

Special Analysis

2022 HPC End Users Perspectives on Vertical/Application Workload Areas and HPC System Software and Middleware

Melissa Riddle, Jaclyn Ludema, and Earl Joseph
March 2023

HYPERION RESEARCH OPINION

Key findings from a recent Hyperion Research study showed that accelerated applications are growing as percent of workload, especially in the industry sector. In addition, many HPC users reported that they have or plan to migrate away from CentOS in response to the changing support model. Insights into the critical factors driving these and other trends are detailed in the 2022 iteration of Hyperion Research's annual MCS end users' study, *2022 HPC Multi-Client Study: Vertical/Application Workload Areas and HPC System Software and Middleware*. Key findings from the report are summarized in this document.

Hyperion Research conducts an annual Multi-Client Study (MCS) to measure and track key trends across the spectrum of the HPC market. The latest iteration of the MCS encompassed 181 HPC end-user sites with 3,830 HPC systems. Reports produced as result of the study span:

- AI and HPDA Usage and Future Technology Trends
- Vertical/Application Workload Areas and Technical Computing System Software and Middleware
- Use of Public/External Clouds for HPC Workloads, Trends, and Drivers
- Processors, Coprocessors/Accelerators, and HPC Budgets
- Trends and Forecasts in HPC Storage and Interconnects

The overall HPC application software sector is projected to grow at a moderate pace from \$5.0 billion in 2021 to \$6.3 billion in 2026, a 5.0% CAGR. The HPC middleware sector is expected to grow slightly faster from \$1.7 billion in 2021 to \$2.3 billion in 2025, a 5.6% CAGR. Most of the vertical/application areas are expected to show strong growth over the next 5 years, resulting in 9 application segments worth over \$1 billion each in 2026. Within that environment, the key trends identified from this HPC end-user base survey include:

- Use of accelerated applications continues to grow as GPU use becomes ubiquitous.
- CentOS usage has declined since last year following the recent business/support model changes, while Ubuntu Linux has risen in popularity as a free alternative.
- Average HPC job sizes are increasing, with about half of applications among study respondents running on multiple nodes.

SELECT KEY FINDINGS

Select key findings in this report include migration away from CentOS following support model changes, rising use of accelerated applications, and increased use of several critical, large-scale applications.

CentOS Usage Declines; Ubuntu Linux Rises

Following Red Hat's change to the CentOS business and support model, Ubuntu Linux has risen in popularity as a free alternative while CentOS usage declines. Today, a quarter of users (26%) consider Ubuntu Linux their primary OS, nearly as many users as historical leader Red Hat Linux (28%).

Most HPC Sites Use C/C++ and Python

C/C++ and Python were the most popular programming languages and models in the study, both used by the majority of survey respondents (85% each). CUDA, MATLAB, and Java were also popular and used by about half of respondents each (51%, 48%, and 48%, respectively).

Slurm Remains the Top Resource Manager

Slurm maintained its lead as the most popular job queuing, resource manager, or scheduling software represented in the study with half of all respondents (50%).

HPC Sites Report Several Critical, Large-Scale Applications

Respondents reported that about half of all their HPC applications (47%) run on multiple nodes. Furthermore, when asked to characterize the top two applications per site, two-thirds (69%) were run on multiple nodes. This suggests that the typical HPC site has at least several critical, large-scale applications, and also many single node jobs.

Accelerated Application Use Rises, Particularly in Industry

Accelerator use has become ubiquitous among HPC sites. Most sites (75%) reported that their #1 most used/important application is accelerated at least some of the time. On average, each HPC site reported 7.5 critical or production applications that use accelerators. Overall, about a third (32%) of HPC cycles are spent on accelerated applications. Industry users reported the highest accelerator usage rates.

