Intersect360 Research White Paper: BRIGHT COMPUTING: MANAGING MULTIPLE PATHS TO INNOVATION

EXECUTIVE SUMMARY

High Performance Computing isn't one thing; it is many things to many people. Over the course of a few years, many organizations began formal programs of data science, mining new and existing data for insights that could translate into top-line knowledge or competitive advantage. Now Artificial intelligence (AI)—or more specifically, machine learning—is the newest workload taking off across the HPC landscape. Intersect360 Research studies have found seventy percent of HPC users are already running machine learning or are working to implement it within the next year; another twenty percent are considering it.

On the technology side, differentiation is on a comeback, and there is a wide range of choice in hardware, software, and even consumption. New processing elements (Intel, AMD, ARM, GPUs, FPGAs, and more), containers (including Kubernetes, Docker, and Singularity), and cloud (mostly hybrid and multi-cloud) have disrupted the HPC industry. Matching diversifying technologies to diversifying workloads is the greatest challenge facing the HPC industry. There is an increasing need for a professional, supported cluster management tool that spans all these dimensions of the new HPC.

That's where Bright Computing comes in. With system management software spanning HPC, analytics, and AI, Bright Computing is filling a necessary role in high-performance segments. The company's core product, Bright Cluster Manager, is the most-cited commercial system management package among surveyed HPC users. Bright Cluster Manager sits across an organization's HPC resources—whether on-premise, in the cloud, or at the edge—and organizes them across workloads.

Fundamentally, Bright Computing helps address the big question in HPC: how to match diverse resources to diverse workloads in a way that is both efficient today and future-proof for tomorrow.

Data science and machine learning? Intel or AMD? GPUs or FPGAs? Docker or Kubernetes? Cloud, on-premise, or edge? AWS or Azure? Bright Cluster Manager lets users decide individually how to incorporate all of these transitions—some or all, mix and match, now or later—in a single HPC cluster environment. With so many independent trends continuing to push HPC forward, Bright Computing is aiming to be the company that helps users pull them all together.

P.O. Box 60296 | Sunnyvale, CA 94088 | Tel. [888] 256-0124
www.Intersect360.com | info@Intersect360.com

MARKET DYNAMICS

Which HPC is Right for You?

By its nature, High Performance Computing (HPC) is transformational. Science begets engineering, ideas beget innovations, and organizations relentlessly pursue the horizon of possibilities. HPC is a dynamic, long-term stable market precisely because we never run out of problems to solve. New discoveries only push back the boundary of what is possible, and there is always more to explore.

This is evident across the HPC industry in terms of both the scope of applications and the types of technologies that are deployed to run them. With the former, traditional scientific and technical computing has been augmented by data analytics and the recent evolution of artificial intelligence, changing what HPC is all about. With the latter, there has been a proliferation of technology elements, operating environments, and delivery mechanisms, changing how HPC is run.

The New HPC: Workloads

"HPC" isn't one thing; it is many things to many people. Intersect360 Research defines HPC as "the use of servers, clusters, and supercomputers—plus associated software, tools, components, storage, and services—for scientific, engineering, or analytical tasks that are particularly intensive in computation, memory usage, or data management." It is important across a wide range of organizations, as follows:[1]

> *HPC is used by scientists and engineers both in research and in production across industry, government, and academia. Within industry, HPC can frequently be distinguished from general business computing in that companies generally will use HPC applications to gain advantage in their core endeavors—e.g., finding oil, designing automobile parts, or protecting clients' investments—as opposed to non-core endeavors such as payroll management or resource planning.*
>
> *Characteristics of HPC applications include, but are not limited to:*
>
> - *Requirements for leading-edge system performance, or ability to address the most demanding problems;*
> - *Requirements for leading-edge scalability;*
> - *Tendency to incorporate, test, and perfect new technologies and methodologies associated with market creation and expansion*

Intersect360 Research relies on this definition because it focuses on use case over specific configurations, and because of its ability to incorporate new types of workloads as the market evolves. Consider the evolution of analytics throughout the big data revolution of the past

