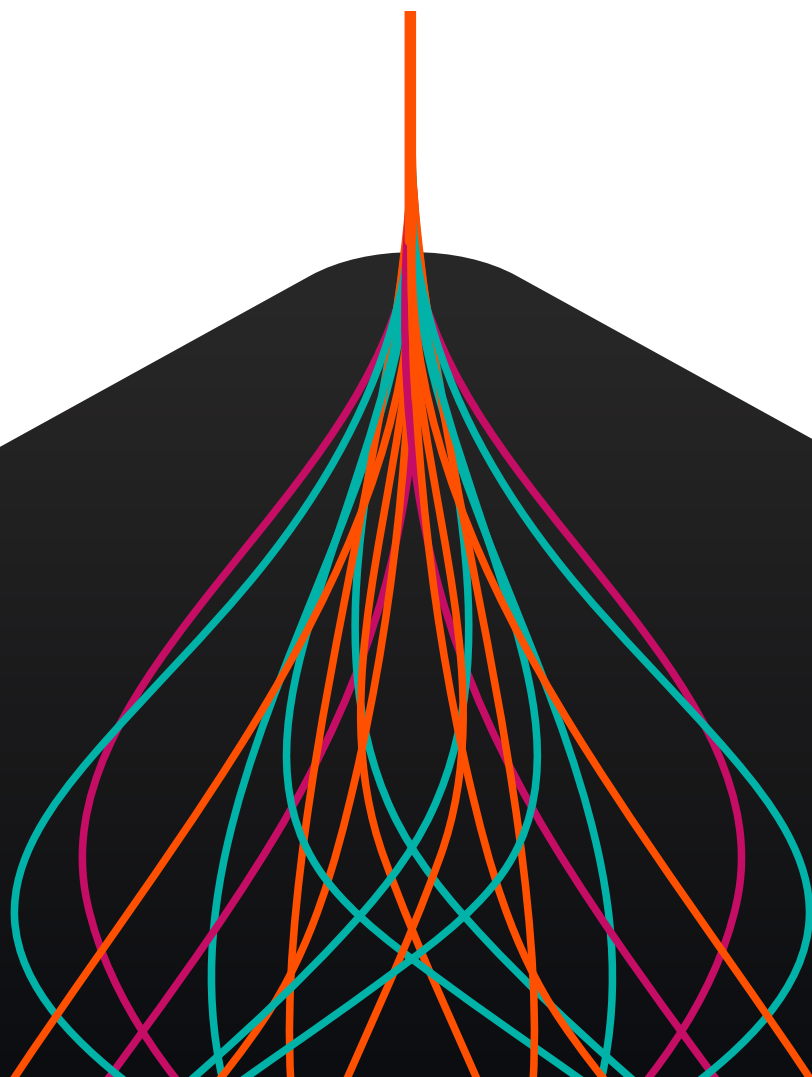# *The Future* of Data

New Principles for the AI Age

Content

A fundamental shift is happening in IT. Artificial intelligence, cyberthreats, modern apps, sustainability, and more are upping the stakes for organizations—and putting tremendous pressure on IT leaders.

To mitigate evolving risks while maximizing the rewards of innovation, a new way of thinking about data infrastructure is required. This ebook will help you make sense of the challenges ahead and prepare for the opportunities of tomorrow.

Introduction

We're at a fundamental inflection point in data infrastructure.

The past decade was focused on digital transformation, which made data the lifeblood of every organization. Old ideas about IT infrastructure needed to be cast aside, and new ways of thinking were required to meet challenges of scale, speed, and complexity. Forward-thinking IT leaders made use of emerging trends, like virtualization and the public cloud, to accelerate their transformation efforts. In many cases, these digital transformations would have been severely limited—or impossible—without these new approaches. The results are such that today, every organization is now a data organization.

IT is once again at the center of a transformation—artificial intelligence (AI). The benefits and use cases seem almost limitless for both saving money (increased efficiencies and productivity, for example) and increasing revenue (e.g., opening new markets and improving customer experiences). Mass adoption of generative AI (GenAI), in particular, is changing the operations of every organization, and the imperative to leverage it is only increasing. AI is already putting tremendous strain on our data infrastructures. In a recent survey of 1,500 IT leaders, **80% said that they are worried that their business will be left behind if their infrastructure can't support AI fast enough[1].**

In this ebook, we'll look at the challenges ahead and explore why legacy data infrastructures won't be enough to meet them. We'll also lay out new principles which we believe are needed to thrive in an AI-focused world—and explore how you can leverage and benefit from them.

