

FAST, SHARED DATA IS A COMPETITIVE ADVANTAGE

Combined data and shared insights *are* the value in our data economy.

Data has become “the new currency” within our economy and is essential to value creation in our world today. The most valuable companies are no longer giants of infrastructure and physical resources, but instead the greatest purveyors of data.¹ Connectivity has given power to companies born in the modern era as data emerges as one of the most important assets. The avalanche of data will distinguish the winners from the losers, and the ability to make sense of information will be the deciding factor.¹

However, data alone is inherently powerless. Like currency, the value of data comes not from its existence or accumulation, but from the opportunity it represents. And, like currency, the potential value of data remains untapped by organizations that don't take purposeful action to invest it.

The actual value of data lies in what can be released from it: intelligence generated through aggregation, correlation, analytics, and real-time combination of different datasets. That is, data becomes more valuable the more it is processed and used. Perhaps most importantly, the value of data appreciates exponentially the more it's shared, when shared the right way and in the right place.

Change is hard, but inaction has consequences. The main barriers to sharing data, and therefore to tapping its value, lie in siloed data, clouds, and organizations.

- **Siloed data.** Data is often spread across the enterprise in discrete silos to meet the needs of specific applications. However, this inhibits modern analytics-driven workflows that demand data be available and delivered quickly as a consistent whole, not fractured in data silos. Additionally, cold data, collected as the economics of storing it improved, had its potential frozen in time when it was captured and forgotten, unintegrated with the objectives of your business.
- **Siloed clouds.** Cloud has earned its place in enterprise IT, but that place has so far been separate and distinct. Separate isn't shared, isn't invested, and isn't valuable. Those data silos withhold data's value from your business, constraining innovation and degrading your users' experience.
- **Siloed organizations.** At the same time, organizational silos, inherited from traditional separation of IT builders and technology operators, prevent the agility of DevOps where the builders *are* the operators and the service is their objective.

To overcome these barriers and release value from data, businesses today are becoming data-centric, and their infrastructures must evolve toward a data-centric architecture. A data-centric architecture is an approach to designing an end-to-end environment — across compute, network, storage, and cloud — that is optimized for ubiquitous and fast consumption of data. A data-centric architecture is characterized by five key pillars:

- Fast, shared data
- On-demand and automated
- Globally reliable and secure
- Hybrid cloud by design
- Constantly on and improving

Data-Centric Architecture

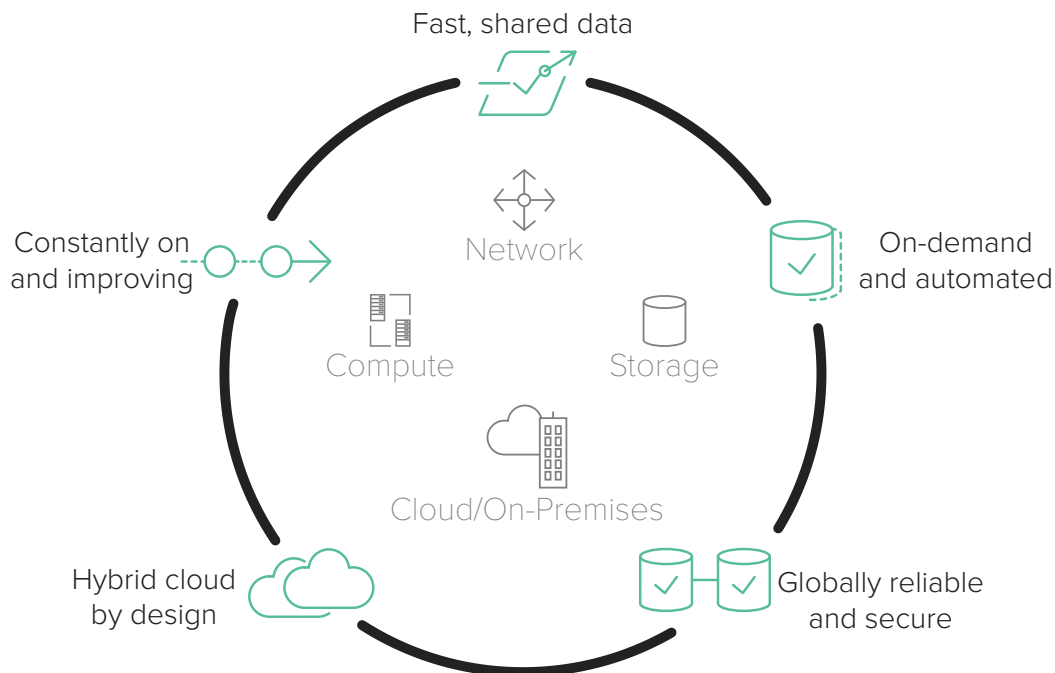


Figure 1. The five defining pillars of a data-centric architecture.

A data-centric architecture enables high-velocity processing of volume data that is accessible to all applications, users, and machines wherever they live, resulting in more efficient and agile processes.

This paper focuses on fast, shared data and the benefits that capability provides to businesses.

The New Data Environment

You might not need maximum velocity for every workload at all times, but data is a system, and you do need an environment that is capable of performing at speeds that will enable machine learning (ML) and artificial intelligence (AI) as part of a data strategy. These technologies are no longer nice-to-haves, but have achieved adoption as common tools within applications and data processing to accelerate the results that come from data operations. Think about enterprise applications you might already have deployed, such as Microsoft®, SAP®, and Oracle® solutions. Or consider your own consumption of data from sources that provide real-time customer or market intelligence, which you use as input to your own business. Your data plan isn't only about how fast you can process information, but also how fast you can ingest and use the data that feeds you, and then, in turn, deliver that data to areas of your business. A lack of performance means that downstream applications can be left waiting, with both machine processes and users experiencing inadequate response times, which can quickly create a perception, or even the reality, that your solution is insufficient for your users. In cases where collected human, sensor, and machine data are combined for new insights and value, fast data handling feeds the entire system, whose value rests on the immediacy of these insights.

