



# From JBOF to Unified AI-ready Enterprise Data Platform

AIC F2026 with BlueField-3 and SupremeRAID™ HE—Unleashing Scalable Gen5 NVMe Performance with Enterprise-Grade Protection

## TABLE OF CONTENTS

Executive Summary .....	1
Architectural Breakthrough .....	2
The Foundation: AIC F2026-01-G5 JBOF .....	2
The Brain: NVIDIA BlueField-3(BF3) DPU .....	3
The Engine: SupremeRAID™ HE .....	3
Architecture Overview .....	4
AIC F2026: Zero-CPU Storage Platform .....	4
Base Configuration – Dual-Canister (2 DPUs + 2 RAID GPUs) ..	4
Scaling to Advanced Dual-Canister (4 DPUs + 4 RAID GPUs) ...	5
Performance Data: Base vs Advanced .....	5
Throughput Performance (1M Random) .....	5
IOPS Performance (4k Random) .....	5
Use Cases: Built for the Modern Datacenter .....	6
Advantage: Why Choose This Solution? .....	6

## Executive Summary

In the era of Generative AI and high-performance computing, legacy storage architectures have become the critical bottleneck. The **AIC F2026 NVMe JBOF**, integrated with **NVIDIA BlueField-3 DPUs** and **SupremeRAID™ HE**, creates a revolutionary AI-Ready Enterprise Data Platform. This architecture eliminates the traditional trade-off between performance and protection, delivering over 8.8 million IOPS and 85 GiB/s throughput in a compact 2U design, while sustaining full RAID 5/6 enterprise-grade resilience. By combining proven reliability with AI-optimized scalability, the F2026 transforms storage into a high-performance enabler for modern AI pipelines.

## The Challenge: Bottlenecks in the AI Pipeline

Modern AI training and inference workloads demand massive throughput while maintaining data integrity. Traditional hardware RAID controllers are a choke point, unable to keep pace with Gen5 NVMe SSDs. Software RAID consumes CPU resources that applications need to run efficiently.

## The Solution: Disaggregation and Offload

By offloading RAID processing to the GPU (SupremeRAID™) and storage networking to the DPU (NVIDIA BlueField-3), the AIC F2026 platform delivers bare-metal NVMe performance with the reliability of an enterprise SAN.

## Architectural Breakthrough

This solution integrates three best-in-class technologies into a unified, high-availability building block.

### The Foundation: AIC F2026-01-G5 JBOF

The F2026-01-G5 is a next-generation 2U, 26-bay PCIe Gen5 NVMe JBOF engineered for high bandwidth, low latency, and enterprise-class availability. Built on the **Broadcom PEX 89144 PCIe Gen5 switch architecture**, it delivers a non-blocking data path within a compact chassis and supports hot-swappable, dual-port Gen5 NVMe SSDs, enabling dense and resilient data access.

With **dual active-active controller canisters** and up to **8 PCIe Gen5 x16 slots for DPUs**, the F2026 offloads I/O-intensive tasks such as NVMe-oF, compression, and encryption from host CPUs to dedicated hardware acceleration engines. This architecture eliminates single points of failure common in legacy designs, while enabling independent scaling of storage and compute with headroom for future acceleration and expansion.



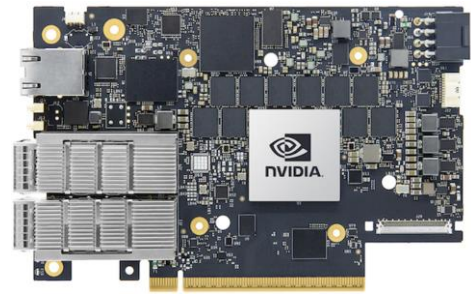
The AIC F2026 is a high-density, dual-controller JBOF designed for the PCIe Gen5 era, engineered for maximum bandwidth and ultra-low latency and enterprise-class resilience.

- **Non-Blocking PCIe Gen5 Performance:** Powered by the Broadcom PEX 89144 Gen5 switch, the dual-path design ensures sustained Gen5 x4 bandwidth across all NVMe SSDs.
- **High-Availability Hardware Architecture:** Equipped with active-active dual controllers (Node-1 / Node-2) and a dual-port PCIe backplane, the system eliminates single points of failure.
- **High-Density NVMe Storage:** Supports up to 26 hot-swappable dual-port Gen5 NVMe SSDs within a compact 2U chassis.

## The Brain: NVIDIA BlueField-3(BF3) DPU

The BlueField-3 (B3220SH) transforms the JBOF into an intelligent NVMe-oF target, enabling advanced storage acceleration and connectivity.

- **Zero-CPU Storage:** Runs the Graid SPDK Target on the DPU's Arm cores, allowing the BF3 to handle all target logic.
- **200G Connectivity:** Provides massive NDR 200Gb/s pipes to feed hungry AI clusters over InfiniBand or Ethernet.
- **Storage Services:** Capable of hosting NAS (SMB/NFS), parallel filesystems, and SDS/Object Storage services.



