Power Tools to Step Up to the Federal Data Challenges

MARKET TRENDS REPORT

govloop

Executive Summary

The world is generating vast quantities of data. The advent of the Internet of Things (IoT) and the spread of sensors for every kind of device imaginable – smart roads and buildings, self-driving vehicles, self-diagnosing machinery, ordinary household appliances – you name it, there is probably an IoT-enabled version of it.

The same applies to scientific data. It doesn't matter if it's tiny wind, water or soil sensors recording information in remote areas, or satellites taking detailed scans of the planet from space. It all contributes to increasingly large pools of raw data.

But having all that data does not equate to processing or understanding it. Nor does it mean that analysts, researchers, scientists, physicians and all the other professionals who could use it have the tools to do so.

Many government agencies rely on supercomputers to support intensive data processing, and cloud computing to store their oceans of data awaiting analysis. But supercomputers are far too costly to be everywhere, and the cloud can ramp up costs quickly. Plus, many federal workers are not working in their offices these days, limiting their access to computing resources.

An alternative to address this shortage of high-performance computing and capacity is the rapid adoption of high-performance workstation PCs.

To learn more about how this evolving class of solutions is filling vital requirements, GovLoop collaborated on this report with officials from Z by HP, which makes the company's line of high-end workstations. We discuss the key capabilities agencies should look for and best practices for putting workstations to work.

By The Numbers

75%

of enterprise-generated data will be created and processed outside a traditional data center or cloud by 2025.

60%

of government AI and data analytics investments will be intended to directly impact real-time operational decisions and outcomes by 2024.

41.6 billion

IoT devices will be connected to the internet by 2025, generating 79.4 zettabytes of data.

180+ zettabytes

the total amount of data forecast to be consumed globally by 2025.

$22.6 billion

the size of the global big data health care analytics market in 2019, which was projected to continue growing at a compounded annual growth rate of about 20%.

50%

of government organizations will establish formal accountability structures for data-sharing, including standards for data structure, quality and timeliness.

"Robust consumer PC demand again drove sales, particularly in regions where governments maintain stay-at-home orders as the COVID-19 pandemic persists. Prior to 2020, consumers had been shifting to a phone-first focus, yet the pandemic reversed this trend."

– Mikako Kitagawa, research director at Gartner

Data-Intense Work Drives New Requirements

The Challenge: Keeping Up With Data in Near Real Time

Federal agencies are wrestling with a fundamental problem with Big Data: The data is being generated at exponentially faster rates every year.

Tommy Gardner, CTO for HP Federal, said this is a key challenge facing data scientists. "You can restrict your data in some sense, or you can run multiple calculations on multiple parts of data, almost in parallel processing mode," he said. But you need to find some way to get your data processed.

Sending data to the cloud can alleviate the problem, but it is an expensive solution since it's bandwidth-intensive, and it erodes time sensitivity.

Some examples of huge data flows in government include:

- Climate readings such as temperature, humidity, precipitation amounts, wind speeds and cloud formation
- Satellite imagery for everything from weather forecasting to military activities and intelligence-gathering
- Generated results, such as applying algorithms to real-world information to create multiple hypothetical scenarios based on changing sets of assumptions

At the same time, agencies can use that data to meet their unique mission requirements if they can harness data science techniques and emerging technologies such as artificial intelligence and virtual reality. But they face the limitations of their legacy IT infrastructure and careful allocation of scarce resources, such as supercomputer availability.

The global pandemic has added another layer of complexity to these challenges by forcing many in the federal government to work remotely. They have not had seamless access to the computing resources integral to doing their jobs. IT teams throughout the government have spent the past year scrambling to find ways to securely deliver capabilities so agencies can continue to meet their mission requirements.

The government now plans to support an elevated level of remote working in the future, even once the pandemic subsides and workers can return to their offices – which means continued emphasis on distributed high-end computing capabilities.

The Solution: High-Performance Workstations

Agencies are finding that high-performance workstations offer a powerful alternative. While workstations can't handle the volume of data that a supercomputer can, high-end workstations can be used to separate data into much smaller subsets, process it and then combine it back into a single query, all from the comfort of one's office or home.

"The high-end workstation is probably one of the most valuable tools that HP makes for the world to do their jobs and to get things done correctly, accurately, systematically and safely," said Gardner. "And at the same time, they can be utilized to make decisions that in the past we just weren't able to make."

These new high-end workstations include such benefits as massive storage, which allows users to store large quantities of data for analysis; very high-end graphics capabilities suitable for image-intensive applications such as design, architecture and video editing; and significantly faster processing times through the use of high-end microprocessors and multiple cores. Such benefits translate into lower costs of operation by minimizing the use of cloud resources and improving workflow. And they can provide useful insights much faster than waiting for supercomputer cycles.

Depending on the user and application, this also opens up the prospect of shifting analysis to the point of collection. For instance, connecting sensors in the field to their workstations allows researchers to collect many different types of data for processing as it is being recorded, leading to timely analysis and the ability to spot trends in real time. Only high-end workstations can support the level of distributed, powerful computing needed for highly scientific applications.

"The edge is really where data is being collected," Gardner said. Moving processing closer to it is critical for the future.

Best Practices in High-End Workstations

Analytical Power

There are complexity gradations and latency factors, to name two, to be considered when assessing investment in high-performance desktops. For instance, smooth animation requires virtually instant data access to create images that flow. Digital video editing, by comparison, requires the same speed of data accessibility but also needs to handle much larger quantities of data. The scale of the data and the complexity of the algorithms to be run by data scientists, the time-sensitivity of the data and the importance of the question(s) to be answered may call for more memory and CPUs working in parallel.

