PURESTORAGE®
Uncomplicate Data Storage, Forever

NVIDIA. | NVIDIA ELITE PARTNER

High-Performance AI for State and Local Government

Simple, scalable AI infrastructure from NVIDIA and Pure Storage.

Artificial intelligence (AI) has the potential to transform state and local government. We're already seeing its potential to improve public safety and citizen services, and to protect critical infrastructure. AI also helps agencies drive better, more informed, data-driven decision-making, and can free up millions of labor hours for critical tasks, leading to increased operational and administrative efficiency.

A recent report estimated that AI could boost productivity by $519 billion a year across national, state, and local governments. According to Gartner, Inc., at least 60% of government artificial intelligence and data and analytics investments will directly impact real-time operational decisions and outcomes. AI requires completely new infrastructure, and the complexities of legacy solutions create roadblocks that prevent agencies from tapping into the full potential of the new era of intelligence.

"Do-It-Yourself" Causes Roadblocks

The IT infrastructure historically used to support legacy data science efforts can't provide data access fast enough to unleash the advantage of GPU-powered AI-based data science. Agencies need solutions to enable AI simply, quickly, and cost effectively. When implementing home-grown AI solutions, agencies 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 pipeline and store massive amounts of data. Legacy servers are not up the task.

- **Complex AI software** requiring careful, precise, and frequent configuration by highly specialized developers.

- **Endless maintenance cycles** demanding constant software patching, tuning, and compiling, along with ongoing tweaking of the AI algorithms and frequent hardware upgrades to keep up with the unending deluge of new data.

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 Quantum[1] Networking
- NVIDIA Software: NVIDIA AI Enterprise, NVIDIA Base Command, NVIDIA NGC, CUDA-X, Magnum IO, RAPIDS, MONAI Label and Core, BioNeMo, FLARE, Holoscan

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
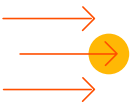
Uncomplicate Your Data

AIRI®, a complete AI infrastructure solution, certified on the NVIDIA DGX BasePOD reference architecture, is the next 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 agency AI results with the industry's most efficient scale-out storage platform. DGX BasePOD is a prescriptive AI infrastructure for organizations, 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's simple to use, enabling you to reap the benefits of AI right away with a fast and efficient infrastructure that meets your AI demands at enterprise scale.

Powered by NVIDIA Base Command software, DGX BasePOD provides the essential foundation for AI development. NVIDIA Base Command is the operating system of the DGX data center, helping agencies speed the ROI of AI and enabling them to leverage the best of NVIDIA software innovation. Agencies 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.

For full technical details, please consult the AIRI Pure Storage NVIDIA DGX BasePOD Reference Architecture.

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 time 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.
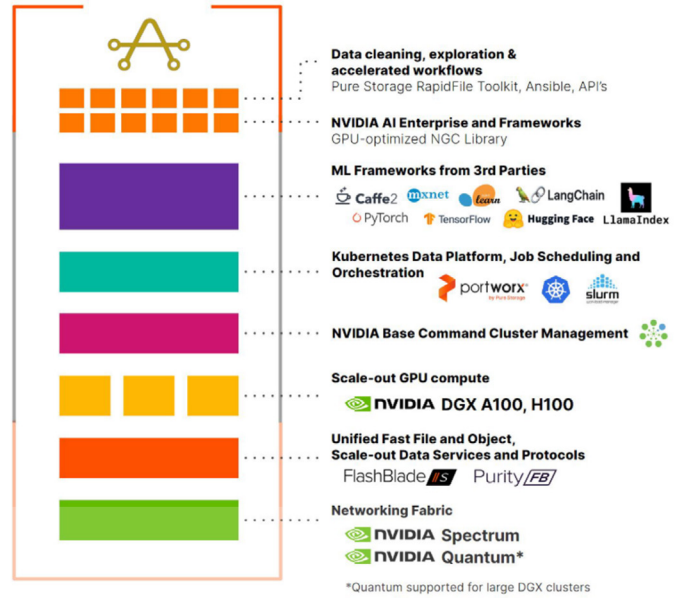
Get Better Economics for All of Your AI Data

Lower TCO by consolidating data silos. Gain performance wherever needed within your AI process and get 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 agencies 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 Storage RapidFile Toolkit provide a strong software foundation to jump-start and accelerate your AI journey.

