

MEET AIRI™

YOUR FIRST AI-READY INFRASTRUCTURE AT SCALE

THE ERA OF AI IS HERE

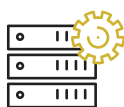
Enterprises will tap the power of AI for faster innovation and a competitive edge. Yet AI requires a completely new infrastructure, and the complexities of legacy solutions are holding enterprises back from moving into the new era of intelligence.

“DO-IT-YOURSELF” OFTEN YOUR ONLY OPTION

AI requires new, modern technologies like GPUs, scale-out flash, and RDMA fabric to move tremendous amounts of data. Unfortunately, too many AI initiatives are stalled with the complexities of a “do-it-yourself” approach using legacy technologies, leading to months of delays and idle time.



NEVER-ENDING CYCLES OF
COMPILING & TUNING OPEN
SOURCE SOFTWARE



MONTHS OF SYSTEM BUILDING
AND TUNING, CONSTANT
MAINTENANCE



LEGACY SOLUTIONS FULL OF
DATA BOTTLENECKS, FROM
STORAGE TO GPU TO APPS

HYPERSCALE MADE SIMPLE AND FAST

AIRI™ is the industry’s first complete AI-ready infrastructure, architected by Pure Storage® and NVIDIA® to extend the power of NVIDIA® DGX™ systems. Powered by FlashBlade™ storage and NVIDIA DGX-1 and DGX-2 servers, AIRI offers enterprises a simple, fast, and future-proof infrastructure capable of growing from AIRI “Mini” to hyperscale – and meeting AI demands at any scale, without downtime.



FROM “MINI” TO HYPERSCALE AIRI

AIRI SIMPLIFIES AI-AT-SCALE

- Reduces racks of complexity into a complete solution
- Data scientists can focus on algorithms, not infrastructure

AI-AT-SCALE IS AN ADVANTAGE

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

AI IS DIVERSE AND EVOLVING

- From CNNs to GANs, AI is constantly changing
- AI is a data pipeline, requiring an integrated platform

50 RACKS UNDER 50 INCHES

AI pushes beyond the reach of legacy technologies like serial CPUs and spinning disks. At the core of AIRI are NVIDIA® DGX-1™ servers and FlashBlade, industry-leading solutions architected for AI. Each replaces the performance of many racks of legacy technology, offering data scientists the power of a large supercomputer for any AI initiative – in a converged solution with a fraction of the physical space and power and cooling costs of legacy infrastructure.

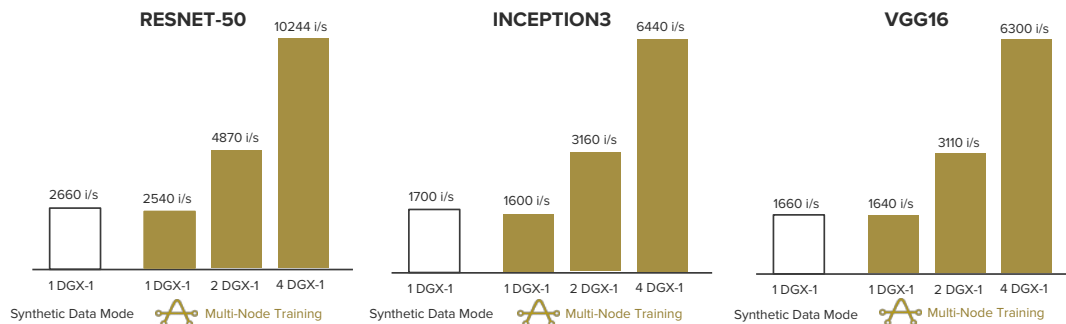


DATA BOTTLENECKS, ELIMINATED

“Do-it-yourself” infrastructure requires constant tuning. As one bottleneck is resolved, another often shows up somewhere else in the system, resulting in weeks to months of lost productivity. AIRI is a complete infrastructure, tuned from software to hardware to keep the GPUs busy for workloads at any scale.