Security

As a general rule, anything valuable enough to warrant high-performance computing capabilities has to be as secure as possible – protected from tampering, exfiltration, data corruption, external monitoring and other threats, including human error, such as phishing. Each workstation used for scientific purposes should have security-inherent features built into the machine, such as hardware-enforced protections and data recovery options.

For truly valuable or highly targeted information, the workstation should also have layers of third-party protection such as anti-malware programs, endpoint threat detection and possibly even encryption. Only a truly high-end workstation can support all of that while still providing enough power for its data processing tasks.

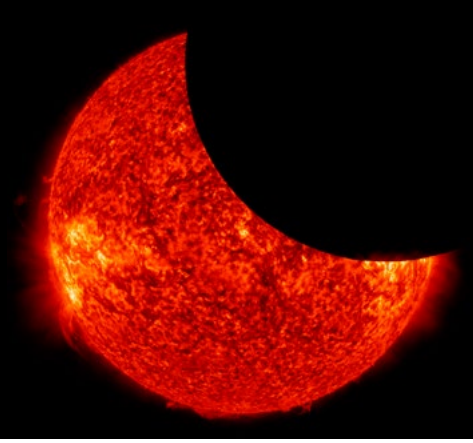
Cost-Effectiveness

This is another aspect of assessing the amount of analytical power required, as well as demonstrating the appropriateness of a high-performance workstation solution as opposed to other alternatives, such as multiple conventional CPUs or the capacity, capabilities and bandwidth costs and considerations of the cloud. Workstations should be fully supported and able to be fixed or replaced quickly in the event of any downtime.

Distributed Power

High-performance workstations in the office can be software-enabled to connect to at-home or mobile devices to provide access for data-intensive work. This also facilitates teams working with shared data. It's also possible to base a workstation at someone's home office to allow them to continue to process data outside of the office.

Case Study: Data-Crunching Delivers Vital Insights



The scientists at NASA's Solar Dynamics Observatory had a problem. The observatory captures images of the sun every 1.3 seconds; the images are used to gain insight into different types of solar variations and how they affect life on Earth. The information is a valuable asset, but more than 18 petabytes of images have been collected, making analysis a huge challenge.

In addition, researchers developed an algorithm that removes errors from the images, such as bad pixels, then archives them. The algorithm is highly accurate, but when there are so many petabytes of images, there still are billions of pixels that have been misclassified as errors.

Removing errors using conventional CPU processing would have taken several years, and supercomputer use at NASA is highly restricted. This didn't give the team a lot of good options.

They eventually decided to process their data using HP high-performance desktop workstations like the Z8 G4, powered by two Quadro RTX 8000 GPUs. The workstations allowed the team to analyze the images and remove all errors in less than a week by enabling the researchers to explore and iterate calculations.

Having the data rapidly accessible to the GPUs maintains responsive workflows, particularly for AI applications that often need fast access to the data. And remote tools allow team members to work from home, utilizing the power of their workstations while maintaining their efficiency. Thanks to quick data processing, the future is quite bright for projects using the information collected by the researchers working at NASA's Solar Dynamics Observatory.

HOW HP CAN HELP

HP has a long history providing state-of-the-art, reliable, high-performance desktop and laptop computers. The Z by HP line of workstations extends this tradition by providing several models of high-performance computers tailored to processing-intensive applications.

Every Z workstation is optimized for the performance best suited to a user's specific needs. And every workstation is fully configured and built with security from the ground up, making them the world's most secure. Additionally, tools like ZCentral Remote Boost empower users to connect their tablet, laptop or even a thin client to a powerful Z by HP workstation back in the office for intense processing.

All Z by HP workstations are fully ready for virtual reality applications that support the world's most demanding scientific and collaborative platforms. And if mobility is required, there is an entire line of portable workstations available, like the impressive ZBook Studio Mobile Workstation.

Learn more about Z by HP workstations at hp.com/data-science.

Conclusion

The internet-enabled digital world is far more complex than it was just a decade ago. Agencies are trying to find answers to challenges that are also getting more complex. How quickly can a vaccine be created for COVID-19? What is the impact of the 2021 wildfire season on climate change around the world? How can the nation's electrical grid be modernized to incorporate renewable resources?

These, and a host of other urgent questions, require analyzing all the information at hand to find feasible, creative solutions to pressing problems.

Providing data scientists, researchers, engineers and others with the analytical tools and capabilities needed to study these challenges and search for answers is critical. Providing the resources will always be an elusive goal – as Gardner said, "demand will always outstrip supply in this area" – yet this is the path to fundamental discoveries. High-end workstations can become a critical step in supporting that journey.

ABOUT HP

HP Inc. creates technology that makes life better for everyone, everywhere. Through our portfolio of printers, PCs, 3D Printing Technologies, mobile devices, solutions, and services, we engineer experiences that amaze and enable workflows that are seamless. Z by HP workstation solutions are designed to empower creators and power users around the world, from artists, to engineers and data scientists and more. More information about Z by HP is available at www.hp.com/z.

ABOUT GOVLOOP

GovLoop's mission is to "connect government to improve government." We aim to inspire public-sector professionals by serving as the knowledge network for government. GovLoop connects more than 300,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to connect and improve government.

For more information about this report, please reach out to info@govloop.com.

1152 15th St. NW Suite 800
Washington, DC 20005

P: (202) 407-7421 | F: (202) 407-7501

www.govloop.com
@GovLoop