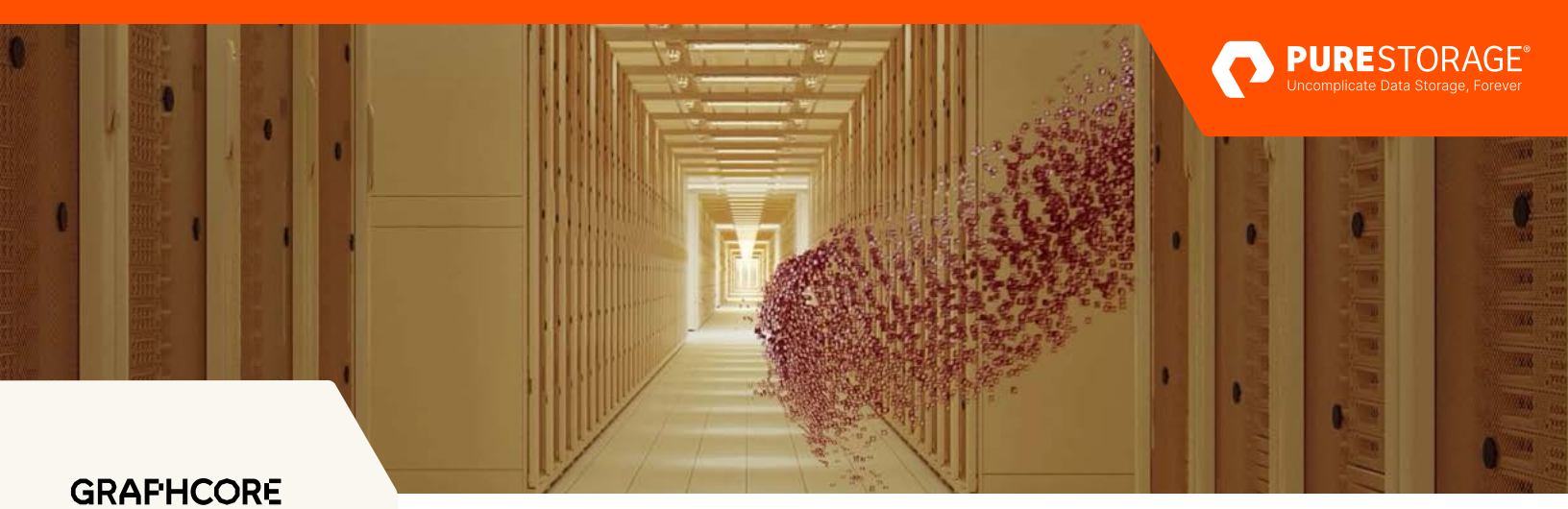

PURESTORAGE®
Uncomplicate Data Storage, Forever

GRAPHCORE

About Customer
Intelligence Processing Unit (IPU) inventor Graphcore is setting new standards for AI compute and helping innovators make breakthroughs in machine intelligence that positively impact everything from drug discovery to decarbonization.
www.graphcore.ai

Geo
EMEA

Industry
Technology

Solution Area
Accelerate Core Applications
Power Artificial Intelligence

Pure Storage® Products in Use
FlashBlade®

Graphcore Breaks the Barriers on AI Exploration

Graphcore's IPUs are disrupting AI, setting the global standard for machine intelligence compute and creating opportunities across everything from drug discovery to decarbonization.

The IPU has a unique architecture specifically designed for AI compute needs. It delivers a huge performance uplift and lets researchers explore new types of work that were previously not possible or functionally challenging using legacy processor architectures such as the GPU.

As a key part of the end-to-end AI optimized technology stack, Pure Storage supports Graphcore's goals to break through the shackles placed on AI exploration.

> "Getting the best from our IPUs demands storage with high performance, low latency, high throughput, and distributed scalability. FlashBlade has delivered on all these fronts."
>
> **STUART CORNELL,**
> SYSTEMS OPERATIONS MANAGER, GRAPHCORE

Impact on Graphcore

Scale on the fly to support new and growing customer demand

Focus on developing world-leading AI accelerators

Rapidly expand presence in new industries

Challenges

AI compute is too complex for legacy storage systems

Storage bottlenecks potentially limit processing performance

Storage integrations can add complexity to IT systems

Results

Delivered high bandwidth, throughput, low latency for AI compute

Simple installation, setup allows IT to set and forget storage

Flexibility to support variety of data types on one platform

Smoothing Access to Data—Innovating With AI at Speed and Scale

The challenge with AI models is that they require a lot of data to train, improve, and refine them. Having the right storage and right configurations is critical to creating an efficient and reliable infrastructure for AI workloads.

Stuart Cornell, Systems Operations Manager at Graphcore, explains, "Graphcore systems deliver extremely high-performance AI compute. To get the most out of our IPU processor, it is important to have storage systems that also meet the unique and demanding requirements of AI, and Pure delivers that."

Delivering a Flawless Storage Experience

Graphcore deploys Pure Storage in Graphcloud—its cloud-based machine-learning (ML) platform that helps customers expand from experimentation and proof of concept projects to production systems. Graphcloud runs on large IPU Pod systems, powered by several hundred IPU processors, which all rely on Pure Storage FlashBlade. The all-flash unified fast file and object (UFFO) platform enables the company to consolidate data across Graphcloud's machine learning operations (MLOps) pipeline to keep the IPU Pods fed.

"Getting the best from our IPUs demands storage with high performance, low latency, high throughput, and distributed scalability," says Cornell. "FlashBlade has delivered on all these fronts. And it's simple. In the best possible way, Pure means I can forget about the storage. This flawless experience also reflects what we want for our customers."

FlashBlade also brings significant power consumption efficiencies to the overall system. FlashBlade has high performance density, delivering an enormous amount of throughput for a small unit. Given Graphcore IPUs are throughput hungry, Pure Storage can replace their legacy storage equipment that would inevitably take up much more space and power.

Engineering for the Next Generation of Compute

Cornell concludes: "Pure is engineering the type of storage technology that AI needs. Graphcore's adoption of Pure brings intelligent compute and storage together to create a converged infrastructure solution capable of serving AI and ML workloads at any scale."

PS2287-02 10/2023