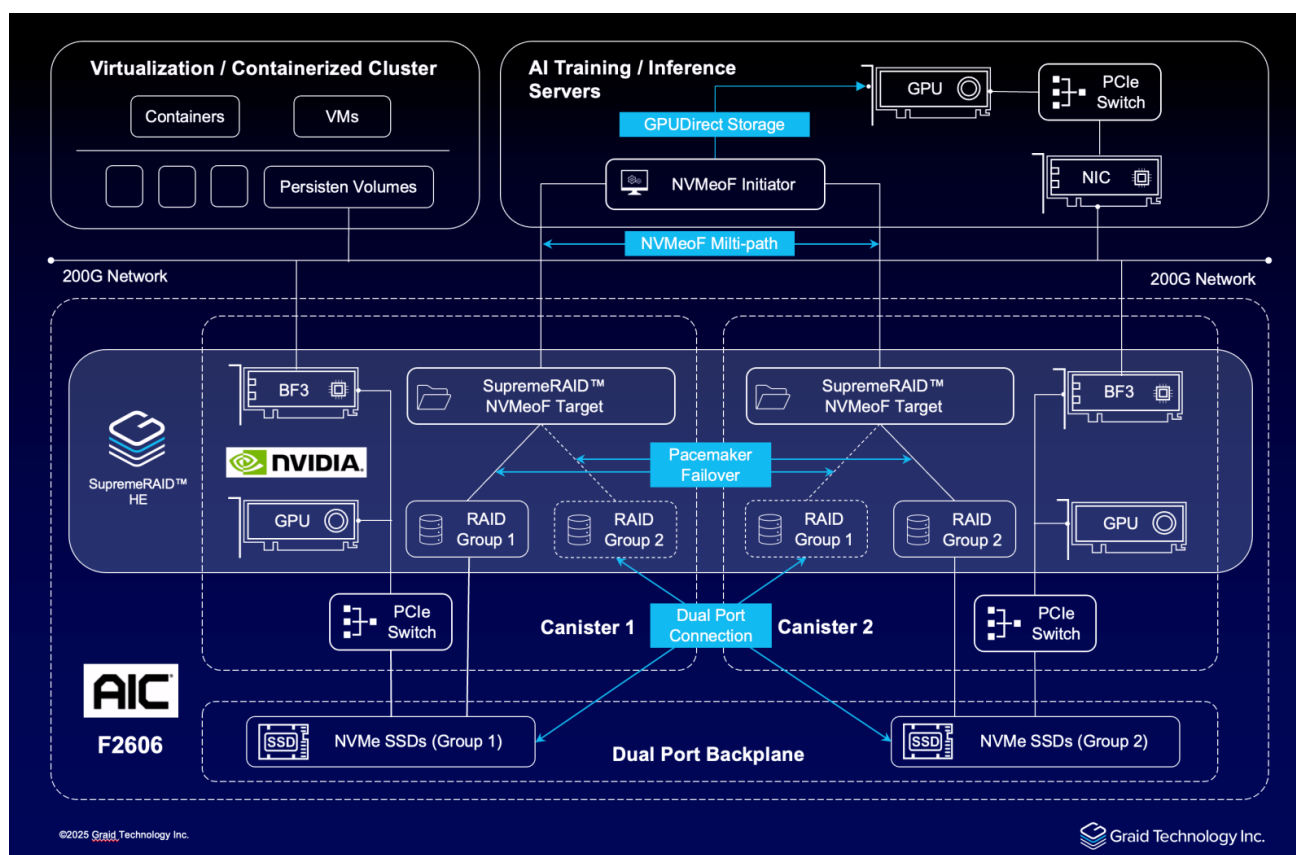
## The Engine: SupremeRAID™ HE

SupremeRAID™ HE (Pro Ampere V2) moves RAID processing from the CPU to a dedicated NVIDIA GPU, delivering accelerated performance and enterprise-class resilience.

- **GPU Models:** Utilizes NVIDIA RTX A1000 (SR-HE-PAM2) for RAID acceleration.
- **Advanced Resilience:** Includes advanced journaling and automatic bad block recovery to safeguard data integrity under extreme workloads.
- **Unmatched Scalability:** Supports RAID levels 0, 1, 5, 6, and 10 with up to 1,023 volumes per array.



## Architecture Overview



### AIC F2026: Zero-CPU Storage Platform

The AIC F2026 is engineered as a Zero-CPU storage platform. Networking and data protection tasks are offloaded from the host CPU onto specialized acceleration engines — NVIDIA BlueField-3 DPUs and SupremeRAID™ GPUs.

### Base Configuration – Dual-Canister (2 DPUs + 2 RAID GPUs)

The standard configuration utilizes a dual-canister (Node-1 and Node-2) design within a single 2U chassis. This setup is engineered to ensure there is no single point of failure while maintain simplicity.