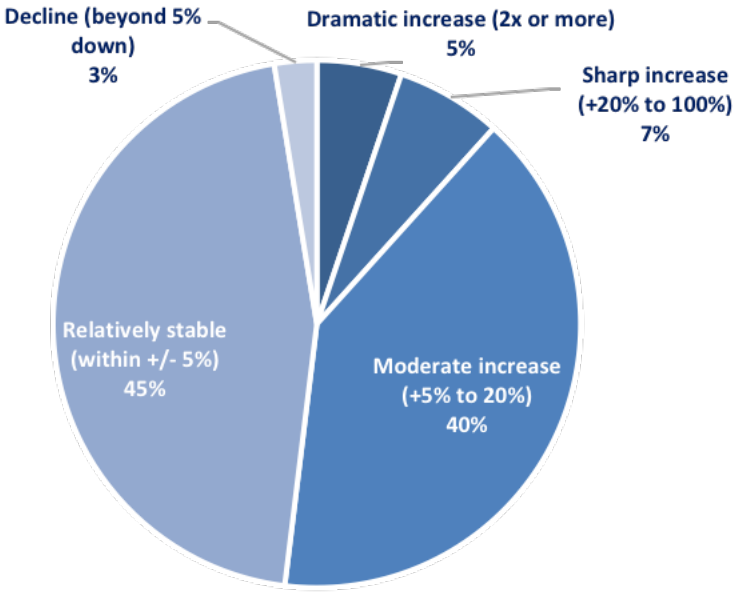
Over the course of a few years, many organizations began formal programs of data science. Today, AI is the newest workload taking off across the HPC landscape.

[1] https://www.intersect360.com/what-is-hpc.

decade. Over the course of a few years, many organizations began formal programs of data science, mining new and existing data for insights that could translate into top-line knowledge or competitive advantage. Repeated studies from Intersect360 Research showed that this did not have a large effect on high-level budgets—organizations tended to rely on infrastructure or personnel that was already in-place or planned; new software was usually free or else a new module of an established enterprise package. Where dimensions of performance or scalability came into play, companies relied on their HPC systems if they had them, or they turned to cloud computing when they did not. HPC largely absorbed this new dimension of scalable computing.

This case study remains relevant as another trend takes hold. Artificial intelligence (AI)—or more specifically, machine learning—is the newest workload taking off across the HPC landscape. Intersect360 Research studies have found seventy percent of HPC users are already running machine learning or are working to implement it within the next year; another twenty percent are considering it. In this case, there is a top-level budgetary effect. The incorporation of machine learning workloads into HPC environments is often accompanied by a significant increase in budget. (See Figure.)

Effect on High-Performance Workload Budget Related to Incorporation of Machine Learning
Intersect360 Research, 2020



The net effect is that HPC users are now addressing a wider range of workloads than ever before. Traditional HPC—scientific and technical computing applications—are still going strong; science has not been solved to completion. Now analytics and machine learning,

which have their own demands, have been added to the perpetual to-do list. HPC users need more from their HPC environments.

The New HPC: Technologies

Against this backdrop of diversifying workloads and demands, HPC users have also seen a diversification of the types of technologies available to meet evolving needs. Time was, a "supercomputer" was a relatively self-contained device, a big piece of custom-designed iron specialized with its own proprietary software stack. The cluster revolution that began in the 1990s disaggregated the hardware into industry-standard servers, connected with industry-standard networks, running industry-standard operating systems. The technology was more accessible, but there was less to differentiate one system from another in terms of performance or competitive advantage.

Today differentiation is on a comeback. The technology stack is still disaggregated, but now there is a wide range of choice in hardware, software, and even consumption models. Some of the major technology trends that have disrupted the industry are:

- *New processing elements:* For more than ten years, Intel has been the dominant provider of CPUs for the servers that form HPC systems. While this is still the case in our most recent survey data, there are nevertheless multiple trends in play. AMD has made remarkable strides with its EPYC processors, which have the same x86 instruction set as Intel Xeon. There are also significant projects based on ARM processors, providing another option to x86 CPUs. Furthermore, many HPC systems are now augmented by co-processors or accelerators. For several years, the most common have been GPUs from NVIDIA, which now are part of over half of HPC systems. AMD is also a player here; its future Radeon GPUs will soon be available in an integrated package with AMD CPUs. FPGAs and custom processing elements for machine learning have also become feasible options.
- *Cloud:* The topic of cloud has been one of the hottest areas of inquiry for Intersect360 Research, and it now represents the largest growth rate among HPC product and services segments. About half of users now make some use of cloud for their HPC workloads,[2] and cloud computing is the fastest-growing revenue segment of the HPC market.[3] However, few of these organizations use cloud exclusively. Most are hybrid clouds, combining on-premise resources with public cloud infrastructure.
- *Containers:* With so many different types of HPC resources available, both on-premise and in the cloud, consistency has become an issue. Faster or specialized resources are of no use if applications do not run as expected. One option is to deploy containers, which are controlled, packaged, portable environments,

Differentiation is on a comeback. There is a wide range of choice in hardware, software, even consumption models.