A Recent Survey of 1,500 IT Leaders Found:

80%

Said that they are worried that their business will be left behind if their infrastructure can't support AI fast enough[1].

Other Forces Compounding IT Challenges Include:

- **Increasing cyberthreats**
- **Modernizing applications**
- **Balancing risks and innovation**
- **Sustainability initiatives**

1 | "The Innovation Race: Reducing Risk and Navigating the AI Frontier for Future Success", from Vanson Bourne and Pure Storage, June, 2024

01 From Terra Firma to Shaky Ground: AI Is Only Part of the Shift

The move to leverage AI (and GenAI, in particular) in organizations represents a fundamental shift in how we both store and capitalize on data. The traditional approach was to select data storage systems for individual use cases and workloads—databases, applications, backup, etc. This strategy has left many organizations with a patchwork of disparate, siloed data stores, both on-premises and in the cloud.

But in order to leverage that same data for GenAI projects, organizations must break down these silos, so that AI models can be fed. Otherwise organizations risk bad or incomplete GenAI models—and bad outcomes. This could mean missing out on the efficiencies and productivity gains that AI can deliver, and losing competitive advantage. This is just one example of how AI is straining on existing data infrastructure.

There will not only be the expense of acquiring the needed infrastructure but also the planning for vastly increased power and space requirements. This growth must be managed in an increasingly cost- and sustainability-focused IT ecosystem, where approving massive expansion projects—or massive public cloud expenses—can't be taken for granted. It's clear that working under existing assumptions about data infrastructure won't get the job done for AI.

Right alongside the shift to support AI, cyberthreats such as ransomware attacks are still on the rise, occurring at rates faster than one every minute. The financial impacts are growing as well, with the average global cost of each breach at almost $4.5M (and more than double that in the U.S.). New threats continue to challenge traditional security models, as AI-based attack techniques become increasingly sophisticated.

Cyber resiliency has become a boardroom-level conversation, given the massive exposure to every organization and potential impacts to brand and

customer experience on top of the financial impact. IT leaders are spending an exorbitant amount of their time and resources on preventing cyber and ransomware attacks. Therefore, data security will remain a top priority for any infrastructure improvements, encompassing both prevention and rapid recovery from attacks.

The Issues Created by Managing the Data That AI Requires Shouldn't Be Taken Lightly:

81%

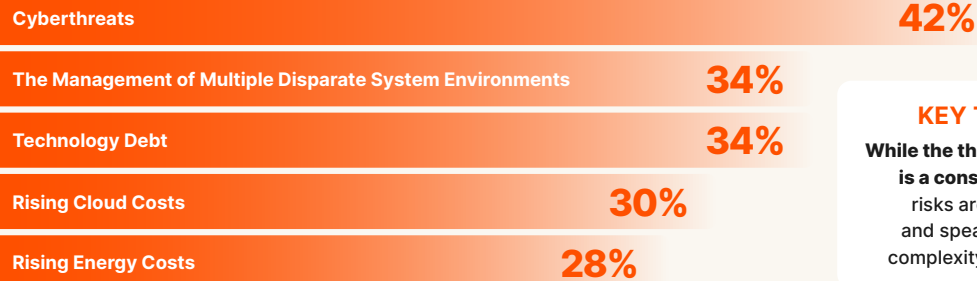Believe that AI-generated data is likely to outgrow their organization's current data centers.[1]

The Need to Balance Risk, Costs, and Innovation Is a Constant Struggle for It Leaders:

86%

of surveyed place the reduction of their organization's risk profile as their top priority.[1]

1 | "The Innovation Race: Reducing Risk and Navigating the AI Frontier for Future Success", from Vanson Bourne and Pure Storage, June, 2024

According to IT Leaders, the Risks Having the Greatest Impact Include:[1]

Cyberthreats	42%
The Management of Multiple Disparate System Environments	34%
Technology Debt	34%
Rising Cloud Costs	30%
Rising Energy Costs	28%

KEY TAKEAWAY

While the threat of cyberattacks is a constant focus, other risks are not far behind and speak to the growing complexity of infrastructure.