The Value of Shared Data

The shareability of your data is key in realizing its value, turning data's *potential* into actionable *insight* and eliminating the silos highlighted earlier. Data is meaningless on its own; its value to your organization depends on your ability to share and analyze it, across clouds and applications, regardless of data type or size.

When you increase the speed of data ingress, processing, and egress from storage, you're on your way to taking advantage of ML and AI advances that have become recognized as critical new enablers in the data-economy value chain. A data-centric architecture that supports this type of data sharing can help prepare your organization for technological advancements and keep you ahead of the curve when it comes to getting the most value out of your data.

How Pure Storage® Products Enable Fast, Shared Data

At Pure, we launched the all-flash array revolution — and fast, shared data is a core principle. Our products deliver performance and agility to realize your fast, shared data needs:

- **FlashArray™:** Enable shared, accelerated data through 100-percent, end-to-end NVMe Express® (NVMe™) DirectFlash technology. Storage subsystems can become a bottleneck when mixed workloads are run concurrently, causing latency and reduced productivity. FlashArray offers flexibility, consistent performance, and low latency, reducing the impact on database performance when running a variety of concurrent workloads is necessary, and preserving your system speed. And, at up to 3 PB effective capacity in 6U, FlashArray is more efficient than ever, while providing 99.9999 percent proven availability for mission-critical applications.²
- **FlashBlade™:** FlashBlade arrays combine fast performance with large capacities for high-demand applications. The result is a data hub with elastic performance and concurrency, delivering performance no matter the workload. FlashBlade features petabytes of capacity and the ability to handle tens of billions of objects or files, which means you can share all your data. Run analytics, simulations, and software development 10x faster.³ Empower data scientists, developers, and the next generation of AI with effortlessly simple and scalable shared, accelerated storage.

- **DirectFlash™:** Remove data-transfer obstacles and add performance at scale. We're pushing beyond the inherent limitations of solid-state drive (SSD)-based all-flash array (AFA) and iSCSI architectures to a place of global flash management from DirectFlash software, hardware, and fabrics. In finally unlocking the true power of flash, you'll get up to 50 percent better latency, 5x greater density, and 15–30 percent more usable capacity from the same flash media.⁴
- **Pure1® cloud-based storage management:** Managing the exponentially growing data in a hybrid cloud world is an expensive and resource-intensive task for IT today. Pure1 powered by META - Artificial Intelligence Engine, helps customers automate data management and maximize their IT investments in the cloud and on-premises at their fingertips. Be it mobile or desktop, Pure1 offers simple, API-driven centralized views of enterprise data, full-stack analytics including VMs, predictive support, intelligent workload planning, hardware simulations and upgrades, capacity, and performance planning to deliver the shared data environment that unleashes the value of an organization's data.

Additionally, Pure offers converged solutions that integrate other data center stack components designed to accelerate meeting your business needs while enabling you to use the latest technologies in your data center evolution. These solutions include:

- **FlashStack™:** Based on 100-percent NVMe FlashArray or FlashBlade, our FlashStack solution delivers maximized performance and reliability across business-critical applications, DevOps, and analytics. Built for the cloud, FlashStack can fully integrate with Cisco®, VMware®, and OpenStack® platforms, in addition to others, so that you're able to share your data.
- **AIRI™:** Built on FlashBlade and NVIDIA® DGX-1™, AIRI extends NVIDIA DGX-1 systems with FlashBlade to run critical training workloads at scale, streamlining time to insight. It's built on state-of-the-art technologies like remote direct memory access (RDMA) over 100 gigabit Ethernet (GbE) and leading software tools.⁵ AIRI is tuned to keep graphics processing units (GPUs) busy and data flowing, intelligently processing your data to gain valuable insights.

Build a Foundation That Consistently Enables Fast, Shared Data

Pure is dedicated to helping you deliver the value of data — your most valuable asset — consistently, quickly, and to every user and application that needs it. Our tools and solutions are designed to evolve, standardize, and optimize at implementation and into the future. Your data-centric architecture should enable you to share and transform data into valuable information and intelligence.

Contact a Pure specialist today to learn more about a data-centric architecture and how it can make fast, shared data available for your business.

© 2019 Pure Storage, Inc. All rights reserved. Pure Storage, Pure1, Pure1 Meta, the P Logo, DirectFlash, FlashArray, FlashBlade, FlashStack, and AI/RI are trademarks or registered trademarks of Pure Storage, Inc. in the U.S. and other countries. All other trademarks are registered marks of their respective owners.

- ¹ Forbes. "Why Data Is the Most Important Currency Used in Commerce Today." March 2018.
www.forbes.com/sites/michelleevans/2018/03/12/why-data-is-the-most-important-currency-used-in-commerce-today/.
- ² Pure Storage. Entry-Level Storage web page. www.purestorage.com/products/entry-level-storage.html.
- ³ Pure Storage. "FlashBlade: Scale-Out Storage for Modern Data." 2018.
www.purestorage.com/content/dam/purestorage/pdf/datasheets/Pure_Storage_FlashBlade_Datasheet_05.pdf.
- ⁴ Pure Storage. DirectFlash web page. www.purestorage.com/products/purity/directflash.html.
- ⁵ Pure Storage. AI Infrastructure web page. www.purestorage.com/products/flashblade/ai-infrastructure.html.