises deployments to increase their budget for cloud resources.

These HPC sites are adding cloud spending to their mix in order to provide their user base the most effective compute resources for HPC jobs, and to better optimize their budgets, skillsets, workflow, and time constraints. With any transition, some sites will be trailblazers, shifting above average portions of their HPC workflow to cloud platforms. Others will tread more lightly, starting small to understand how best to take advantage of what the cloud offers. The Hyperion Research five-year cloud forecast contained in this report highlights that the market for HPC spending in the cloud will outpace the on-premises market (17.6% compared to 6.9% five-year CAGR, respectively), yet will remain a smaller portion of the overall HPC market over the next five years.

As cloud providers and on-premises vendors alike work to provide tools and capabilities to ease access to cloud resources, Hyperion Research expects cloud adoption to grow but not take over the on-premises HPC market. The HPC on-premises market is not projected to dramatically change from the steady 6%-8% five-year CAGR seen in years past, but the use of HPC cloud resources is causing a fundamental shift that will alter the future of the HPC market overall in the coming decades.

KEY FORECAST DEFINITIONS

This fifth iteration of the Hyperion Research cloud forecast covers HPC end user spending on third-party cloud resources and does not include the spending on HPC systems by cloud service providers (CSPs). In addition, this forecast includes only HPC end users who also have (or had) on-premises HPC systems. It does not include HPC end users who have only ever run workloads in the cloud.

This forecast doesn't include a new set of buyers who recently started doing HPC work exclusively in the cloud. This "cloud born" or "cloud native" group of end users may be significant and impact the overall market size. These users are not currently included in our forecast numbers for these reasons:

- Many of these users may not know that they are running HPC jobs in the cloud because of their unfamiliarity with what constitutes HPC workloads, based on the definitions Hyperion Research uses to categorize workloads and users.
- These users are difficult to identify because CSPs do not track information about specific workloads running on their platform.
- Recent research suggests that many of the cloud born users are emergent SMBs that require HPC but have workflow and skillset characteristics that line up well with public cloud usage.

Hyperion Research is working to quantify this group of end users to include in future forecasts.

Note: this forecast covers only spending on third-party (CSP) cloud resources for public, private and hybrid cloud configurations and does not include on-premises private clouds.

2022 HPC CLOUD FORECAST

Given the shift in buying behavior for the cloud, Hyperion Research expects the cloud market for HPC to grow at a pace more than twice that of the on-premises server market. Projecting out to 2026, the cloud market for HPC workloads is expected to approach \$11.5 billion USD, growing to nearly half the size of the on-premises HPC server market. While each segment in the HPC ecosystem is adopting cloud resources, a few key sectors and application areas are fueling this high growth:

- Artificial Intelligence and other data-intensive workloads are being run at a higher rate in the cloud due to the wider availability of GPUs (for sites without large GPU installations), the access to public data sets, and the expertise of CSPs for AI workloads.
- Manufacturing and the FSI (Economics and Financial) sectors both are expected to outpace the overall cloud market in growth over the forecast period.
- The government sector, specifically government labs, are expected to aggressively increase their cloud usage over the forecast period as they work to understand how and where the cloud can optimize their compute infrastructure.
- Sites that have smaller primary HPC systems, those in the workgroup or departmental segment, have shown strong adoption of cloud resources for HPC jobs in the past few years.

As previously noted, the HPC cloud forecast represents what users are spending to run their HPC workloads in the cloud. This encompasses all HPC resources (e.g., compute, storage, networking, file systems, application licenses).

The table below shows 2020 and 2021 actual revenues alongside the full cloud forecast from 2022 to 2026, highlighted by the cloud market breaking the \$10 billion USD in a year by 2026. For 2022, Hyperion Research expects the market to grow 23% YOY, reaching \$6.3 billion USD. The five-year CAGR is 17.6%.

TABLE 1

HPC Cloud Forecast 2020-2026

(\$M USD)	2020	2021	2022	2023	2024	2025	2026	CAGR 2021-2026
HPC Cloud Forecast	4,300	5,100	6,304	7,369	8,511	9,873	11,453	17.6%

Source: Hyperion Research, 2022

HPC Cloud Forecast by Sector

Today and for the near future, the industrial sector of the HPC ecosystem represents the largest portion of the cloud ecosystem, representing nearly 2/3 of all revenue generated by end user spending on cloud resources for HPC. However, the fastest growing sector is the government sector, which is expected to more than double their annual spend on cloud by 2024.

The slowest growing sector is the academic sector. Although the academic sector spends the least on cloud resources, they are still using the cloud heavily. Many academic sites have access to free or donated cycles for research, as well as grants and other funding avenues for running HPC workloads in the cloud.

