Successful AI Applications Rely Heavily on Making the Right Storage Decisions

September 2020

Artificial intelligence has become a transformative technology that enables enterprises to unlock huge economic value while optimizing the use of existing IT and business resources. Storage is an important requirement for bringing AI applications to market, so decision-makers need to do their homework on the right storage solutions and suppliers.

Artificial intelligence is a game-changing technology that is no longer the next new thing talked about at conferences or residing in the fertile imaginations of data scientists. Instead, it is a practical, pragmatic and actionable tool that is creating substantial, demonstrable economic value.

Management consultancy PwC puts the total economic impact of AI at an eye-popping $15.7 trillion between now and 2030,[1] and research from McKinsey found that most enterprises already are using AI in some capacity and even more plan to increase their AI investments over the next few years.[2]

AI's growing footprint on the global enterprise landscape has been supported by important technology advances in hardware and software, including storage. Many exciting new AI applications in industries ranging from gaming to gene splicing have been able to deliver substantial economic and operational value in no small part due to the availability of fast, resilient, intelligent and scalable storage.

The sheer size of AI data sets—combined with the need for reliability, scalability, responsiveness and availability—makes the choice of data storage platforms a critical step in your organization's journey to truly transformative AI applications.

What transformative AI applications have in common

AI, in both its deep learning and machine learning iterations as well as others, is all about data. AI takes massive amounts of data—often seemingly disparate types and formats of data from many

---

1 "Sizing the Prize—Global Artificial Intelligence Study: Exploiting the AI Revolution," PwC, 2017

2 "AI-Based Storage Helping Companies Get More Out of Their Data," ITPro Today, March 10, 2020

different sources—and learns human behavior and reactions to deliver actionable insights in the form of groundbreaking applications.

To do that, organizations must have the right data storage solutions in place to support AI applications. Those solutions must deliver key functionality such as:

- High capacity that is easily and quickly scaled as requirements expand.

- Very high bandwidth that can support shared storage across multiple servers.

- Low latency to process very large data sets without expensive and inefficient bottlenecks that cause unnecessary wait times to access and process data.

- Powerful data management features such as deduplication, compression, immutable snapshots and more.

- Reliable and fast data protection.

- Tight integration with on-premises and cloud infrastructure.

Using the right storage helps organizations across industries develop AI applications that make a real difference. For instance:

- **Healthcare.** Real-world applications for the healthcare sector include clinical decision support, high-resolution radiological imaging, analysis of cardiac disease risk factors and post-surgery monitoring of patients in rehabilitation and other facilities.

- **Financial services.** Data scientists and quants (quantitative analysts) are becoming rock stars in the financial world, using AI to do everything from automated high-frequency trading, portfolio analysis and predictive modeling to deep sector analysis.

- **Automotive/connected vehicles.** Whether you're talking about driverless long-haul trucking, autonomous cars or intelligent fleet management, AI processes, analyzes and takes real-time actions based on very large data sets collected in real time from vehicle-based computer systems and road sensors.

Of course, these use cases are just a smattering of exciting new AI applications, with untold numbers still to be developed. But what they all have in common is a need for world-class storage solutions that deliver capacity, performance, low latency, resilience, security and scalability.

Storage for AI training and AI inference

Two areas where storage is key in AI application development are AI training and AI inference. Data scientists and AI developers may be more familiar with these terms than IT practitioners, so let's look at a brief overview of each and why storage is important.

In AI training, a machine learning algorithm is created as the foundation for a framework. That framework is used by data scientists and software engineers to build and train AI or machine learning models; in some sense, the model learns the behavior of what it is modeling. In contrast, AI inference uses an AI-trained machine learning algorithm to build prediction models for advanced decision-making.

In both cases, having a lot of high-capacity, lightning-fast and resilient storage is essential. AI training algorithms and AI inference engines are at the heart of building AI applications, and training and inference are greatly enabled by having the right storage infrastructure.

> Storage solutions for AI applications are being asked to do more than they did even just a few short years ago.

What storage solutions for AI applications must deliver

Storage solutions for AI applications are being asked to do more than they did even just a few short years ago. To feed AI engines with lots and lots of data at blinding speeds, data sets and training batch sizes are growing larger and more unstructured. Storage for AI must also be delivered with the highest possible reliability, with absolute data integrity and the ability to scale capacity to previously unimagined capacities.

