PURESTORAGE®

WHITE PAPER

Your Data, Your AI: Building the Foundation for Tomorrow

Legacy storage is holding your AI back—it's time to fix it.

Contents

Introduction

Enterprise IT stands at a pivotal crossroads. The convergence of unprecedented data growth, escalating AI workloads, and mounting pressures around cost, security, and performance has exposed the limitations of legacy infrastructure. Traditional storage systems, once sufficient for static workloads and siloed data, are now straining under the weight of dynamic, data-intensive AI operations.

Traditional Storage Faces a Nontraditional Challenge

Artificial intelligence is no longer the future—it's the present. From generative AI and large language models (LLMs) to real-time decision-making, AI is rapidly transforming business models, customer experiences, and competitive landscapes. But unlocking AI's full value demands a modern foundation that can break down silos, automate operations, and deliver seamless, secure access to data at scale.

This white paper explores the unique demands of AI workloads and what it takes to build infrastructure that can handle them. It explains why traditional storage falls short, what's replacing it, and how the radically different Pure Storage® approach delivers the performance, agility, and resilience enterprises need to thrive in the AI era.

Storage Built for AI

Generative AI (GenAI) is transforming organizations' workflows and offerings. More and more companies are shifting from AI proof of concept to full-scale implementation across the enterprise, leaving behind their old operational models in favor of new AI-driven innovations that improve customer service, boost productivity, and expand capabilities. Is it any wonder that their legacy infrastructure can't keep up? Traditional storage was built for a world that we've moved beyond.

IT buyers identify GenAI as the top area where they expect IT spending to increase this year.[1] As companies invest more time and resources into their AI initiatives, having the right technical infrastructure supporting them becomes increasingly critical. This is especially challenging because AI applications draw on enormous datastores, often including data in many different formats. When your GenAI applications can access and fully utilize your organization's data instead of a nonspecific data set (such as a public LLM might use), you gain richer, more valuable results. But that's hard to do when your data is siloed.

In a traditional storage infrastructure that has grown over time, data is often fragmented, which makes it difficult to access and integrate the diverse data sets that AI applications need. Without the ability to easily access all relevant data, your AI deployment is at a disadvantage. Instead of the real-time, data-driven decision-making your business needs, you might be left with unsatisfying and incomplete answers. An abundance of data with no way to truly gain insights is a common challenge, and it can make your business more vulnerable to poor decision-making and inefficiencies. You need real insights, not screens full of numbers.

With enough effort, you might be able to consolidate and organize all that data. But without a modern solution, data sprawl will slowly but surely encroach again. It's similar to decluttering your house. You can enjoy your clean living room for a few days, but if you don't have clear and consistent places to put your shoes, keys, and mail, it's only a matter of time before everything gets messy again.

With multiple copies of the same data spread across traditional storage systems, managing copy data becomes even more complex. There's a cruel irony here: a successful AI initiative requires addressing copy data management. However, that same work also makes it harder to implement AI. By complicating operations and increasing overhead, you reduce your agility, one of the most important elements of a successful AI implementation.

As your AI operations expand, these issues grow alongside them. With more users accessing and contributing to the development of your AI tools, it becomes more important that only authorized users can access the data fueling your AI initiative. But the challenges go beyond security. Traditional systems have a rigid way of provisioning, with preset resources that may or may not be a good fit for evolving AI workloads. At one extreme, your datastores may grow beyond the capacity of your storage system, creating performance bottlenecks. On the other hand, you might allocate too many resources for a workload, resulting in inefficiencies and wasted capacity. And because data in traditional storage systems is so siloed, you might experience both problems simultaneously.

These data management and maintenance tasks take time, energy, and resources. Some estimates suggest that companies spend up to 80% of their IT budgets simply maintaining legacy systems.[2] Your budget, not to mention your IT team, wasn't built to handle this kind of pressure.

Finding the Right Storage Approach for AI

Rigid traditional storage is ill-suited to the needs of today's dynamic, AI-driven environments. But once you've chosen to leave behind your traditional storage, the next step is determining the path forward. New technologies may offer increased performance and other advancements, but it would be a mistake to simply upgrade to a newer solution with the same underlying architecture that caused issues. Your AI workloads require a solution built for today's problems and today's opportunities, one that takes an entirely new approach.