Managing a diverse panoply of tools as a cohesive, efficient solution can be a tremendous challenge.

[2] Intersect360 Research, HPC User Budget Map data, 2020.

[3] Intersect360 Research HPC market forecast data, 2020.

including operating system, middleware, and application, that can be moved from one system to another, for improved reliability and reproducibility of results. Early, public-sector container deployments mostly used Docker, which is open-source; Kubernetes is common among commercial sites, and some HPC specialists prefer Singularity.

While all these choices provide the opportunity for tailored HPC suited to a wide range of workloads, in practice, managing such a diverse panoply of tools as a cohesive, efficient solution can be a tremendous challenge. HPC users want the latest technologies, but they also need to manage ongoing workflows, to keep costs from spiraling out of control, and to maintain a sense of investment protection. Matching diversifying technologies to diversifying workloads is the greatest challenge facing the HPC industry at large.

INTERSECT360 RESEARCH ANALYSIS

Bright Cluster Manager: Flexibility for the New HPC

The primary challenge in adapting to the new HPC—new workloads and new technologies—is a question of management. How can I manage disparate high-performance technologies, whether on-premise, in the cloud, or at the edge, in an efficient way? Clusters have always come with management tools, but mostly they are open-source point-products that were not developed to encapsulate such a wide array. There is an increasing need for a professional, supported cluster management tool that spans all these dimensions of the new HPC.

That's where Bright Computing comes in. With system management software spanning HPC, data analytics, and AI, Bright Computing is filling a necessary role in high-performance segments. The company's core product, Bright Cluster Manager, is the most-cited commercial system management package among surveyed HPC users.[4]

Bright Cluster Manager sits across an organization's HPC resources, spanning core, cloud, and edge, and organizes them across workloads. A choice of user interfaces—command-line or graphical—helps administrators monitor and optimize applications, both in terms of cost and time. Moreover, certain features of Bright Cluster Manager speak directly to the current trends in HPC.

Bright Cluster Manager supports mixed environments: mixed processing elements, mixed architectures, and mixed operating systems. Administrators can track and manage the utilization of accelerators, including both GPUs and FPGAs, monitoring which workloads make the best use of each environment. This helps support an environment in which each workload gets assigned to its optimal resource, shortening time-to-solution and reducing system overhead. When applications need to be moved between resources, Bright Cluster

With system management software spanning HPC, data analytics, and AI, Bright Computing is filling a necessary role in high-performance segments.

Bright Cluster Manager sits across an organization's HPC resources, spanning core, cloud, and edge, and organizes them across workloads.

[4] Intersect360 Research, *HPC User Site Census: Middleware and Developer Tools*, 2019.

Manager provides the underpinnings of Kubernetes, Docker, and Singularity containers, all managed through the same interface. (See image below.)

