Graid Technology Inc.

SupremeRAID™ AE (AI Edition)

Achieve Uninterrupted AI Performance —Prevent GPU Bottlenecks from Drive Failures

Learn more at www.graidtech.com/product/supremeraid-ae/

SupremeRAID™ AE (AI Edition) is designed to empower enterprises running GPU servers and AI workloads with unmatched data management capabilities. Built to meet the demands of modern AI applications, SupremeRAID™ AE eliminates data transfer bottlenecks and optimizes GPU performance, enabling faster AI model training, inference, and streamlined data workflows.

Built for the Demands of Modern AI

- **Intelligent Data Offload Engine:** By offloading data tasks to optimize GPU utilization, SupremeRAID™ AE enhances system efficiency, allowing for faster, more effective processing of complex AI workloads.
- **Flexible GPU Deployment:** SupremeRAID™ AE provides the flexibility to begin with shared GPUs and easily transition to dedicated configurations as workloads expand. This adaptability ensures cost-effective scalability and prepares enterprises for future growth.
- **Enterprise-Grade Data Protection:** Designed to maintain uninterrupted access to critical datasets, SupremeRAID™ AE prevents workflow disruptions and safeguards data integrity during intensive AI training and inference processes.
- **GPUDirect Storage (GDS) Support:** SupremeRAID™ AE enables direct data transfers from NVMe drives to GPU memory, bypassing host memory to eliminate bottlenecks. This seamless integration accelerates workflows with near-zero latency, ensuring faster model training and inference for AI applications.
- **NVMe-oF Support:** With the ability to scale storage across multiple hosts, SupremeRAID™ AE supports the massive datasets required for modern AI models while preserving performance, even as data volumes grow.
- **Clustering AI Storage Compatibility:** SupremeRAID™ AE seamlessly integrates with BeeGFS, Lustre, and Ceph distributed file systems, simplifying large-scale dataset management and reducing storage costs without the need for data migration.

SupremeRAID™ AE Redefine the Boundaries of AI Performance

Accelerate Results

GPUDirect Storage (GDS) support ensures faster data throughput and near-zero latency, accelerating model training and inference for AI workloads.

Maximize GPU Performance

The Intelligent Data Offload Engine optimizes workloads, enabling full utilization of GPU resources for peak performance.

Protect Critical Data

Our resilient enterprise RAID protection safeguards your datasets, ensuring uninterrupted access during intensive AI processes.

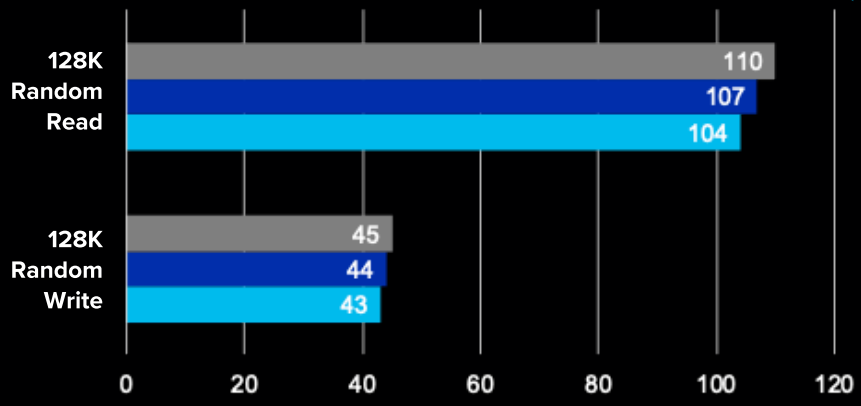
Scale Seamlessly

Flexible deployment options and GDS support allow for smooth scalability to meet your current and future performance demands.

SupremeRAID™ AE (AI Edition)

Benchmark Performance Numbers

- **Prevents GPU idling** caused by drive failures, ensuring uninterrupted workloads.
- **Achieves over 95% of raw NVMe performance** from GPU with RAID protection.



	128K Random Write	128K Random Read
RAID5 Theoretical (9 NVMe)	45 GiB/s	110 GiB/s
fio	44 GiB/s	107 GiB/s
gdsio	43 GiB/s	104 GiB/s

Hardware
Server Model: Supermicro SYS-821GE-TNHR, CPU: Intel(R) Xeon(R) Platinum 8462Y+ *2, Memory: M321R8GA0BB0-CQKZJ 64GB DDR5 4400 MT/s *32, GPU: NVIDIA H100 *8, NVMe Drive: SAMSUNG PM1743 3.84T MZWLO3T8HCLS-00A07 *9

Software
OS: Ubuntu 22.04.5 LTS, Kernel: 5.15.0-131-generic, SupremeRAID™ AE Driver: 1.6.1-38, Benchmark Tool: fio-3.30, gdsio-1.11 on H100*8 in GDS compatible mode

RAID Configuration
RAID5 with SAMSUNG PM1743 3.84T *9

SupremeRAID™ AE

Software & Hardware Specs

Supported RAID Levels:
RAID 0, 1, 5, 6, 10

Max Physical Drives:
32

Max Drive Groups:
Linux: 8

Max Virtual Drives per Drive Group:
Linux: 1023

Max Drive Group Size:
Defined by physical drive size

Supported Enterprise GPUs:
NVIDIA Ampere, Ada, Lovelace, Hopper, and Blackwell series

Supported NVMe SSDs:
Dapustor, FADU, Hagiwara, Kingston Technologies, KIOXIA, Memblaze, Micron, Petaio, Phison, Samsung, Scaleflux, Seagate, Solidigm, Western Digital

OS Support:
RHEL 9
Ubuntu 20.04 / 22.04 / 24.04

For questions about product specs, email us at info@graidtech.com

SupremeRAID™ AE

Achieve Uninterrupted AI Performance —Prevent GPU Bottlenecks from Drive Failures

Built for the Demands of Modern AI.
SupremeRAID™ AE (AI Edition) revolutionizes GPU-based RAID with GPUDirect Storage support and an Intelligent Data Offload Engine, delivering peak NVMe performance and efficiency for AI and GPU-driven workloads. Contact us today to unlock the full potential of your AI infrastructure! Email **info@graidtech.com** to get started today.

SupremeRAID™ AE

Powering AI with Unmatched Performance & Resiliency

Graid Technology Inc.

Contact Us: info@graidtech.com

Learn more at www.graidtech.com/product/supremeraid-ae/

05132025