As you map out your storage strategy, consider the following:

- What capabilities would enable maximum agility for new and volatile AI initiatives?
- What solution would empower you to consolidate all your data to use it most effectively for AI-driven decision-making?
- How can you empower artificial intelligence for IT operations (AIOps) with reliability, flexibility, and high performance?

The following subsections outline the key features of a modern storage platform for AI.

Unified Architecture for Operational Simplicity

AI workloads require access to a wide range of data types, often spread across different systems. This need strikes at one of the most significant weaknesses of traditional infrastructure. Legacy environments rely on separate storage silos for block, file, and object data, creating operational complexity and slowing AI initiatives. A unified storage platform with native multiprotocol support eliminates these silos by allowing all data to be managed through a single architecture. This simplifies operations, accelerates data access, and reduces the burden of maintaining multiple systems, empowering your teams to move faster with AI and AIOps.

Automation and Agility for a Changing Landscape

AI-driven initiatives demand extreme agility. Adopting a storage platform that embraces automation and offers built-in flexibility allows you to keep pace with evolving workloads. The ideal platform provides autonomous configuration and self-tuning capabilities to scale performance without manual intervention as capacity grows.

Your team shouldn't have to spend time on the same routine maintenance tasks they were performing 20 years ago. Instead, they should be implementing those ambitious AI initiatives while your storage platform adjusts dynamically.

Comprehensive Observability with AIOps

To manage and optimize AI workloads and the data that fuels them, you need robust AIOps. With unified observability and modern AIOps, you gain real-time visibility into who's accessing your data and how they're using it. Your storage platform should also give you instant insight into your environment's holistic health and performance, alongside automatic problem identification and proactive, AI-driven resolution. That capability should go beyond identifying anomalies and user access issues; your storage platform should be able to respond to the needs of AI workloads in real time, from adjusting resource allocation to tuning performance.

Hybrid-cloud Compatibility for AI at Scale

AI initiatives rely on fast, flexible access to large volumes of data, both on premises and in cloud environments. A storage platform built for the hybrid cloud can deliver that access. Such a platform would support seamless data movement, consistent performance, and centralized management across your entire infrastructure. This improves resource efficiency, simplifies operations, and allows AI models to handle training and inference where it makes the most sense, all without adding complexity or cost.

Flexible Capacity and Scalability

Because AI workloads have unpredictable capacity demands, the ideal storage platform for your AI initiatives can quickly and automatically adjust storage resources without manual overhead. That means scaling up when your workloads require it and scaling down just as easily, helping you avoid the costly overprovisioning often associated with traditional storage platforms.

You may not start your AI initiatives with large volumes of data, but as your data needs grow, you need a solution that can offer simple, nondisruptive expansion. Ultimately, for your most complex AI workloads, your platform should be able to easily expand to deliver high performance with petabytes of data and billions of individual data items.

High Availability and Continuity

High availability and durability are critical requirements for AI workloads due to the need for continuous, uninterrupted access to immense stores of data. The right storage platform can offer continuous operation for real-time, AI-driven decision-making. Robust loss protection is another important piece of the puzzle, ensuring that no matter what happens with your hardware, software, or operations, you can retrieve the critical data that fuels your AI initiatives.

Your AI initiatives will likely grow more critical over time, and it is all but sure that AI workloads and tools will grow more complex and capable. Google Gemini, for example, doubled the number of tokens it could process in only a four-month span.[3] In the same way, your storage platform must be able to keep pace with your growing data needs. Ideally, your solution would offer easy software upgrades and seamless hardware upgrades and replacements so you can keep growing your capabilities without the financial disruption of a new purchase or the operational disruption of downtime around an upgrade.

Why Choose the Pure Storage Platform for AI?

AI initiatives may be some of the most complex, technical work your teams have ever handled. There's no need to increase their burden by requiring them to use a cumbersome, outdated storage platform. An Enterprise Data Cloud approach utilizing the unified, as-a-service Pure Storage platform can smooth their path and yours.