TABLE 2

HPC Cloud Forecast by Sector

(\$M USD)	2020	2021	2022	2023	2024	2025	2026	CAGR 2021-2026
Academia	215	242	284	316	347	382	421	11.7%
Government	1,093	1,337	1,641	1,929	2,246	2,625	3,242	19.4%
Industry	2,992	3,521	4,378	5,124	5,919	6,866	7,790	17.2%
Total	4,300	5,100	6,304	7,369	8,511	9,873	11,453	17.6%

Source: Hyperion Research, 2022

HPC Cloud Forecast by Vertical

Digging deeper into the forecast, the table below expands the HPC cloud forecast into the 13 verticals tracked by Hyperion Research in all forecast areas. By 2026 both Bio-Sciences and CAE (the manufacturing segment) are expected to spend roughly \$2.4 billion USD on cloud resources for HPC workloads, making them the two largest segments in the HPC cloud sector. Both verticals were early adopters of HPC in the cloud and are expected to continue their strong adoption of HPC in the cloud. Other segments projected to increase utilization of HPC cloud resources include Weather, Government Lab, FSI (Economics/Financial), and Geosciences, which are all expected to grow at rates higher than the overall market.

TABLE 3

HPC Cloud Forecast by Vertical 2020-2026

(\$M USD)	2020	2021	2022	2023	2024	2025	2026	CAGR 2021-2026
Bio-Sciences	1,298	1,439	1,600	1,812	2,036	2,290	2,337	10.2%
CAE	795	957	1,325	1,556	1,807	2,106	2,456	20.7%
Chemical Engineering	108	128	158	185	213	248	287	17.6%
DCC & Distribution	244	289	357	417	481	557	645	17.4%
Economics/Financial	248	315	409	502	591	700	828	21.3%
EDA	317	372	455	527	602	692	794	16.4%
Geosciences	270	328	414	495	585	694	824	20.3%
Mechanical Design	21	24	28	32	35	38	42	11.7%
Defense	327	391	484	567	657	764	888	17.8%
Government Lab	304	364	455	537	626	734	1,088	24.5%
University/Academic	215	242	284	316	347	382	421	11.7%
Weather	65	140	189	241	303	383	485	28.1%
Other	88	111	145	183	228	286	358	26.5%
Total	4,300	5,100	6,304	7,369	8,511	9,873	11,453	17.6%

Source: Hyperion Research, 2022

HPC Cloud Forecast by Geographic Region

Geographically, North America HPC users spend the most on cloud resources for HPC workloads today and over the forecast period. However, the HPC cloud growth rate of APAC, which includes Japan here, is expected to outpace North America and EMEA, reaching more than \$3.5 billion USD in 2026. The table below summarizes the global geographic regional HPC cloud spending forecast.

TABLE 4

HPC Cloud Forecast by Geographic Region

(\$M USD)	2020	2021	2022	2023	2024	2025	2026	CAGR 2021-2026
North America	1,733	2,000	2,404	2,766	3,167	3,713	4,405	17.1%
EMEA	1,098	1,328	1,658	1,958	2,261	2,570	2,862	16.6%
APAC	1,321	1,584	1,979	2,337	2,728	3,169	3,688	18.4%
ROW	148	188	263	308	355	421	498	21.5%
Total	4,300	5,100	6,304	7,369	8,511	9,873	11,453	17.6%

Source: Hyperion Research, 2022

FUTURE OUTLOOK

The HPC cloud market is undergoing a fundamental shift right now, with users changing the way they view and acquire cloud resources. For many sites cloud resources have become a key component in future roadmaps of resource procurements. Clouds are used to address peaks in workload demand and to handle a subset of workloads that are more useful to be run in the cloud than on-premises. Given this shift, the HPC cloud market is expected to more than double in per year market revenue by 2026, reaching nearly \$11.5 billion USD.

As user sites continue to expand their utilization of cloud resources, it will become more critical for the managers to understand how and where the cloud can work to address their compute needs. The cloud may not be the optimal solution for every workload, and some workloads can and should remain in on-premises systems. Providers and vendors alike should focus on ease of use and management tools to enable HPC users to take full advantage of both the cloud and their on-premises deployments, such as smarter job allocators and simple user interfaces. Additionally, the continued support of key applications and better transparency in cost estimates for the cloud will provide solutions that may entice users to increase their cloud utilization and understand more deeply how to incorporate cloud effectively into their compute resource pool.

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

Copyright Notice

Copyright 2022 Hyperion Research LLC. Reproduction is forbidden unless authorized. All rights reserved. Visit www.HyperionResearch.com to learn more. Please contact 612.812.5798 and/or email info@hyperionres.com for information on reprints, additional copies, web rights, or quoting permission.