

MARCH 2026

The 2026 SSD Crisis: Implications for Enterprise AI Infrastructure

Simon Robinson, Principal Analyst

SSD crisis overview

AI's insatiable appetite for data is driving profound impacts for storage infrastructure and will create significant pricing pressures and supply shortages of storage media through 2026 and beyond. Although much of the initial focus is around pricing and availability of NAND Flash and enterprise solid-state drives (SSDs), shortages are anticipated across the entire storage and memory hardware landscape.

Omdia now estimate SSDs to be almost **13x the price of HDDs on a \$/GB basis through 2026, more than double the ratio in 2023.**

For IT infrastructure leaders, the timing of these industry pressures is less than ideal. Organizations globally are looking to ramp up their AI investment, and data growth is accelerating. At first blush, these challenges threaten to impact an organization's AI deployment plans. Yet, this isn't inevitable. By taking a more strategic view of the broader data and storage environment, and making some smart technology decisions, this supply crisis can be an opportunity for organizations to build a modern enterprise data architecture that is optimized for the AI era.

Analysis: The perfect storm is driving a new kind of storage shortage

Significant supply constraints in the memory and storage market emerged quickly in late 2025, and the resulting impact on storage media prices will be a significant factor driving storage infrastructure decisions in 2026 and beyond. Though supply issues are not uncommon, this particular shortage looks different: a perfect storm of factors are combining to create unprecedented challenges. In essence, limitations in the supply of NAND Flash have coincided with a massive spike in demand for fast storage, driven primarily by the massive data requirements of AI and accelerated computing, particularly as the popularity of foundational LLM-based services from the likes of Google, OpenAI and Microsoft increases, and as capabilities such as agentic AI grow increasingly sophisticated.

To address this booming demand, hyperscalers and other large cloud and AI providers have placed enormous orders of DRAM and NAND Flash, significantly impacting overall supply and costs of all forms of memory and storage, including SSDs, hard disk drives (HDDs), and controllers. Multiple NAND originators have already stated that their entire capacity for 2026 has been reserved, and most have stated they have no plans to significantly boost production capacity.

These challenges are validated by research findings and data from Omdia, which anticipates rapid price increases in enterprise SSDs (eSSDs) over the coming quarters, and very likely into 2027, as well as price increases for HDDs.¹

Multiple NAND originators have now stated that their entire capacity for 2026 has already been reserved.

Notably, the relative cost of SSD capacity compared to HDD on a \$/GB basis is now widening at a considerable rate; a reversal of the prevailing trend over recent years. Omdia now estimates eSSDs to reach almost 13x the price of enterprise HDDs on a \$/GB basis through 2026, more than double the ratio of under 6x experienced in 2023. This gap is likely to widen further over the coming months.

Additionally, in a 2025 Omdia survey of IT decision-makers, three-quarters of respondents said they anticipated their data growth rate over the next 24 months to be higher or significantly higher than over the previous 12 months. The top drivers for data growth were AI initiatives creating more data, and the requirement to retain this data for longer (see Figure 1).²

Consequently, the next phase of the AI revolution will not only fundamentally reshape the data infrastructure underpinning the handful of hyperscalers building gigantic foundational models but will have a meaningful knock-on effect for all organizations looking to build and deploy storage infrastructure at any level of scale, for any workload.

Defining a viable path forward and why data architecture matters

This supply crisis presents challenges for IT leaders charged with building a data infrastructure that can fully take advantage of new technologies and innovation, particularly around AI and high-performance computing. Some might feel they are simply forced to accept much higher costs, but this might reduce or even destroy the project's ROI. Deferring or even cancelling an AI initiative or new workload deployment is similarly undesirable.

