



Embracing the Data-Driven Future

RESEARCH BRIEF



Introduction

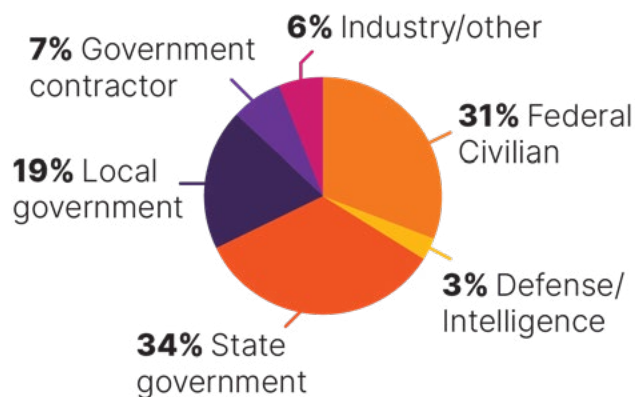
In the past decade, government agencies have experienced a massive influx of data, much of it unstructured. Understanding where all of that data resides, managing it and protecting it are gargantuan tasks, but if that data is safe, accounted for, accessible and managed efficiently, agencies can reap its value and use it for advanced processes such as high-level analytics and DevOps.

These challenges and benefits are causing some agencies to reevaluate their existing data management tools, technologies and processes. For many agencies, this reevaluation is coming at the right time, as they also aim to modernize IT infrastructure, move more workloads to the cloud and embrace cutting-edge technologies.

To learn more about how agencies can get ready for a data-centric future, GovLoop teamed with Pure Storage to survey 71 federal, state and local government workers about their agencies' data strategies.

The results point to what agencies want and need to manage, secure and make the most of their data. You'll also hear from experts at Pure Storage about what agencies can do to address challenges and embrace a data driven culture through a storage-as-a-service (SaaS) solution.

I work for...



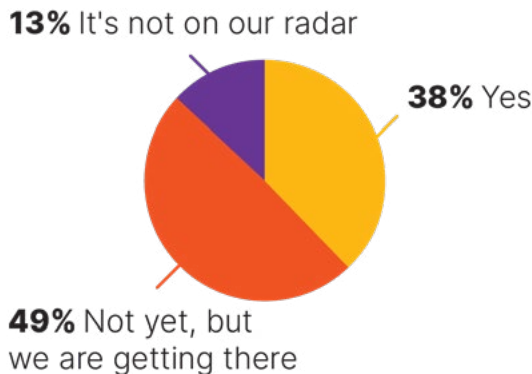
The Data-Driven Agency

Agencies are generating so much data, so fast, that it's overwhelming. Harnessed correctly, that data is invaluable for increasing efficiency and performance, better understanding and serving constituents, identifying fraud and security breaches, preparing for disasters, and much more.

Data is so important to the federal government that the nation's chief data officers formed their own group in early 2020. It focuses on implementing the Federal Data Strategy, which aims to build a culture that values data; provides guidance on governing, managing and protecting data; and determines the most efficient and appropriate data use.

These advances show clearly that federal agencies today know the value of data. In fact, the survey found that the vast majority of agencies either have a data-driven culture today or are making progress in that direction (see Figure 1).

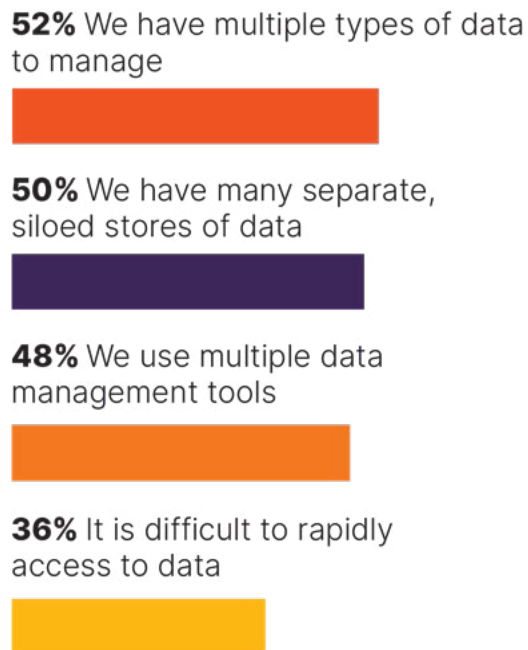
Figure 1: **According to Gartner, a data-driven culture is one that can examine and organize data with the goal of better serving users and speeding up decision-making. In your opinion, does your agency have a data-driven culture?**