Notably, one of the biggest threats to any transformation initiative—technical debt—was cited by 34% as the risk with the most impact.

Modernizing the application stack is a good example of the innovation that IT is trying to achieve. Both existing apps and next-generation AI, analytics, and media apps are being reimagined from the traditional monolithic systems of the past to a combination of containers and microservices. These cloud-native applications are often more resilient and flexible than traditional apps. They can more easily scale (both within the data center and across a hybrid cloud model), as well as process and interpret data in real time. These advantages help boost organizational agility and reduce costs. But these application changes also mean that traditional data storage strategies must adapt into one of delivering data services. For example, data services for these modern workloads need to be fully orchestrated and automated via APIs, which often requires new storage architectures.

In addition, organizational sustainability initiatives are increasing globally, with a focus on how IT can reduce its environmental impact. While sustainability regulations are currently more prevalent in Europe and Asia, the Americas are seeing increased sustainability focus due to organizational ESG initiatives. And even in the U.S., regulations are more likely to increase in the future. Local utilities are already placing regional limits on data center expansion as they face pressures from, among other things, the growth of electric vehicles and outdated power infrastructure. AI in particular is often singled out as a workload that consumes massive amounts of both energy and data center space. These factors are driving the search for evermore efficient IT infrastructure, especially data storage. Even if an organization is sustainable in other areas, inefficient data centers can damage reputation with customers, investors, and employees. In other words, efficiency is no longer just a "nice to have", it's becoming a necessity.

They Also Recognize That the Focus on Managing Risks Can't Come at the Expense of Innovating[1]:

6/10

58%
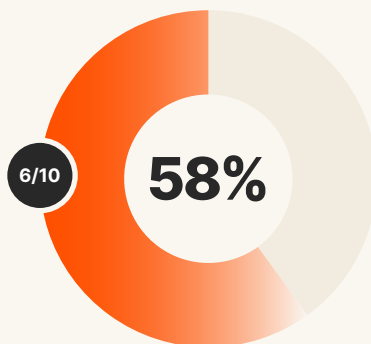
58% said that **current business needs** require more time spent on innovation.

Whether the main pressure you're feeling is AI, security, modern apps, or sustainability—or all of them at once—things are changing rapidly, and IT needs to think differently in order to respond to organizational needs and accelerate innovation. **Change must come in many areas of the IT stack, but it is felt most keenly in data infrastructure.** And what will be the next challenge? If your data infrastructure isn't designed to be agile, you'll be limited in your ability to respond to whatever comes next.

1 | IT Leaders were asked to select their top three risks. Percentage figures are the number of respondents who selected that risk.

02 Old Ideas about Data Storage Will Hold You Back

IT leaders recognize that data storage is complex.

Many assume that only marginal improvements can be made over time, based on experiences they've had with traditional legacy storage solutions. But the assumptions commonly held about storage are largely driven by old ideas that are baked into these legacy systems. Legacy data storage is rigid, inflexible, inefficient, and complex, and thus is ill-suited to handle IT challenges such as GenAI. Simply adding more of the same legacy data storage, entrenched as it is in the old ways of thinking, will only make projects more difficult and resource-draining, limiting outcomes and potentially derailing projects completely.

How Legacy Storage Hinders IT Goals

The more complex a storage system is to set up and manage, the larger the impact will be on staff,
delaying projects, increasing costs, and tying up resources that otherwise could be tasked to AI projects.

Static, Vertical Storage Configurations

- **Traditionally, storage has been designed and implemented for each workload, use case, feature need, and location.** The ability to scale and share data across these vertical storage systems is very limited, hindering projects like AI that require consolidated, scalable data stores. Static configurations also reduce the ability to apply automation, making data management even more complex and limiting the accessibility of data. For example, data services for modern application development are often managed outside the traditional workload and storage environment, adding even more complexity for both developers and IT staff. Data security often has to be applied on each system via manual patching, adding layers of complexity, increasing the potential for human error, and presenting a larger cyber "attack surface"—all of which can lead to vulnerabilities. And inherent inefficiencies in both the design and implementation of these legacy systems will make it difficult to meet goals around lower energy use and increased sustainability.