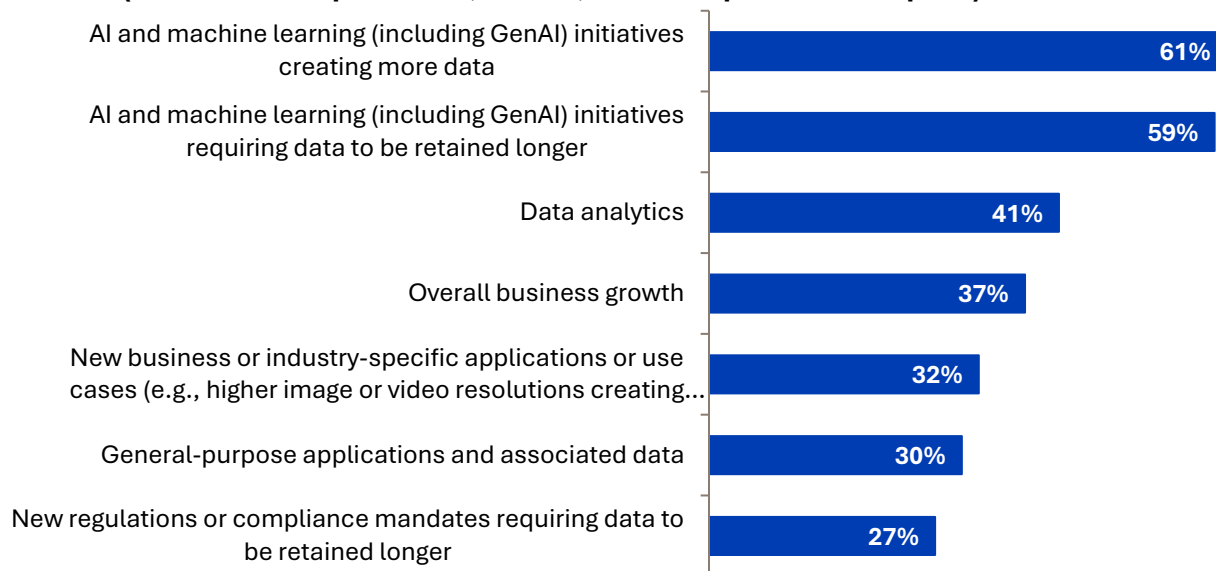
There is another option, however, and that is for IT infrastructure leaders to take a more strategic view of the overall data storage environment. By utilizing innovative software technologies that span their broader storage footprint—such as Hammerspace—organizations can more effectively leverage all of their storage assets, across both Flash and HDD media, and, crucially, tap into under-utilized capacity they have already deployed and paid for.

¹ Source: Omdia Brief, [SSD & HDD Storage Market Dynamics and Pricing Brief – January 2026](#), January 2026.

² Source: Enterprise Strategy Group (now Omdia) Research Report, [The Critical Role of Storage in Building an Enterprise AI Infrastructure](#), September 2025.

Figure 1. Top factors driving data growth

What are the top factors that will drive data growth at your organization over the next 24 months? (Percent of respondents, N=350, three responses accepted)



Source: Omdia

Recommendations for successfully navigating the SSD crisis

Though the SSD crisis presents IT leaders with a potential challenge, especially when faced with rapidly growing data volumes, it also offers an opportunity to consider alternate approaches that may actually help to more effectively manage storage costs, as well as their broader data environment, over the longer term. In an era of falling storage costs and plentiful supply, it is too easy to “throw more hardware” at a capacity issue. Though this may alleviate a particular requirement in the short term, in reality, it drives down overall storage utilization, increases storage fragmentation, and compounds already high levels of storage complexity.

Moreover, it makes the task of managing data in the aggregate much more difficult. This is exactly the wrong moment to be facing such challenges since success in the AI era requires organization to be able to understand, access, and orchestrate actions across all their data in a fundamentally more fluid and dynamic manner.

So, aside from purchasing large amounts of (much more expensive) new capacity, what practical steps can organizations take today to both satisfy their short-term storage capacity demands as well and transform to a more-strategic data architecture posture? Omdia recommends evaluating the following approaches:

- **Make your existing storage assets work harder.** Organizations typically have significant amounts of deployed storage capacity that is already paid for but is not fully utilized. Spanning multiple media and deployment types (e.g., DAS, NAS, object), this underused capacity can amount to a substantial total that may enable organizations to partially or completely avoid the need to buy more capacity if it can be easily made available to additional workloads and applications. Notably, organizations that have deployed large amounts of GPU servers may find substantial amounts of underutilized NVMe storage

within those servers. This could be harnessed in aggregate to provide both capacity and high-performance storage for AI workloads.

- **(Re)consider tiering to free up SSD capacity.** Tiering data across multiple media types, such as SSDs and HDDs, is certainly not a new notion but has perhaps fallen out of favor in recent years as organizations began to standardize on all-flash architectures. With SSD prices vastly outpacing HDDs, the ability to move older or colder data from SSDs to a more cost-effective alternative may be a compelling way to free up existing SSD capacity, potentially avoiding or delaying the need to buy additional flash-based capacity. As well as tiering to lower-cost alternatives on premises, such as HDD-based object storage or NAS, organizations should also consider low-cost cloud-based solutions.
- **Make every stored bit count, with efficient technologies.** The digital revolution has seen data volumes explode, but these capacity challenges are often exacerbated by unfettered growth of data copies unnecessarily consuming capacity and further pushing up costs. AI workloads can further add to the challenge here, as additional copies of data are required to populate models in multiple venues (e.g., in the cloud). Accordingly, as well as utilizing well-established data reduction technologies such as deduplication and compression, organizations should consider global namespace technologies that enable them to access data from multiple locations without creating additional copies.

