Empowering Today's Data Scientist

January 2020

Written by: Linn Huang, Research Vice President, Devices & Displays

Introduction

According to IDC's Worldwide Global DataSphere Forecast, 2019–2023, the amount of data that is created, captured, replicated, and consumed globally is projected to grow at a 26% compound annual growth rate (CAGR) for 2018–2023 — from 33ZB to 103ZB. For the modern organization, the ability to effectively mine that datasphere for critical business insight could be the difference between building sustainable competitive advantages and facing a long march toward obsoletion.

Consequently, companies are increasingly hiring data scientists with aspirations of making smarter decisions over the long run. These data scientists must be equipped with top-of-the-line tools that can deliver real-time performance at the speed of business.

Enter workstations, a product category that has stood for performance and reliability. The workstation market has grown rapidly of late, largely driven by new data science workloads. From 2015 to 2019, workstations grew on average 10% a year, compared with -2% for the traditional PC market.

In a 2018 IDC survey of U.S. IT decision makers (ITDMs), 62% planned to expand their workstation installed base, citing artificial intelligence, machine learning, or other data science technology as the primary driver for that expansion.

In a 2019 IDC survey of U.S. ITDMs, 95% of respondents who cited data science as a primary workstation application intended to grow their workstation count in the coming years. The truth becomes more evident over time: Data scientists work best on workstations.
The Clear Benefits of Workstations

Workstations are born from a market made up of users who all have highly technical needs. These compute offerings are typically higher performance and optimized to meet technical workloads in rigorous conditions, with a high degree of configurability to meet the unique demands of each customer’s workload.

From a performance standpoint, workstations can take advantage of server-grade CPUs with high core-count Intel Xeon processors, as well as powerful graphics cards such as NVIDIA’s Quadro RTX 8000. IT buyers that find that level of hardware stack overserving of their needs can easily find an appropriate value mark given the configurability and deal-driven nature of this market. In short, workstations are built for workloads, not mass-produced for markets.

Workstations also provide unmatched reliability. Products go through thorough hardware certification to ensure they can meet the rigors of performance in trying conditions. Workstations are also software certified to ensure an optimized, consistent performance of select applications. These products are also typically backed by additional service and support layers.

Key Trends in Data Workloads

IT managers should consider two trends when thinking about data scientists and their workloads. First, cloud computing is a disrupter, but it has its limits with data science. The cloud provides companies significant benefits in scale and flexibility. It could streamline end users’ workloads, allowing for greater platform agnosticism and collaboration. However, cloud computing will always lag in terms of latency given where the compute is happening and the networks the data must traverse. As such, prudent IT managers will keep significant performance at the desks of their critical data scientists.

In a 2018 IDC survey about cloud and AI adoption, 80% of adopters reported repatriating workloads from the public cloud. Organizations are learning which workloads are best served by the cloud and which workloads are best run locally, such as real-time data capture and processing. By 2023, 57% of the aforementioned 103ZB will occur at the endpoint. The fastest-growing location of data creation will be at the edge (+33% 2018–2023 CAGR). Consequently, even as cloud computing changes the computing paradigm, data scientists will still do their best work on workstations.

Second, device choices are becoming an increasingly important part of the employee experience. In a 2018 IDC survey of U.S. workers, 63% said the technology they are provisioned impacts their productivity, and 62% said it impacts their satisfaction levels. The technology deployed has a direct impact on an organization’s output and ability to retain talent.

In a 2019 IDC survey of U.S. ITDMs, fleet modernization was rated a top 5 IT priority. Organizations are increasingly preparing to support a modern workforce with modern tools. In the same survey, more than three-quarters of respondents cited performance as a top 3 purchase criterion, with more than half ranking it first over price, brand, support, etc. If your organization relies on its data scientists and you want to keep them solvent, productive, and happy, workstations are the prudent choice.
**Considering Z by HP Workstations**

The HP ZBook 17, HP Z4, and HP Z8 Data Science Workstations are integrated solutions with the data scientist in mind. The Z8, which offers multi-GPU power, can support dual NVIDIA Quadro RTX 8000 cards for high performance. With NVIDIA RAPIDS pre-installed and compatibility with cuDF, cuML, cuGraph, and Dask, organizations can significantly shorten the time from deployment to when data scientists are at full speed — analyzing their data science workloads and providing business critical insights in real time. The Z portfolio offers a wide ecosystem and channel support.