- **Redundant Controllers:** Each node (canister) is equipped with its own NVIDIA BlueField-3 DPU and SupremeRAID™ HE GPU.
- **Shared Storage Access:** All 26 NVMe SSDs are dual-ported and connected via a Dual-Port PCIe Gen5 Backplane. This allows both nodes to physically access the same pool of drives simultaneously.
- **Failover Logic:** The system uses Pacemaker and Graid SPDK Target software. If Node-1 encounters a fault, the storage services and RAID volumes are automatically migrated to Node-2.

## Scaling to Advanced Dual-Canister (4 DPUs + 4 RAID GPUs)

For maximum-performance environments, each canister can be scaled to accommodate two DPUs and two RAID GPUs, yielding a total of four DPUs and four RAID GPUs across the dual-canister architecture.

- **Linear Performance Scaling:**

Expanding to the advanced configuration, each canister doubles its DPUs and GPUs, delivering near-linear gains in throughput and IOPS.

- **Scale-Out Throughput:**

As the dual-canister system expands, throughput and IOPS scale near-linearly. In the advanced dual-canister setup (4 DPUs + 4 RAID GPUs), read throughput reaches **85.5 GiB/s** with **8.87 million IOPS**, effectively doubling the performance of the base configuration.

- **Fabric Redundancy:**

In the advanced dual-canister environment, NVMe-oF Multi-pathing allows client servers to maintain a consistent data path across the entire DPU fabric. Even if a path or a canister goes offline, the multi-path initiator reroutes traffic to the active DPUs without interrupting the application.

## Performance Data: Base vs Advanced

Performance was tested from an SPDK Initiator using Kioxia CM7 NVMe SSDs over an NDR 200G network. The results demonstrate near-linear scalability, proving that throughput and IOPS grow directly with the addition of BlueField-3 DPUs.

### Throughput Performance (1M Random)

Metric	Base (2 DPUs + 2 RAID GPUs)	Advanced (4 DPUs + 4 RAID GPUs)
1M Random Read	42.5 GiB/s	85.5 GiB/s
1M Random Write	16.1 GiB/s	32.9 GiB/s

Analysis: Moving from the **Base Configuration** to the **Advanced Dual-Canister Configuration** precisely doubles throughput, confirming that the architecture scales seamlessly with added hardware to saturate theoretical maximum speeds over InfiniBand.

### IOPS Performance (4k Random)

Metric	Base (2 DPUs + 2 RAID GPUs)	Advanced (4 DPUs + 4 RAID GPUs)
4k Random Read	4.56 Million IOPS	8.87 Million IOPS
4k Random Write	1.09 Million IOPS	2.17 Million IOPS

Analysis: Scaled to 4 DPUs with 4 RAID GPUs rivals massive rack-scale storage arrays while maintaining a compact 2U footprint. Even with RAID 5 parity overhead, the system sustains over 2 million write IOPS due to SupremeRAID™ GPU offloading.

## Use Cases: Built for the Modern Datacenter

- **AI & Machine Learning (GPUDirect Storage)**

The platform is GPUDirect Storage (GDS) Ready. The Graid SPDK Target with NVMe over RDMA enables direct GPU access to RAID volumes, bypassing the host CPU to deliver ultra-low latency for AI/ML workloads.

- **Enterprise-Grade High Availability**

The system provides robust HA capability to ensure there is no single point of failure. SupremeRAID™ array migration, NVMe-oF multi-pathing, and Pacemaker failover ensure seamless failover and constant data availability even during controller maintenance or hardware failure.

- **Scalable Cloud-Native Storage**

Supporting up to 1,023 volumes per array, the platform provides the massive scalability required for modern multi-tenant environments, Software Defined Storage (SDS), and Kubernetes.

## Advantage: Why Choose This Solution?

The AIC F2026 AI Data Platform provides several industry-leading benefits:

- **Extreme Performance:** Unlocks the full potential of Gen5 NVMe, delivering wire-speed performance (up to 85 GB/s) and millions of IOPS.
- **High Availability:** Enterprise-ready design with redundant controllers and automated failover ensures no single point of failure.
- **Simple & Flexible Deployment:** Architecture can be easily scaled and configured (2-node vs 4-node) based on specific performance and protection level requirements.
- **Resource Efficiency:** Full offload to DPU and GPU ensures host CPU cycles are preserved for primary AI and application workloads.
- **Advanced Data Integrity:** Combines GPU-accelerated RAID 5/6 with advanced journaling to protect data without the typical performance penalties of software RAID

## Conclusion

The AIC F2026, empowered by NVIDIA BlueField-3 and SupremeRAID™ HE, is more than a JBoF—it is a comprehensive AI-Ready Enterprise Data Platform. It provides enterprise-grade high availability while delivering the raw performance required to power the next generation of AI innovation.

## More Information

Please refer to AIC website: [www.aicpc.com](http://www.aicpc.com)

Please refer to Graid Technology website: <https://graidtech.com/>