“Every agency is a data-driven agency, whether they recognize it or not,” said Nick Psaki, Federal Principal Engineer at Pure Storage. “The key is finding a way to make sure the data is available, accessible and secure so agencies can get the most value from it.”

Ensuring that data is fully available and accessible can be a major barrier for many agencies in achieving a truly data-driven culture. Not only is data often spread among different on-premise environments and multiple clouds, but the expanding variety of data types — image data, data in legacy databases, sensor data and web data, for example — complicates the situation. According to the survey, more than half of agencies consider having multiple types of data to manage as a top challenge (see Figure 2).

Figure 2: **What are your biggest data management challenges? Choose up to three.**



“Every agency is a data-driven agency, whether they recognize it or not. The key is finding a way to make sure the data is available, accessible and secure so agencies can get the most value from it.”

- Nick Psaki, Federal Principal Engineer at Pure Storage

Silos happen as a function of how an organization grows over time. Although this is a natural occurrence, silos prevent agencies from embracing more modern, analytics-driven workflows, which require that data be available and accessible at all times. Eliminating those silos requires a different approach — one that centers on a data-centric architecture for storage and is designed for sharing data and the hybrid cloud world.

Embracing the cloud is also key to creating a data-driven culture. Storage optimized for hybrid cloud environments can help agencies improve performance and take advantage of greater scalability, flexibility and cost efficiencies. When considering cloud-based storage options, insist on high availability, nondisruptive upgrades, data-at-rest encryption, and data portability and application mobility across both on-premise and cloud environments.

These capabilities are part of what makes up a data-centric architecture, which centers around a shared set of data services. This allows legacy and modern web-scale applications to share data without boundaries.

One way of enforcing a data-driven culture is by using the self-SaaS model, which provides data architects, managers and administrators with a way to consolidate data and enable data services from a centralized pool. It also helps

eliminate many traditional challenges, including provisioning, maintenance and patching, sizing, human error, and cost. With a true subscription model, agencies can scale capacity up and down without disrupting operations while getting consistent support.

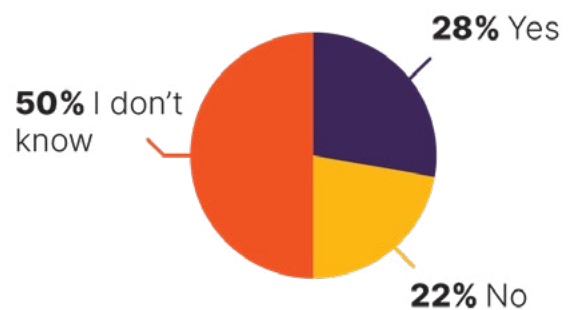
The survey found that just over a quarter of agencies use storage-as-a-service (see Figure 3), but Psaki didn't find that surprising. In fact, he said that many more probably were using it but didn't know it. “If you're doing it right, your users don't know the difference, which is why they may not realize they are actually using it,” he said.

A truly data-driven organization, with a data-centric architecture, gives an organization the ideal “360 degrees of freedom,” Psaki said.

“You can put your data where you want, when and how you want, without worrying about compatibility, interoperability, availability or durability, and you will be ensuring that your agency is ready for whatever comes next,” he said.

The goal is to have a seamless environment, in which agencies spend less time worrying about the mechanics of managing data and more time deriving value from it.