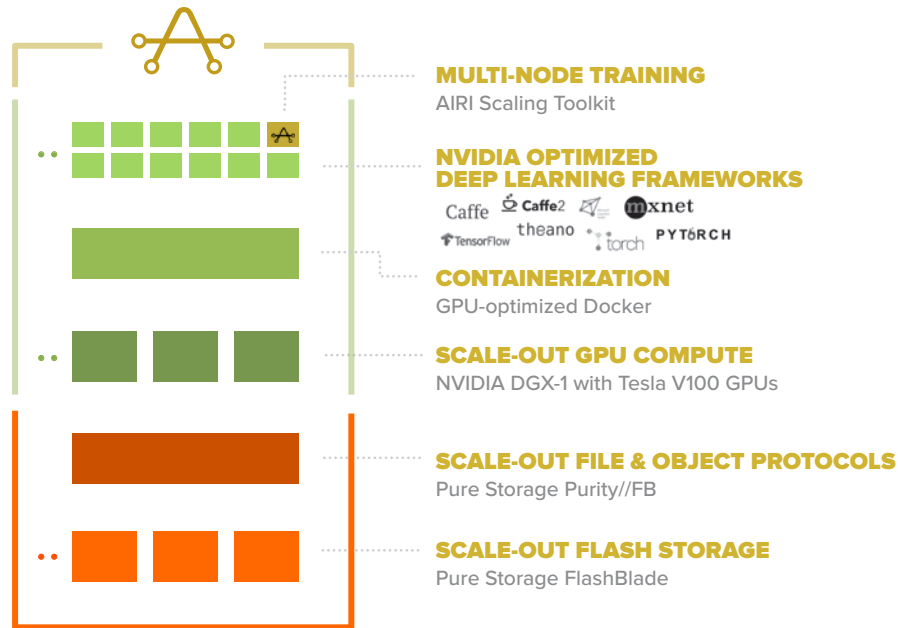
AIRI DELIVERS LINEAR, SCALE-OUT PERFORMANCE

KEEPING GPUs BUSY WITH TENSORFLOW AND 100Gb ETHERNET WITH RDMA



AIRI TECHNOLOGY STACK

EXTENDING THE POWER OF NVIDIA DGX-1 SYSTEMS FOR AI-AT-SCALE



AIRI is built with a **complete software stack** enabling data scientists to get up and running in a few hours, not weeks or months. **AIRI Scaling Toolkit** allows users to run their first multi-DGX training within hours, with a few commands, and to slash training time from weeks to days for their most critical AI workloads.