Siloed, Inconsistent Experiences

- **Legacy storage offers disparate product lines that are "platforms" in name only.** Different UIs, management systems and architectures have been cobbled together, and are often rooted in decades-old HDD systems and various technology acquisitions. Cloud storage implementations from legacy vendors often have vastly different management and functionality compared to on-premises implementations, severely limiting the benefits of a hybrid-cloud strategy. This legacy approach adds complexity, inefficiency and cost, and reduces scalability.

Complex Provisioning and Configuration

- **The legacy storage world is one of "trouble tickets:" complicated procedures that require IT specialists to manage each step for IT's internal customers when they need access to data infrastructure.** As new workloads and ways to leverage your organization's data increase so will the burden on IT staff to manage the provisioning, configuration, and security of data storage. This becomes even more limiting in the context of AI and application modernization projects, throwing up barriers to success.

Manual Processes to Optimize Performance and Maintain Storage

- **Time-consuming processes are also required to optimize performance, maintain data storage uptime, and integrate data with other systems.** This is the complex, behind-the-scenes work that legacy storage requires from IT to make sure things don't break, slow down, or can't scale with your needs. The varied tools and interfaces across a vendor's product portfolio only complicate this work, requiring teams of specialists and tying up resources that could be working on more strategic projects, like AI.

Constant and Disruptive Upgrades

- **Here is where the rigidity of legacy storage is felt most keenly.** Storage systems that cannot scale require repeated, costly and disruptive infrastructure upgrades, which tie up resources and drive up costs. This planned obsolescence is built into legacy storage and is part of the legacy storage vendor business model, even if limited upgrades may be offered on paper. And often there is no ability to scale between different products (even from the same vendor) as needs evolve.

New Thinking
Required to Tackle Modern Challenges

AI-Ready Infrastructure

Cyber Resiliency

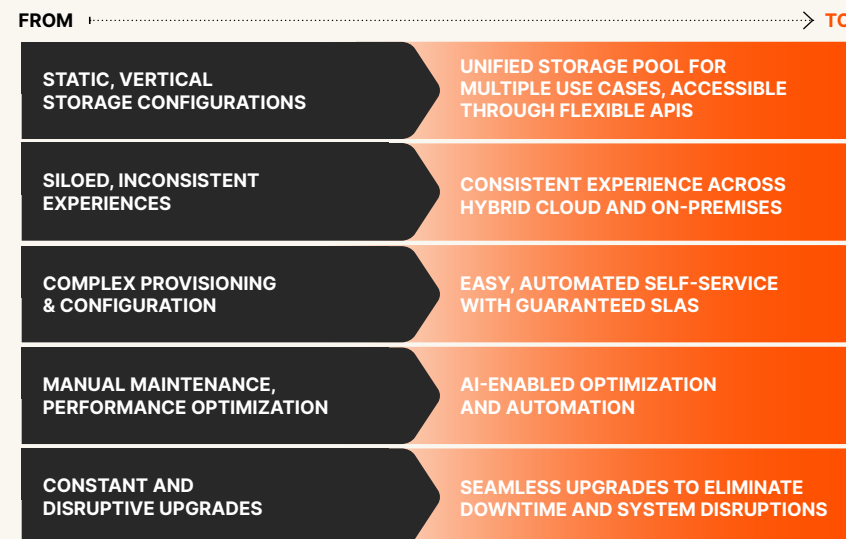**Application Modernization**

Sustainability

03 New Principles of a Modern Data Platform

You must fundamentally challenge your assumptions about data infrastructure to meet the challenges of today and tomorrow. You must rethink its role, transforming from a disparate collection of products in a data center to a strategic enabler spanning from the edge to the cloud, via a cohesive, consistent data platform. A new set of modern principles, for a modern data platform, is needed to thrive in this new world of IT.

- **A unified storage pool** solves the problems associated with traditional legacy static and vertical storage systems. A common, powerful, and highly efficient data platform that works across multiple use cases and is accessible through flexible APIs is needed to deliver the simplified data consolidation at scale organizations require. This approach is crucial to successful AI projects. It also helps with modern applications, so that API-driven data services can coexist on the same storage platform as traditional workloads. An efficient platform delivers scalable performance for different workloads from mission-critical to backup use cases and solves for the space and power limitations of traditional legacy storage.