An Array Designed for AI

Although the Pure Storage platform offers much more than high-performance hardware, arrays are a good place to start. The Pure Storage FlashBlade//EXA™ brings agility, scalability, and exceptional performance for AI workloads. Among the many storage array options available with the Pure Storage platform, FlashBlade//EXA stands out as an AI powerhouse, maximizing performance density and empowering you to accelerate all stages of your AI workloads. Production, inference, and training happen at the speed of business. With the Pure Storage platform, you can truly make the most of your AI investment—eliminating idle graphics processing units (GPUs) caused by a storage solution that can't keep up.

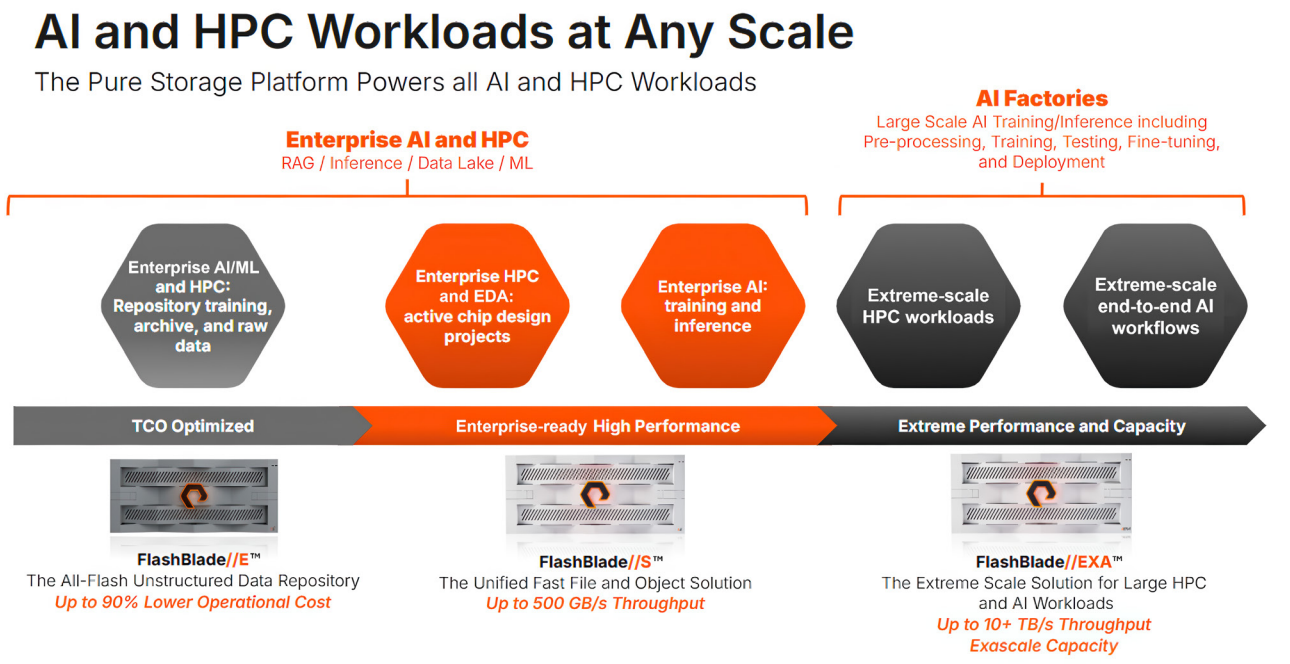

FIGURE 2 AI and high-performance computing powered by the Pure Storage platform

Management Reimagined

By choosing the Pure Storage platform, you're replacing the many data silos of traditional storage with a single unified platform with one control plane. The Pure Storage platform includes AIOps from the ground up with Pure1®. Pure1 uses machine learning to process data, extract insights, and streamline IT operations. It offers intelligent workload placement, security recommendations, and performance troubleshooting, among other services. When your platform can handle this much traditional IT work without human intervention, your IT teams have the freedom to focus on higher-level tasks. With maximum visibility, increased automation, and easier management, our customers have seen as much as a 75% reduction in storage management time.[4]

Our Portworx® platform further eases the management burden by simplifying and automating container management and protection for Kubernetes workloads. By making common data sets accessible to all data scientists, no matter where the data is located, Portworx provides the agility and data mobility that AI workloads so desperately need. Like the rest of the Pure Storage platform, it takes an as-a-service approach that enables easy scalability as you plan for larger AI implementations.