Conclusion

Massively constrained storage supply and soaring costs might be an unintended consequence of the AI boom, but the current crisis is forcing the broader industry to revisit long-standing assumptions around the economics of the storage industry; namely, that supply will always be plentiful, and prices will only fall. For IT leaders, the phrase often attributed to Winston Churchill, “Never let a good crisis go to waste,” seems particularly relevant in this context. Rather than taking a tactical approach to data storage that exacerbates the challenges created by underutilized data silos, perhaps this is the time to take a more strategic view.

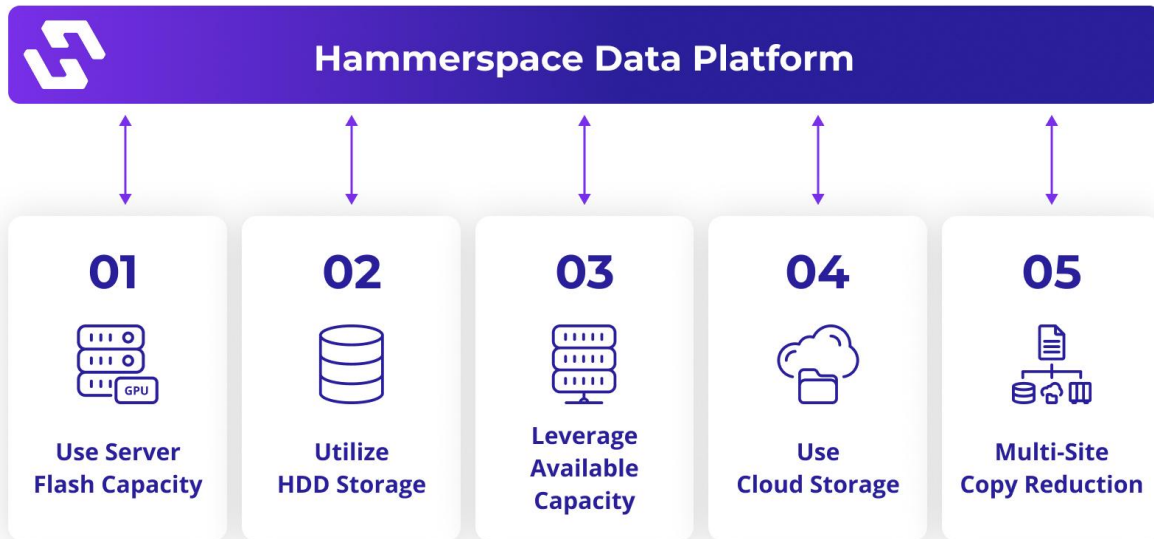
By embracing innovative software, developed by companies such as Hammerspace, that transform these silos of trapped capacity into a unified data pool, organizations will likely find they are better positioned to navigate the current supply crisis: deferring or even avoiding new Capex purchases, reducing dependence on volatile components, better utilizing existing capacity, and all the while not compromising on their AI ambitions. Indeed, as data continues to demonstrate its pivotal role in determining AI success, organizations might also be able to leverage storage as a critical enabler, rather than having to overcome it as a constant hindrance.

About Hammerspace

Hammerspace is a high-performance data platform for AI, HPC, and unstructured data, and it can help IT organizations survive the SSD crisis in five unique ways (see Figure 2). Learn more about how Hammerspace can help and get your survival guide [here](#).

Figure 2. Overview of Hammerspace data platform

Access all data via a Global Namespace using standard file and object protocols.
Parallel file system architecture delivers performance and scale for AI/HPC.



Source: Hammerspace

Copyright notice and disclaimer

The Omdia research, data, and information referenced herein (the “Omdia Materials”) are the copyrighted property of TechTarget, Inc. and its subsidiaries or affiliates (together “Informa TechTarget”) or its third-party data providers and represent data, research, opinions, or viewpoints published by Informa TechTarget and are not representations of fact.

The Omdia Materials reflect information and opinions from the original publication date and not from the date of this document. The information and opinions expressed in the Omdia Materials are subject to change without notice, and Informa TechTarget does not have any duty or responsibility to update the Omdia Materials or this publication as a result.

Omdia Materials are delivered on an “as-is” and “as-available” basis. No representation or warranty, express or implied, is made as to the fairness, accuracy, completeness, or correctness of the information, opinions, and conclusions contained in Omdia Materials.

To the maximum extent permitted by law, Informa TechTarget and its affiliates, officers, directors, employees, agents, and third-party data

Get in touch: www.ondia.com askananalyst@ondia.com