- **A consistent experience** is delivered across hybrid cloud and on-premises data infrastructure, instead of siloed and inconsistent experiences found in legacy storage. A common architecture and interface solves this, so data can be more easily managed and accessed regardless of data type, workload, or performance need. This empowers IT staff to manage larger, more varied data stores with much less effort and improves data security, as well.

- **Easy, automated self-service** with guaranteed SLAs, as opposed to complex, manual provisioning and configuration. This simplicity should extend across all data stores and services, including those used in modern, container-based applications.

New Principles of a Modern Data Platform

FROM ⟶ TO

FROM	TO
STATIC, VERTICAL STORAGE CONFIGURATIONS	UNIFIED STORAGE POOL FOR MULTIPLE USE CASES, ACCESSIBLE THROUGH FLEXIBLE APIS
SILOED, INCONSISTENT EXPERIENCES	CONSISTENT EXPERIENCE ACROSS HYBRID CLOUD AND ON-PREMISES
COMPLEX PROVISIONING & CONFIGURATION	EASY, AUTOMATED SELF-SERVICE WITH GUARANTEED SLAS
MANUAL MAINTENANCE, PERFORMANCE OPTIMIZATION	AI-ENABLED OPTIMIZATION AND AUTOMATION
CONSTANT AND DISRUPTIVE UPGRADES	SEAMLESS UPGRADES TO ELIMINATE DOWNTIME AND SYSTEM DISRUPTIONS

The focus should be on empowering self-service to simplify operations and drive ease of use. Then, data maintenance becomes almost invisible, enabling generalists to manage it all. This approach is enriched when as-a-service SLAs are guaranteed, further reducing the burden on IT staff.

- **AI-enabled optimization and automation** versus manual and time-consuming processes for optimizing performance, maintaining data uptime, enhancing data security, and integrating with other systems. In other words, solving for the complex, behind-the-scenes work that legacy storage requires from IT to make sure things don't break, slow down, or can't grow with your needs.

- **Eliminating downtime and system disruptions** with a data platform that's designed to scale and be upgraded continuously without the constant, disruptive upgrades required by legacy storage. After all, hyperscalers don't take your systems offline for an upgrade. None of your infrastructure should be constrained like that—whether in the cloud or on-premises.

04 How You Can Benefit from the New Principles

By adopting the new principles of modern data storage, you can realize the simplicity, agility, performance, reliability, and efficiency needed to tackle the challenges of today, even as you prepare your organization for excellence in the future. When you focus on these principles while you are building out your IT strategy, you can increase your organization's competitiveness and fuel both top-line growth and bottom-line savings.

This new approach is far simpler than that of traditional legacy storage, which is mired in old assumptions. It also delivers you significant benefits.

Cost Optimization

- A consistent, automated, and easy-to-use interface helps reduce costs associated with staffing, since you can accelerate the move toward IT generalists instead of needing specialists to manage your data storage infrastructure.

- An efficient platform means that you use significantly less space and power, saving on operational costs.

- Eliminating downtime and increasing the reliability of your infrastructure increases staff efficiency, reduces maintenance costs, and avoids other costs associated with downtime which can affect both your brand and your bottom line.

Innovation and Agility

- Scalability for both capacity and performance provides the agility you need to meet not only the challenges of the AI age, but also other changes that may come—both near- and long-term.

- Staff time can be freed up for more value-add projects, increasing your speed of innovation.

- You can leverage new technology as soon as it's available, further accelerating your innovation and reducing your technical debt.

Sustainability

- Efficiency also enables the flexibility to avoid any regulatory or other limitations placed on data center expansion and energy use, especially pertinent to massive AI projects in an era of increased focus on sustainability.

All of these benefits combine to provide a solid foundation that will increase your organization's agility and ability to achieve goals and compete well into the future.

05 How to Transform Your Data Infrastructure

Rethinking your own data infrastructure may seem like a daunting task, but it can be simplified by taking a phased approach. It's important to remember that your data infrastructure is in service of your overall IT and business strategy. Clearly understanding your business and IT priorities will allow you to phase in this transformation in a way that supports—and doesn't disrupt—the business, even on an accelerated time scale.