Scale and Upgrade Painlessly

As your datastores continue to grow and AI technologies get more complex (and more valuable), you need a storage infrastructure that scales along with you. With the Pure Storage platform, your storage pool can evolve and expand with user demands. At the same time, Pure Fusion™ automated workload orchestration ensures efficient utilization of storage resources, supports dynamic provisioning, and enables proactive security and compliance management. Unlike traditional storage, the Pure Storage platform is designed specifically for AI, analytics, and other next-gen workloads with ever-increasing data demands.

Larger upgrades aren't going to be a problem, either. The Pure Storage platform is future-proof: Evergreen® lets you invest once and stay current forever. With Pure Storage, you can rest assured that you always have the most modern storage technology—reliable and highly available—without worrying about major software upgrades or disruptive hardware upgrades. Evergreen//One™ for AI expands on that promise with performance guarantees based on the maximum bandwidth requirements of your GPUs and accelerators. Instead of bottlenecking on legacy storage, your AI implementation can scale up seamlessly as it evolves or scale down as you iterate on specific workloads.

Future-driven Resilience

All of us expect the services we use every day to just ***work,*** and you should be able to expect the same of your storage. To keep data flowing to AI workloads, Pure Storage gives you unprecedented availability and a guarantee of up to 99.9999% uptime with monitored and managed storage as a service (STaaS) SLAs.[5] Even during unforeseen disruptions, your critical AI workloads can keep running. The Pure Storage platform offers a highly resilient architecture built to withstand threats from bad actors, natural disasters, and everything in between.

Best-in-class Partnerships

To help you achieve your AI goals, we've embraced partnerships with other companies delivering AI excellence. Through close collaboration with NVIDIA®, the Pure Storage platform is certified as a high-performance data storage platform for NVIDIA Partner Network Cloud Partners. Together, we offer an AI-ready infrastructure (AIRI®) reference architecture validated and certified with NVIDIA technology. Other results of these partnerships include prevalidated designs with Cisco for easy deployment, scalable AI performance solutions with CoreWeave, and support for multiple vector databases ideal for AI workloads. The Pure Storage platform goes beyond storage to deliver a truly AI-ready infrastructure, including the best of the broader AI ecosystem.

The Bottom Line

Executives overwhelmingly believe in GenAI's potential, with 87% expecting it to boost revenue within three years.[6] But those windfalls won't just happen. You must strategically transform your IT infrastructure to take full advantage of the GenAI revolution. With the high-performance hardware, streamlined management tools, easy scalability, unprecedented availability, and partner ecosystems that future workloads demand, the Pure Storage platform is the right choice to help you realize AI's full potential.

Conclusion

AI is reshaping the enterprise at every level, from how data is managed to how decisions are made. But to fully realize the transformative power of AI, organizations must first overcome the architectural and operational constraints of legacy infrastructure. Traditional storage, designed for a pre-AI world, can't meet the speed, scale, or intelligence today's workloads demand.

The Pure Storage platform delivers a future-ready foundation for AI—a unified, automated, and intelligent data architecture that simplifies operations, accelerates performance, and ensures security and availability at scale. With built-in AIOps, hybrid-cloud compatibility, and seamless scalability, it turns data into a competitive advantage rather than a costly liability—fueling innovation while freeing IT teams from a firefighting mindset.

In a world where data is the new fuel and AI is the engine of growth, your infrastructure should empower progress, not hinder it. With the Pure Storage platform, you're not just upgrading your storage—you're building a smarter, more agile future.

Additional Resources

- Explore the Pure Storage platform.
- Learn more about the Enterprise Data Cloud.
- Read about how the Pure Storage platform supports databases and other solutions.
- Learn how to build your Enterprise Data Cloud with the Pure Storage platform.

1 | https://www.bcg.com/publications/2025/ai-shifts-it-budgets-to-growth-investments

2 | https://www.fastcompany.com/91285955/how-ai-will-end-the-era-of-risking-it-all-to-modernize

3 | https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work

4 | http://purestorage.com/customers/halifax-health.html

5 | https://www.purestorage.com/content/dam/pdf/en/datasheets/ds-pure-storage-purity.pdf

6 | https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work

ID2389-01-en-08/25