

SOLUTION BRIEF



SupremeRAID™: The Optimal RAID Solution for AI & Machine Learning on Arm



ARM IP

- + Physical IP
- + Security IP
- + Subsystems
- + System IP

OVERVIEW & GOAL

SupremeRAID™ is the optimal RAID solution for AI and machine learning on Arm CPUs, utilizing GPU-based technology for accelerated I/O, enhanced read/write performance, over 50% increased storage capacity, and full CPU access. Our mission is to provide cutting-edge IT storage infrastructure for NVMe and NVMeoF SSDs, ensuring top performance, speed, ease of use, flexibility, and total cost of ownership (TCO) for future high-performance workloads.

Arm-based servers use processors built on Arm architecture, which are known for their energy efficiency and low power consumption. Arm servers are an excellent choice for energy efficiency, cost-effectiveness, and scalability scenarios. They are well-suited for cloud computing, edge computing, IoT devices, and environments with fluctuating or growing workloads.



APPLICATION AREA

- + Aerospace and Defence
- + Artificial Intelligence
- + Server and Infrastructure

By offloading RAID Infrastructure operations to a GPU, SupremeRAID™ enables the Arm-based server to excel at its primary task and not be burdened by secondary tasks.

In this brief, we compare the RAID5 performance of SupremeRAID™ SR-1000 with Linux MD RAID on an Arm CPU based server platform, focusing on 4K random read/write and 1M sequential read/write.

Hardware Specifications

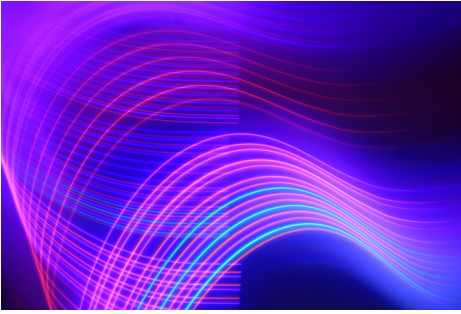
- Server: Giga Computing R182-P91
- CPU: Ampere(R) Altra(R) Max Processor Neoverse-N1 128 Cores *2
- Memory: Micron Technology 18ASF2G72PDZ-3G2R1 DDR4 3200MT/s 16GB * 16
- NVMe: Solidigm P5620 SSDPF2KE032T1 3.2TB * 12
- RAID: SR-BUN-1000-04-FD32 * 1

Hardware Topology

- Solidigm P5620 SSDPF2KE032T1 3.2TB * 6 on CPU1
- Solidigm P5620 SSDPF2KE032T1 3.2TB * 6 on CPU2
- SupremeRAID™ SR-1000 on CPU2

Software

- OS: Ubuntu 20.04.5 LTS



CHALLENGE

RAID is a necessity, but traditional RAID is an I/O bottleneck, leaving little to no ROI on investment in AI/ML infrastructure.

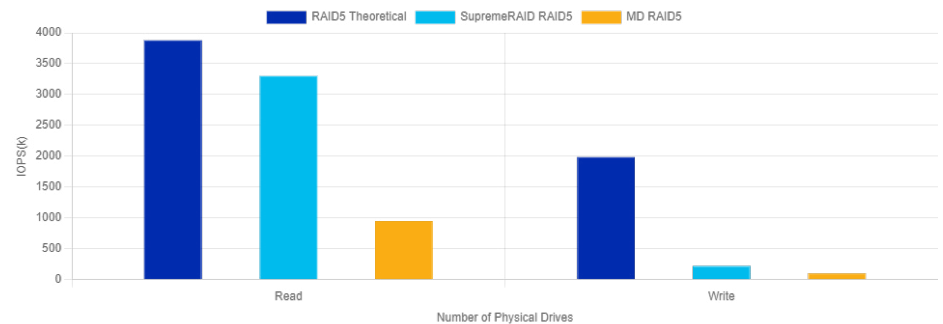
- NVMe storage is used everywhere, with capacities and performance constantly increasing.
- NVMe storage performance is evolving too fast to be fully utilized by existing storage architectures.
- Software-based RAID competes with other applications for CPU power, slowing the whole system down.
- Advances in NVMe technology are moving so fast that ASIC-based RAID can't deliver the full performance capability of the SSDs – bottlenecking AI/ML deployment.

SOLUTION & BENEFITS

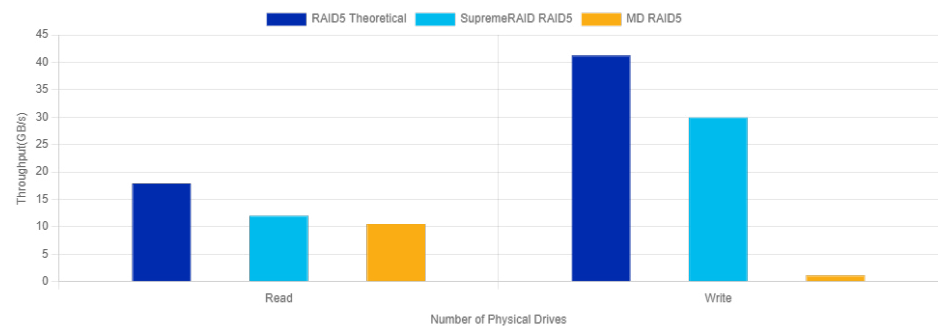
GPU-based SupremeRAID™ eliminates the RAID bottleneck for faster I/O and improved cycle and memory management on Arm servers.

- SupremeRAID™ removes performance bottlenecks by working outside the data path, allowing devices to achieve their maximum throughput.
- By redirecting RAID functions and parity calculations to the GPU card, Arm CPU cycles and memory can be redirected to running more applications and tasks.
- SupremeRAID™ delivers accelerated I/O by utilizing patented out-of-path RAID technology that redirects parity and RAID calculations to the GPU while allowing user data to pass directly from the PCIe bus to the NVMe devices.

Results:



SupremeRAID and MD RAID5 Random IOPS Comparison



SupremeRAID and MD RAID5 Sequential Write Comparison

SUPPORTING INFORMATION

Watch the video [10 Ways AI and Machine Learning Benefit from Faster IO with SupremeRAID™](#)

How does AI & Machine Learning Benefit from Faster I/O with SupremeRAID™?

- Faster data loading for model training
- Reduced model training times
- Improved model iteration
- Enhanced data preprocessing & feature engineering
- Optimized data pipelines
- Improved real-time inference

-
- Improved scalability without bottlenecks
 - Optimized hardware accelerator usage
 - Enhanced user experience
 - More efficient resource utilization

FURTHER DETAILS

- [SupremeRAID™ for AI Machine Learning](#)
- [SupremeRAID™ for NVMe RAID](#)
- [SupremeRAID™ How it Works](#)
- [Graid Technology Whitepapers](#)
- [Drivers & Documentation](#)
- [Graid Technology LinkedIn](#)
- [Graid Technology Youtube](#)
- [Graid Technology X](#)
- [Graid Technology Facebook](#)