Figure 3: Are you using storage-as-a-service today?



Turning Performance to 11

Whether it's speeding how remote workers access data more quickly, generating insights for faster decision-making, processing constituent requests faster, driving digital transformation or simply improving productivity, agencies today understand the need for fast access to data and data services. They also need the ability to manage that data, wherever it resides. Achieving those goals requires superior performance, including high throughput and low latency.

According to the survey, nearly half of respondents cited the ability to rapidly access data as a major challenge, along with security. Other challenges include data quality concerns, legacy systems and software, and multiple types of data to manage (see Figure 4).

Figure 4: What are your agency's biggest challenges surrounding data today? Choose up to three.

55% Ability to rapidly access data



36% Security concerns



32% Data quality concerns



32% Legacy systems and software



27% More types of data to manage



All agencies have siloed stores of data, the result of decades of growth and change. While these silos have always presented problems, they are even more challenging today, as agencies look for ways to corral and use all data to meet mission goals and solve problems. One of the best ways to eliminate those silos is to migrate as much to the cloud as possible, making sure to combine it with infrastructure that ensures consistency across the environment.

Agencies today also are dealing with massive amounts of different types of data, many of which use different data services under the covers. The data also is stored in different formats, including solid state drives and hard, magnetic and optical disks. The variety of data and storage media can complicate accessing and using data. The solution is standardizing on a single set of storage capabilities that addresses all data types.

Over time, agencies also have adopted various tools to address specific issues. In fact, it's not uncommon for agencies to have dozens of tools to perform tasks such as backup, compression, deduplication, replication, snapshots, and synchronous and asynchronous replication. By standardizing on a single tool that provides all required data management capabilities, agencies can not only simplify the environment and save money, but increase the speed and efficiency of data management processes.

All of these challenges — siloed data stores, different data types and an array of tools — make it difficult to achieve the kind of performance agencies require to meet mission needs today. One way to address these issues while improving speed and performance is by adopting software-defined storage, where software controls storage-related tasks and is separated from physical storage hardware. The survey found that agencies are beginning to understand the benefits of software-defined

storage: Nearly 40% of respondents said they already use it to some extent (see Figure 5).

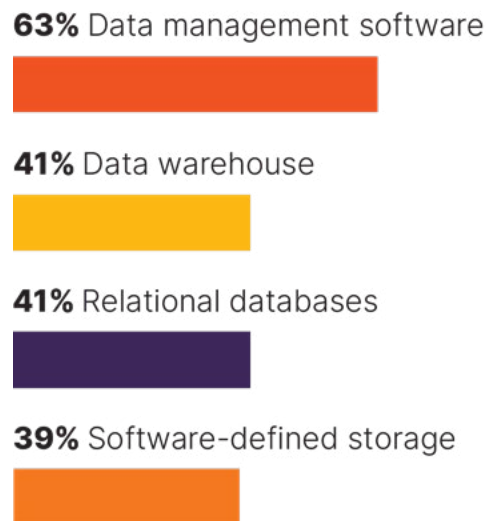
“Even if you think of a storage array as a piece of hardware, what really makes it work is the software,” Psaki said. With the ability to use any commodity compute and storage media on top of the software, agencies can grow and scale the data service platform without worrying about the individual components, he added.

Better yet, agencies can combine software-defined storage with data management software. This combination of technologies allows agencies to distribute and transform data into whatever capability they need, without the latency that plagues other approaches.

Together, these capabilities can deliver the performance agencies need today. They do this by reducing latency significantly. Pure Storage’s FlashArray, for example, delivers sub-millisecond latency at hundreds of thousands of input/output operations per second (IOPS), even in mixed workload environments. This provides high availability with nondisruptive operations, and significantly reduces unplanned downtime.

“You have to be able to interchange data in and out of the application at ever-increasing levels of performance to get the maximum value out of your data,” Psaki said. “It’s about making access to data easy and efficient, which requires the data service itself to move very quickly.”

