

SOLUTION BRIEF

AIRI AI Infrastructure Re-Imagined, Optimized, Ready for Enterprise

Power your AI initiatives with the industry leading efficiency of AIRI[®], NVIDIA DGX BasePOD certified AI Reference architecture that delivers enterprise level performance, simplicity and scalability.

The use of artificial intelligence (AI) is critical to being competitive in many different industries. As one of the fastest growing and largest areas of enterprise technology investment and innovation, AI supports use cases in industries, ranging from healthcare and life sciences to energy, telecommunications, semiconductor and chip manufacturing, automotive, financial services, and beyond.

“Do-It-Yourself” Causes Roadblocks

The IT infrastructure historically used to support legacy data science efforts can't provide access to the right data fast enough to unleash the advantage of GPU-powered AI-based data science. Cloud is an option, but as AI projects grow to production scale, it can be cost prohibitive. Organizations are looking for solutions to enable AI simply, quickly, and cost effectively and storage is often a bottleneck.

When implementing a DIY approach to AI infrastructure, organizations contend with a variety of roadblocks, including:

- **Unexpected hardware expenses** due to the high-performance, dedicated, and optimized hardware required to successfully process each stage of the AI workflows and store massive amounts of data. Legacy servers are not up to the task.
- **Complex AI software** can consume developer cycles spent on software stack engineering, troubleshooting, and optimization.
- **Long deployment times** and operational complexity of AI infrastructure make achieving ROI on AI initiatives difficult at best.



AIRI Technology Stack:

Pure Storage FlashBlade//S™

Pure Storage Software:
Purity//FB OE, Pure1®
Management, Portworx
Kubernetes Data Platform, Pure
RapidFile Toolkit.

NVIDIA DGX H100, NVIDIA DGX
A100, NVIDIA DGX BasePOD

NVIDIA Spectrum and Quantum¹
Networking

NVIDIA Software: NVIDIA AI
Enterprise, NVIDIA Base
Command, NVIDIA NGC™,
CUDA-X, Magnum IO, RAPIDS

AIRI Simplifies AI-at-scale

Reduces the complexity of
integration with a proven
complete DGX BasePOD
Certified solution.

Data scientists can focus on
algorithms and outcomes, not
infrastructure.

AI-at-scale Is an Advantage

More compute = faster training
More data = higher accuracy
AIRI makes it simpler and faster
to run multi-node training



Uncomplicate Your Data

AIRI, a complete AI infrastructure solution, certified on the NVIDIA DGX BasePOD™ reference architecture, is the latest evolution of the industry's first complete AI-ready infrastructure. Architected by Pure Storage® and NVIDIA, and consisting of the latest NVIDIA DGX™ systems, NVIDIA networking and Pure Storage FlashBlade//S, AIRI is a complete hardware and software solution that maximizes AI results with the industry's most efficient scale-out storage platform. DGX BasePOD is a prescriptive AI infrastructure for enterprises, eliminating the design challenges, lengthy deployment cycle and management complexity traditionally associated with scaling AI infrastructure. DGX BasePOD certification ensures a validated, proven full-stack solution that is simple-to-use, enabling you to reap the benefits of AI right away with a fast and efficient infrastructure that meet your AI demands at enterprise scale.

Powered by NVIDIA Base Command™ software, DGX BasePOD provides the essential foundation for AI development optimized for enterprise. NVIDIA Base Command is the operating system of the DGX data center, helping organizations speed the ROI of AI. Base Command powers the DGX platform, enabling organizations to leverage the best of NVIDIA software innovation. Enterprises can tap into the full potential of their AI infrastructure with a proven platform that includes workflow management, enterprise-grade cluster management, libraries that accelerate compute, storage architected for AI, network infrastructure, and system software optimized for running AI workloads.

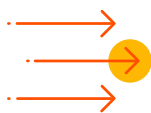
FlashBlade//S is the ideal data storage platform for AI, as it was purpose-built from the ground up for modern, unstructured workloads and accelerates AI processes with the most efficient storage platform at every step of your data pipeline. A centralized data platform in a deep learning architecture increases the productivity of data scientists and makes scaling and operations simpler and more agile for the data architect.

Delivering the Next Generation AI Platform



Simplify AI Infrastructure

AIRI is a pre-validated reference architecture with full stack integration. It can be set up, deployed, and managed quickly as an end-to-end AI workflow solution. Enable your data teams to be more productive, by spending more time delivering insights and less on waiting for results or worrying about their infrastructure.



Maximize Efficiency and Productivity

Pure Storage DirectFlash technology maximizes storage efficiency and training performance. Get higher capacity in a smaller physical and power footprint, while maximizing DGX and storage performance. Predictable high performance AI infrastructure shortens training, enables you to iterate faster, and minimizes time-to-insight.



Get Better Economics for All of Your AI Data

Lower TCO by consolidating data silos, providing performance wherever needed within your AI process and providing all-flash features for AI content repositories at disk economics. A simple and consistent experience for all of your AI data means you can reduce operational expenses in addition to saving costs from efficiency and productivity benefits.



AIRI Technology Stack

AIRI represents a combination of groundbreaking solutions from two industry leaders. All-flash storage architected for AI from Pure Storage is paired with the latest AI hardware and software from NVIDIA. AIRI is optimized out-of-the-box and can seamlessly scale storage and AI compute performance, allowing organizations to begin their AI journey at any scale and grow and adapt as their needs evolve.

NVIDIA Base Command software and NVIDIA AI Enterprise software coupled with the Purity//FB Operating Environment and Pure RapidFile Toolkit provide a strong software foundation to jump-start and accelerate your AI journey.



Pure Storage: Enabling AI Success with NVIDIA

Pure Storage has worked with NVIDIA to help organizations exceed expectations in their AI initiatives across a variety of industries.

The following are just a few of the areas Pure and NVIDIA have enabled AI success:

- Advancing the national security mission of the United States and other nations' intelligence agencies (private references).
- Increasing data processing speeds by 2x and expanding GPU usage from 30% to 80%, maximized efficiency and management of data used to develop and train [Chunkbuk Technopark's](#) AI models.
- Accelerated genetic research and medical imaging 7x at the Center for AI in Medicine at [Chang Gung Memorial Hospital](#).
- Accelerated medical image processing which increased customer satisfaction and market competitiveness at [Olympus](#), a leading MedTech provider.
- Advanced AI analysis of large data sets at [NavInfo](#) to produce real-time maps and navigation models used by car manufacturers worldwide.
- Delivered significant expansion in the volume data science research projects for the [Gwangju Institute of Science and Technology \(GIST\)](#).
- Accelerated EDA to support the pace of chip development where for each nanometer a die is reduced in size, the volume of data required to be processed in the development almost doubles from the prior size (private references).

Learn More

AIRI from Pure Storage powered by NVIDIA accelerates and simplifies the process of deploying and running complex AI infrastructure, so our customers can focus on their core mission and discover critical insights and outcomes faster.

[Explore the AIRI solution](#)

purestorage.com

800.379.PURE