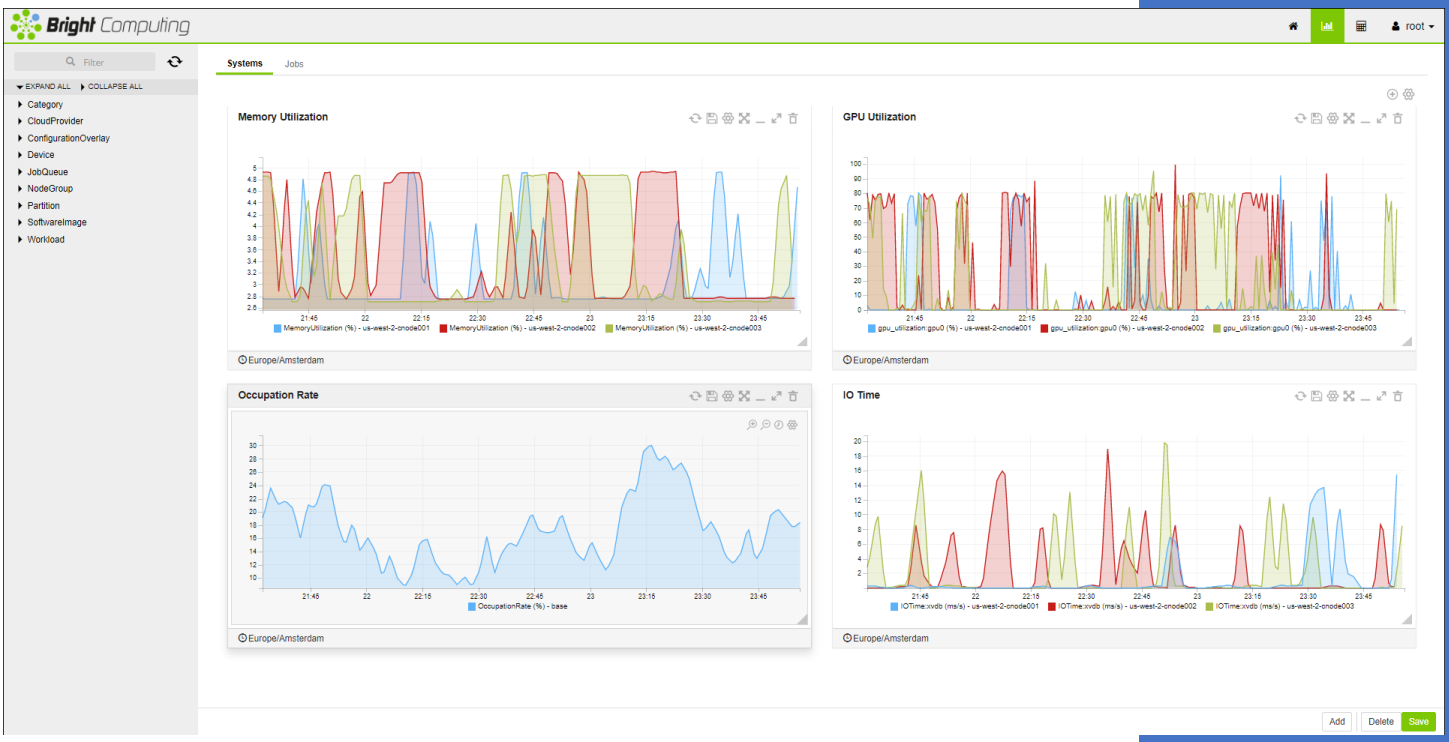


Image: Screenshot from Bright Cluster Manager (Bright Computing, 2020)

Most importantly, the purview spans the hybrid cloud, which is an increasingly common model across enterprise computing. Bright Cluster Manager can dynamically allocate instances from Amazon Web Services (AWS) or Microsoft Azure, or any public cloud running OpenStack tools. The public cloud instances provisioned by Bright Computing are imaged identically to the those running on-premise, so that on-premise workloads can be migrated without modification. By managing these resources through the same user interface as on-premise systems, administrators can handle spikes in demand or shifts in deadlines while still optimizing total cost.

Finding the Right HPC

Incorporating Bright Cluster Manager addresses the fundamental challenge of matching resources to workloads. When an organization has multiple types of HPC systems with different operating environments—for example, a technical computing system with an open-source job manager such as Slurm, and a separate cluster with Kubernetes containers primarily for analytics or machine learning—there can be a question of which system has jurisdiction or priority for new jobs, particularly when reaching for other resources like cloud

instances or edge computing data. Bright Cluster Manager sits atop all these platforms to organize resources across workloads.

Fundamentally, Bright Computing helps address the big question in HPC: how to match diverse resources to diverse workloads in a way that is both efficient today and future-proof for tomorrow. Data science and machine learning? Intel or AMD? GPUs or FPGAs? Docker or Kubernetes? Cloud or on-premise? AWS or Azure? Bright Cluster Manager lets users decide individually how to incorporate all of these transitions—some or all, mix and match, now or later—in a single HPC cluster environment.

With so many independent trends continuing to push HPC forward, Bright Computing is aiming to be the company that helps users pull them all together.

For more information about Bright Computing solutions for HPC, visit https://www.brightcomputing.com/product-offerings/bright-cluster-manager-for-hpc.

P.O. Box 60296 | Sunnyvale, CA 94088 | Tel. [888] 256-0124
www.Intersect360.com | info@Intersect360.com