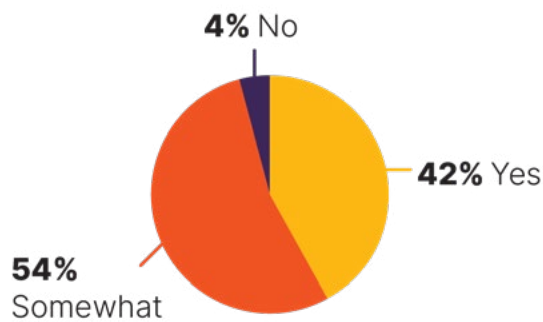
Figure 5: What tools and methods does your agency use today to manage data? Select all that apply.



Fighting Cybercrime on the Data Front

Cybercriminals have always been one step ahead, and they continue to find new ways to extort organizations through sophisticated malware like ransomware. The latest attack vector is data itself. In this scenario, hackers infiltrate networks months in advance of launching an attack, using the time to access as many systems as possible, including backups, and gather the credentials they need to encrypt data in a way that can't be unencrypted. Meanwhile, organizations back up their data as usual, not realizing that the data has already been compromised and encrypted.

Figure 6: **Do you believe your agency's data is well-protected against ransomware and other cybersecurity-related attacks?**



Some agencies have experienced this type of ransomware, and it's not going to stop anytime soon. One recent study noted that ransomware attacks against government agencies tripled between 2018 and 2019. Those kinds of numbers explain why agencies are concerned. Our survey confirmed this, finding that less than half of respondents were fully confident that their agency's data is well protected against ransomware and other cybersecurity-related attacks (see Figure 6). Security was also one of the biggest concerns agencies have about pandemic-related remote access (see Figure 7).

Agencies have valid reasons to be worried. Not only do they typically have multiple data repositories and applications spread between on-premise and cloud locations, but many of those data stores and applications predate modern technologies and the cloud, making them difficult to secure effectively. And these challenges have multiplied tenfold since federal employees began working remotely, requiring them to access data and applications from multiple locations.

Surmounting these challenges requires a different approach to data protection — one that ensures fast recovery and valid, data copies that haven't been compromised. Fast recovery is particularly important when it comes to recovering from a ransomware scenario.

Figure 7: **What have your agency's biggest challenges been in the data management and storage arena during the current pandemic?**

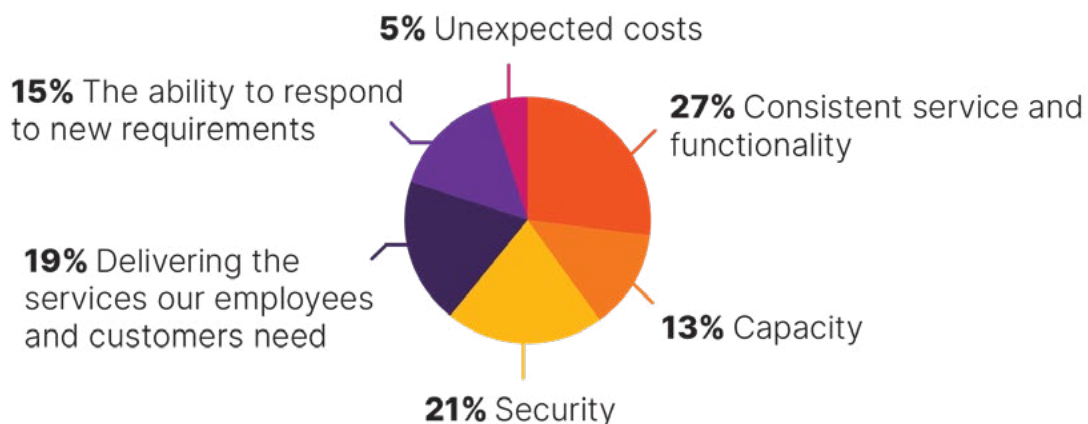




Figure 8: Which of these tools and methods do you consider an investment priority in the coming year?