Core technical requirements for storage for AI applications include:

- Very high capacity that can be easily and quickly scaled as new AI requirements dictate and as data sets and batch sizes become larger than ever.

- All-flash or hybrid flash arrays supporting hundreds or even thousands of drives.

- Scale-out storage delivering high IOPS.

- High throughput for shared storage delivering data to multiple GPU servers for data acceleration.

- Feature-rich data protection, enabling data protection, backup, recovery, archiving, snapshots, deduplication, compression and more—all automated and aligned with flexible policy management.

- Support for a range of protocols, including Fibre Channel, Ethernet, Serial Attached SCSI (SAS) and InfiniBand. Ideally, this is handled through support for NVMe over Fabrics, running on the high-speed PCIe bus.

Of course, having great storage technology is not enough. It is critical for enterprise decision-makers to evaluate the supplier behind the storage solution to ensure it has the R&D resources, financial staying power, technology vision and broad ecosystem of hardware and software partners to deliver a full range of storage hardware and software to support AI applications.

Lenovo storage solutions for AI applications

To meet the demanding and often variable storage requirements for developing and deploying AI applications, organizations need to select a storage solutions partner with a broad and deep product line, a proven track record, substantial R&D resources and an understanding of how infrastructure properly supports AI applications.

Lenovo is a global leader in data storage, delivering a range of hardware and software solutions for small and large AI initiatives alike. Lenovo's flash storage, storage management software and purpose-built turnkey systems help enable AI applications that take advantage of very large, complex data sets.

Lenovo's AI storage portfolio centers on the DM series of hybrid flash solutions. DM solutions use a unified, scale-out architecture that supports both file and block storage, delivering high performance, low latency and massive capacity scalability. Lenovo DM storage is designed for nondisruptive expansion, giving organizations the ability to add and replace arrays and components without disturbing production environments.

DM storage also comes with MetroCluster, a data protection tool that synchronously mirrors data between locations for nonstop data availability—a must for many large-scale AI development projects that cannot afford the risk of storage downtime. And, since many AI applications are developed and deployed in the cloud, DM storage solutions are tightly integrated with the industry-standard OpenStack cloud management platform.

> Lenovo offers storage and other hardware solutions across a wide range of price points and functionality, organizations benefit from the "democratization of AI."

Lenovo's ONTAP data management platform is ideal for AI applications because of its advanced data handling capabilities, including data protection, data migration, storage tiering, archiving and restore. The Cloud Volumes ONTAP platform seamlessly integrates with and replicates data to multiple clouds, including Amazon Web Services, Microsoft Azure and IBM Cloud.

Lenovo also offers a useful reference architecture for AI training workloads, built on Lenovo ThinkSystem entry-level compute and storage infrastructure, along with high-performance GPUs aligned with traditional CPUs. The architecture is pre-tested and validated, supporting scale-out storage and optimized data management to streamline AI workload development.

Additionally, Lenovo's Distributed Storage Solution for IBM Spectrum Scale (DSS-G) is a software-defined storage architecture for scalable file storage and object storage often used in AI application development and other high-performance computing environments. DSS-G solutions can meet a range of requirements across organizations, from department-level research teams to entire enterprises requiring hundreds and hundreds of petabytes of capacity and high bandwidth from their parallel file systems.

Another important point to keep in mind is that, since Lenovo offers storage and other hardware solutions across a wide range of price points and functionality, organizations benefit from the "democratization of AI," meaning they can build relatively lower priced solutions initially but then rapidly and broadly scale them over time.

The great thing about AI is that it doesn't require organizations to boil the ocean: AI applications can start relatively minor in scope and grow as conditions and opportunities allow, and Lenovo storage solutions are available for even small, budget-limited AI development projects.

Summary

As AI applications expand in number and in their ability to drive business value, enterprises need to ensure they have the right storage infrastructure in place to deliver the capacity, performance, resilience and intelligence to support AI initiatives.

That means organizations looking for solutions that enable the development of transformative applications using AI, deep learning and machine learning should select storage options from a proven, reputable technology partner that understands how to deliver best-of-breed solutions. Lenovo provides a full line of storage arrays in both all-flash and hybrid-flash configurations, as well as related technologies such as servers and data management platforms.

For more information on how Lenovo storage supports important AI applications, please visit: www.lenovo.com/us/en/data-center/storage/c/storage

7

This content was commissioned by Lenovo and produced by TechTarget Inc.