Z by HP is also among the more highly trusted brands in enterprise computing. Before a Z product is launched, it undergoes 350,000 hours of lab testing, according to HP, to ensure optimal, consistent performance without degradation. Organizations at the frontier of data science should take note of this commitment to quality given that data software certification is undefined. This has allowed HP to be the market leader of desktop workstations since 2010, according to IDC research, demonstrating a proven track record.

Adjacently, the company also boasts a strong suite of security and software offerings, giving customers peace of mind in deployment. HP’s security stack ensures that a company’s confidential and sensitive data is protected at the endpoint, while HP Multi-Factor Authenticate protects the perimeter from unwanted intruders. For example, HP Sure Sense uses AI to help protect against cyberthreats both old and new, while HP Sure Run can keep scientists running even as their applications and processes are under malware attack.

HP also has a wide scope of other products to support data scientists beyond the box. Curved or 4K displays allow data-intensive workers to visualize more data more clearly than ever. HP also can support an organization’s printing needs for whenever digital images need to be transferred to paper or 3D printing for rapid prototyping. Z by HP’s collaborative solutions for the modern workforce include HP ZCentral Remote Boost (formerly HP Remote Graphics Software), which allows customers to access a workstation cluster remotely. The company also works with third-party software, such as OmniSci and NVIDIA, and curates a library of useful applications through its developer’s portal.

In short, HP can significantly improve the holistic work experience of data scientists.

**Challenges**

Z by HP's challenges within the workstation and data science space reflect the broader market’s challenges. Namely, the climb up the hardware stack presents a corresponding climb in up-front costs. The innovation that Z by HP puts into its workstations above commercial PCs in order to give organizations peace of mind for mission-critical scenarios costs those organizations a premium.

Yet, many organizations discover that overall total cost of ownership (TCO) favors workstations over the long haul. For starters, critical applications, firmware and drivers, and hardware are all designed for optimal performance. Additionally, HP invests heavily in building a suite of support services and a brand that communicates reliability. More efficient uptime and less downtime help ensure a better business outcome. More importantly, with the increase in data security and privacy concerns, many organizations are trying to keep their sensitive data on-premises under their control.

For customers who continue to struggle with whether the up-front premium sufficiently balances against long-term TCO, HP offers not only Device as a Service (DaaS) to help spread costs, reduce in-house oversight, and liberate IT resources but also financing programs for more budget flexibility.
**Conclusion**

Data is fast becoming today's business currency. Organizations that cannot manage and optimize data will face a fate similar to analog-only businesses. Tomorrow's titans will share the ability to cultivate rich business insight from an ocean of data. As a result, hiring and keeping top data scientists and providing them with the right tools for their job must be an imperative for any forward-thinking organization.

Ultimately, the "keeping" end of that bargain must revolve around a consistent commitment to the best tools. Balancing the need to keep data scientists at the forefront of innovation and the complexity of modern IT will be a struggle that's a common denominator for most companies. Z Data Science workstations provide data scientists and business leaders with the opportunity for power at the edge, reliability, and simplicity of deployment to get critical workloads running at real time.

**About the Analyst**

**Linn Huang, Research Vice President, Devices & Displays**

Linn Huang tracks market trends and industry developments that impact the worldwide and U.S. markets for PCs, thin clients, and monitors. He participates in cross-research streams that cover all device categories.
The Z by HP Data Science Solution, powered by NVIDIA® Quadro RTX™ GPUs, provides Data Science teams with unprecedented processing capabilities in an on-premises, single-box solution that combines powerful GPU + CPU compute with the NVIDIA GPU-accelerated data science AI software stack. From data wrangling to artificial intelligence to neural network implementations, Z is the affordable, future-ready solution with optional cloud connectivity to solve data challenges and enable Data Science teams to deliver the insights needed to move business forward. Learn more at www.hp.com/datascience.