56% Data management software



42% Master data management



31% Data transformation tools



23% Data modeling



A ransomware attack doesn't respond well to typical data recovery in which you might have to restore a few files or a single application, database or service. An agency might need to restore all of its data, which could take days or weeks. And during that time, its applications, data or services will not be available.

Ensuring that data is fully recoverable as fast as possible requires modern tools that are built from the ground up with security at their core. In other words,

security should be a native and integrated component of today's systems. Agencies understand this. The survey found that many agencies are considering investing in more modern tools that emphasize security (see Figure 8).

One key capability is a snapshot function that prevents attackers from compromising backups. For example, Pure Storage's FlashBlade includes SafeMode, which, when enabled, prevents even administrators from deleting snapshots. If hackers get admin credentials and can log into the array, they still can't delete the snapshots protecting your backup data. It's impossible for anybody to override them.

With this method, organizations can delete without worry the primary compromised data that hackers encrypted. Once deleted, storage administrators simply point the backup software to the metadata catalog that has been housed in the SafeMode snapshot, making it immutable. This allows business to get back to normal very quickly.

Ransomware may not be avoidable, but adopting more modern technology can go a long way.

"To defeat ransomware, you really have to make security an integral and transparent part of how a system operates," Psaki said. "It's the only way to stay a few steps ahead of cybercriminals."

How Pure Storage Helps

Pure Storage's solutions deliver a modern data experience — one that is simple, efficient, cloud-like, and enables data availability everywhere.

Solutions include:

- *FlashBlade*: A cloud-era flash array for harnessing the power of data at scale
- *FlashArray*: An all-flash, 100% Non-Volatile Memory Express shared accelerated storage designed for mainstream enterprise deployments
- *Pure1*: Cloud-based management and predictive analytics

- *Cloud Block Store on Amazon Web Services and Microsoft Azure cloud platforms*: Software-defined storage and flash management purpose-built to power Pure Storage's shared accelerated storage
- *Evergreen Storage*: A subscription-based approach to storage that ensures continuous upgrades