Data cleaning, exploration & accelerated workflows
Pure Storage RapidFile Toolkit, Ansible, API's

NVIDIA AI Enterprise and Frameworks
GPU-optimized NGC Library

ML Frameworks from 3rd Parties
Caffe2, mxnet, learn, LangChain, PyTorch, TensorFlow, Hugging Face, LlamaIndex

Kubernetes Data Platform, Job Scheduling and Orchestration
portworx, slurm

NVIDIA Base Command Cluster Management

Scale-out GPU compute
NVIDIA DGX A100, H100

Unified Fast File and Object, Scale-out Data Services and Protocols
FlashBlade//S, Purity//FB

Networking Fabric
NVIDIA Spectrum
NVIDIA Quantum*

*Quantum supported for large DGX clusters

Maximum Performance for Key Agency Use Cases

The combination of the latest hardware and software from NVIDIA and the massive parallelism of FlashBlade//S, delivers maximum performance for key government agency use cases such as:

- **Bolstering Cybersecurity:** Automating prevention, protection, and investigation of threats, safeguarding critical infrastructure through integration of AI into cybersecurity efforts, and enhancing response to incidents and threats.

- **Optimizing Public Safety:** Analyzing complex data sets to uncover crime patterns and aid in predictive policing, leveraging AI-enhanced video surveillance data to assist in monitoring and analyzing vast volumes of footage, and providing other law enforcement and first responder insights.

- **Propelling Data Analytics:** Accelerating data collection and analysis, for quicker insights and more informed decision-making.

- **Simplifying Workflows:** Automating and streamlining repetitive tasks such as invoicing, inventory management, data entry, and document processing; reducing paperwork; and improving response times.

- **Leveraging Predictive Analytics:** Analyzing extensive datasets to make predictions, optimize services, and initiate appropriate actions, such as forecasting energy consumption patterns for optimal resource planning.

- **Improving Operations:** Leveraging AI for expenditure analysis, process enhancements, and budget planning.

- **Enhancing Citizen Services:** improving services by enhancing efficiency, accessibility, and responsiveness through chat bots, voice assistants, process and form automation, and protecting citizen data and transactions.

AIRI can optimize all these agency use cases and many more. With virtualized GPUs and petabyte-scale storage, AIRI can support multiple use cases on a single platform.

Sustainable AI Adoption

AI adoption is on the rise across agencies, yet most organizations lack the necessary infrastructure to handle the high-performance data demands and energy requirements essential for maximizing its benefits. In fact, legacy systems often cannot support the massive AI data pipelines required to get the most from machine learning models. And while AI brings immense promise, its impact on energy requirements can be surprising. These hidden costs of AI pose a challenge to the successful implementation of critical government initiatives, including those aimed at achieving environmental goals.

Pure Storage conducted a survey that found that for 88% of those who have adopted AI, the need for computing power has surged dramatically. In fact, nearly three in four IT buyers (73%) whose organizations have implemented AI were not completely prepared for the energy requirements. As a result, nearly all respondents (96%) already have, or plan to, update their IT infrastructure, including 29% who say it has or will require a complete overhaul.

Right now, the vast majority of data in data centers—over 80%—remains trapped on magnetic hard disks. Shifting to flash-optimized systems is a preliminary step organizations can take to reduce power usage between five and ten times. Pure Storage arrays use up to 85% less power than competitive all-flash systems. Pure Storage delivers a unified all-flash platform that is more sustainable than any other enterprise data storage technology and is the only vendor dedicated to reducing data center carbon emissions and energy footprints, absorbing customers' power and rack space costs.

Accelerate and Simplify AI Infrastructure for Faster Mission Outcomes

As AI continues to accelerate in adoption, IT teams require a meaningful data strategy to ensure they can efficiently and effectively operationalize AI through the right infrastructure. AIRI accelerates and simplifies the process of deploying and running complex AI infrastructure, allowing agency data teams to concentrate on their core mission, gain crucial insights, and achieve faster outcomes.

Additional Resources

- Learn more about the AIRI solution.
- Read the analyst report: Meeting the Energy and Data Challenges of AI Adoption.
- Watch the NVIDIA and Pure Storage joint webinar on Data Center Modernization for Scaling AI.
- Explore Pure Storage solutions for state and local governments.

1 Quantum networking supported for larger DGX clusters

purestorage.com

800.379.PURE

PURESTORAGE®
Uncomplicate Data Storage, Forever

PS2579-01-en 03/24