Pure Storage has worked with thousands of your peers to build the right data infrastructure for their organizations. From those engagements, we have curated these best practices to help you get started.

Phase 1

Prework: Discovery, Identify Stakeholders and Executive Sponsorship

This may seem like an obvious first step, but waiting until the end of a scoping project to identify stakeholders, understand their priorities and get their buy-in can derail even the best plans.

- **Make sure you know the business objectives** for the next one to three years. Tie your high-level objectives into the business strategy.

- **Identify goals and "wish lists"** with application owners, data scientists, operational teams, and other stakeholders

- **Identify and engage critical Decision-makers—**Direct management, up to CFO as needed—not with detailed strategy yet, but to understand business goals better and gauge executive interest.

- **Designate an executive sponsor** and an additional champion (such as a business owner).

Phase 2

Strategy and Goals Definition

This is the most crucial phase. This up front planning will allow you to set expectations with your key stakeholders, minimize scope creep, and accelerate implementation throughout the project.

- **Business Goals—**How does the data strategy tie into the company's business objectives? How will an investment in a modern data platform drive the company to strategically differentiate, grow revenue, and/or drive efficiency?

- **Architectural Strategy—**How will you evaluate deploying the data platform across a hybrid cloud environment. Where will you develop? Where will you test? Where will you run production?

- **Project Strategy—**What are your time frames? Build out a high level schedule. Who is on the main project team? Start identifying the key team members.

Phase 3

Current Infrastructure Assessment

Start by identifying the systems that can most benefit from streamlining and a platform-based upgrade.

- Discovery is important. Evaluating your current environment is critical. Where is data siloed? What kind of systems capabilities currently exist? How are applications currently architected, and which ones (existing and planned) benefit most from modernization?

- **Start with the most acute pain points.** Where are there performance or reliability issues that are causing business disruption? Where does your organization currently have technical debt? Where does the organization have resource/talent debt requiring a more streamlined approach to management of systems? What are the true costs of running your legacy infrastructure?

- **Think toward the future, especially around AI.** For modern applications requiring high performance (for example, AI model training and RAG), do you have the capabilities required to be successful? Do you understand what all the teams have planned, and what they will require?

Phase 4

Removing Roadblocks to Success

Here are some key questions to answer, especially when evaluating the vendors and partners that will help turn your vision into reality.

- **Consider the financial roadblocks to your project.** What is the best way to phase in your investments? Are providers that you are evaluating ready to go through thorough proof of concept testing (including onsite), to stand behind their commitments? What is your evaluation strategy and proof of value plan?

- **What are your budgetary constraints?** These can range from when other assets may be fully depreciated, to where there may be pressures on upfront costs driving an interest in on-demand/as-a-service consumption?

- **Identify where the efficiencies and savings will come from in your investment.** How will data infrastructure consolidation unlock your siloed data? Where will it streamline the management of your infrastructure or simplify post-investment support and service? Which providers can streamline data migrations and commit to an Evergreen approach where your investments never go obsolete?

- **Anchor your champions and decision makers on total cost of ownership (TCO) vs simply the upfront costs.** Will a solution help you reduce your data center footprint or address resource challenges you have (due to overly complex data management), or drive improved power and cooling efficiencies?

- **Is any downtime associated with storage upgrades, even if you are consuming it as a service?** Or if you're making a traditional infrastructure purchase, will you be able to avoid storage rebuys when you need new capabilities, performance, capacity, or efficiencies? All of these and more are critical considerations for executives beyond upfront investment costs and drive ongoing savings that will not only help in financial justification, but also keep resources available for future innovation.

Summary

Infrastructure went through a significant transformation over the last decade and must undergo an even more significant one now with AI—as well as continued focus on cybersecurity, risk and cost reduction, innovation, and sustainability. Data infrastructure is a crucial part of this transformation, and legacy storage based on old ideas is not up to these new challenges.

It's time for a new approach. Data storage that isn't designed from the ground up to tackle these new issues will hold you back during this transformation and during the transformations to come. You'll need a storage platform that is designed for the challenges you're facing today, and more importantly, for the challenges of tomorrow —to power you successfully through the changes, so your organization can thrive.

To learn more about how your peers are responding to these challenges, read the Innovations Race survey report. Get more ideas and information on your own data future here.

[Learn more about the Future of Data]