In-House and Open Source Applications Comprise the Majority of HPC Workloads

Respondents reported that purchased ISV software represented only about a fifth (19%) of all HPC jobs. Industry sites devoted nearly half of their run time (46%) to in-house applications while academia favored open source (50%). Compared to the overall workload distribution, the top applications were more likely to be in-house (52.9%) and less likely to be open source (29.1%). This reflects HPC sites' increased investments of time and effort into their top applications as compared to their HPC workloads as a whole.

Top Application Run Times are Highly Variable

When asked to give the typical run times of their top two HPC applications, answers were highly variable. The mean was 144 hours while the median was 7.5 hours. The most popular response (21%) was between 10 and 25 hours and the vast majority of applications (87%) had run times of 100 hours

or less. However, there were a handful of applications (4%) with extremely lengthy run times exceeding 2,000 hours.

AI, ML, DL, and HPDA Frameworks are Widely Used

The most popular AI, ML, DL, and data intensive frameworks were TensorFlow, PyTorch, Jupyter Notebook, SQL, CUDA, Hadoop, and Spark with each cited by about half of respondents or more. On average, each respondent selected 3.8 AI/ML/DL frameworks and 2.4 data intensive frameworks. This highlights not only the rise of AI-HPDA but also the spread of their related software applications and packages.

Most Applications are Still Exclusively Run On-Premises, Although Cloud Usage is Rising

Respondents reported that the majority of all applications (65%) are run exclusively on-premises, a decrease compared to last year when three-fourths of applications (74%) were reported exclusively on-premises. Meanwhile, both hybrid and exclusive cloud usage has increased. When respondents characterized the top two HPC applications at their site, most of these applications (70%) were run exclusively on-premises as well.

HPC Users are Optimistic about Budgets; Prioritize Hardware and Storage

Most surveyed users (63%) expect HPC budgets to increase over the next year, including half of users (52%) who expect their budgets to increase by at least 5%. The highest budget increases were expected in HPC hardware, add-on storage, and public cloud. On average, application software is expected to increase 3.9% while middleware is expected to increase 3.1%.

In addition, the vast majority of users (90%) reported a willingness to pay a 10-15% premium on system price for their desired system attributes.

- The most desirable attributes included better processors, larger and faster memory, higher performance external I/O and storage, and better density, power, and cooling.
- A fifth (21%) were willing to pay a premium for better system management or monitoring software and a tenth (11%) were willing to pay for integrated middleware providing ease of use and ease of development.

Many HPC Users Plan to Use Edge Computing

A quarter of HPC users (28%) either currently employ edge computing or expect to within two years. Top motivators include improving real-time data collection and processing, accelerating HPC applications, access to IoT devices for data collection, and a wider range of sensor data.

AI Popularity Continues to Rise, Especially in Industry

Virtually all users (94%) reported plans to use AI methodologies. In industry, AI was about equal to modeling and simulation as a percent of HPC workload (31% vs. 34%). Furthermore, AI-specific software licenses such as Databricks and Anaconda are almost exclusively being purchased by industry sites, with a fifth of industry participants (21%) reporting that they currently have such a license.

FUTURE OUTLOOK

The short-term expectation is that 2022 is the beginning of a multi-year recovery as the pandemic's impact subsides in the face of increasingly effective countermeasures that improve health, economies, and vendor productivity. Based on representations Hyperion Research obtained from HPC buyers and vendors, the pandemic's primary impacts to date on the HPC server market have been reducing the production and global distribution of parts, along with eliminating most conferences and other in-person sales opportunities needed to fill pipelines.

The overall broader HPC on-premises market is expected to grow to reach \$40 billion by 2026, with HPC servers representing \$20 billion, add-on storage \$9 billion, middleware \$2 billion, applications software \$6 billion and repair services \$2 billion. The five-year CAGR for the overall market is 6.4% with HPC software at 5.0% CAGR and HPC middleware at 5.6% CAGR. By 2026, the HPC middleware market is expected to be approximately the same size as the HPC service market (\$2.3 billion each).

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 2023 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.