CUSTOMER HIGHLIGHT

LEADING SOLUTIONS PROVIDER BRINGING AI TO ALL ENTERPRISES

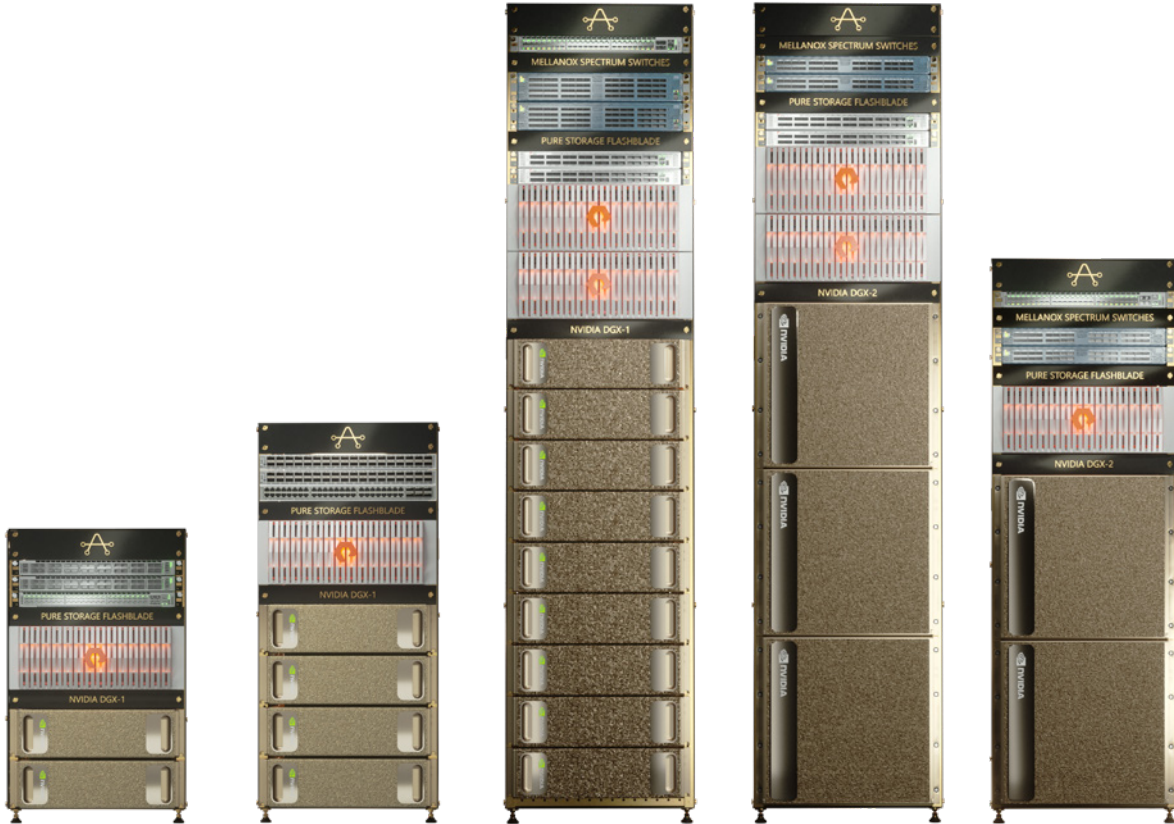
ElementAI is an artificial intelligence solutions provider that gives organizations unparalleled access to cutting-edge technology. It is founded by leading AI experts, including Yoshua Bengio, who is widely considered to be one of the three pioneers of deep learning.

ELEMENT^{AI}

“AIRI represents an exciting breakthrough for AI adoption in the enterprise, **shattering the barrier of infrastructure complexities** and clearing the path to jumpstart any organization’s AI initiative. AIRI is built with the same core solutions ElementAI uses extensively both internally and with customers – the NVIDIA® DGX-1™ and Pure Storage FlashBlade.”

Jeremy Barnes, Chief Architect, ElementAI

SPECIFICATIONS



AIRI "MINI"

2x NVIDIA® DGX-1™
SYSTEMS
2 PFLOPS of DL
Performance

1x PURE
FLASHBLADE
7x 17TB Blades

CONVERGED
FABRIC
2x 100Gb Ethernet or
EDR InfiniBand

AIRI

4x NVIDIA® DGX-1™
SYSTEMS
4 PFLOPS of DL
Performance

1x PURE
FLASHBLADE
15x 17TB Blades

CONVERGED
FABRIC
2x 100Gb Ethernet or
EDR InfiniBand

9x NVIDIA® DGX-1™
SYSTEMS
9 PFLOPS of DL
Performance

2x PURE
FLASHBLADE WITH
MULTI-CHASSIS
30x 17TB Blades

CONVERGED
FABRIC
2x 100Gb Ethernet or
EDR InfiniBand

HYPERSCALE AIRI

3x NVIDIA® DGX-2™
SYSTEMS
6 PFLOPS of DL
Performance

2x PURE
FLASHBLADE WITH
MULTI-CHASSIS
30x 17TB Blades

CONVERGED
FABRIC
2x 100Gb Ethernet or
EDR InfiniBand

2x NVIDIA® DGX-2™
SYSTEMS
4 PFLOPS of DL
Performance

1x PURE
FLASHBLADE
15x 17TB Blades

CONVERGED
FABRIC
2x 100Gb Ethernet or
EDR InfiniBand

VISIT [PURESTORAGE.COM/AIRI](https://www.purestorage.com/airi) TO LEARN MORE

© 2019 Pure Storage, Inc. All rights reserved. Pure Storage, FlashBlade, AIRI, and the "P" logo are trademarks or registered trademarks of Pure Storage in the U.S. and other countries. NVIDIA and DGX are trademarks of NVIDIA, Inc. All other trademarks are the property of their respective owners. ps_sb4p_ai-ready-infrastructure_ltr_03