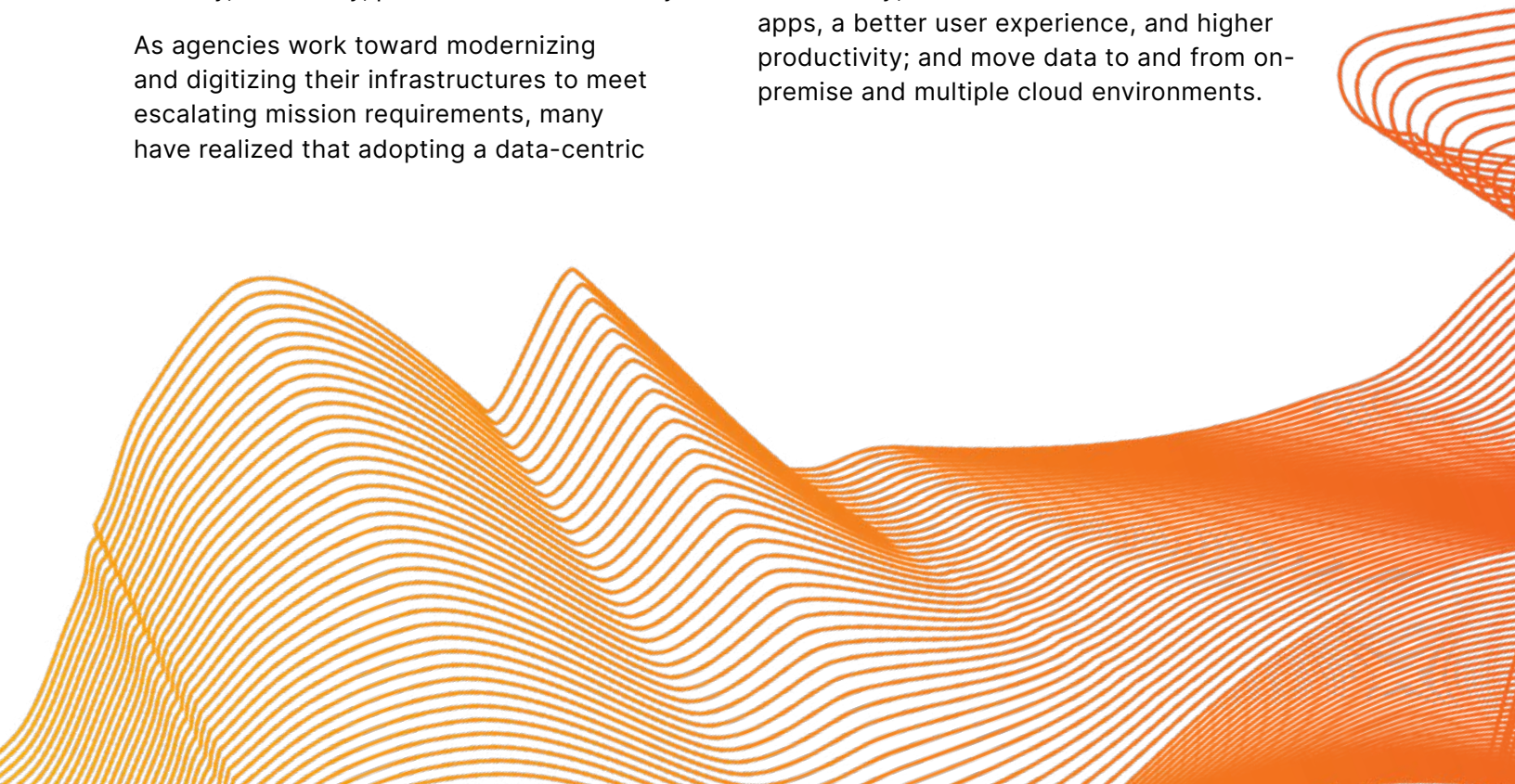
Pure Storage serves more than 50 federal agencies, and has been named a leader in solid-state arrays by Gartner. Its solutions are fully validated for Federal Information Processing Standard 140-2 and Advanced Encryption Standard-256 data-at-rest encryption, the U.S. government's highest security standard.

Conclusion

Agencies today understand the value of data and place increasing value on advanced capabilities such as artificial intelligence, machine learning, predictive analytics and DevOps. At the same time, they understand the importance of data visibility, availability, performance and security.

As agencies work toward modernizing and digitizing their infrastructures to meet escalating mission requirements, many have realized that adopting a data-centric

architecture that can share and deliver data on demand helps further those goals. Important capabilities in a data-centric architecture include the ability to consolidate and simplify all apps to a single storage tier and integrate them easily; access data in real time for faster apps, a better user experience, and higher productivity; and move data to and from on-premise and multiple cloud environments.





About Pure Storage

Pure Storage helps innovators build a better world with data. Pure's data solutions enable SaaS companies, cloud service providers, and enterprise and public sector customers to deliver real-time, secure data to power their mission-critical production, DevOps, and modern analytics environments in a multi-cloud environment. One of the fastest growing enterprise IT companies in history, Pure Storage enables customers to quickly adopt next-generation technologies, including artificial intelligence and machine learning, to help maximize the value of their data for competitive advantage. And with a Satmetrix-certified NPS customer satisfaction score in the top one percent of B2B companies, Pure's ever-expanding list of customers are among the happiest in the world.

To learn more visit www.purestorage.com



About GovLoop

GovLoop's mission is 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 the public sector.

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

www.govloop.com | [@GovLoop](https://twitter.com/GovLoop)



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

P (202) 407-7421 | F (202) 407-7501
www.govloop.com
[@GovLoop](https://twitter.com/GovLoop)

