TECH BRIEF

# Western Digital Ultrastar® DC SN650 and SupremeRAID™ SR-1010 Set the New Standard in NVMe™ RAID

## Existing Technologies & Challenges

In performing RAID computations, the user has historically had the following two options:

- Software RAID (e.g., mdadm on Linux®)
- Hardware RAID (e.g., a RAID Controller Card)

**Software RAID:** Software RAID provides an independent solution that can work with multiple media types (HDD or SSD) and protocols (SATA, SAS, NVMe). The challenge with Software RAID is generally poor performance with a high cost for CPU resources.

**Hardware RAID:** The challenge with these RAID adapters is that they can only be used with their native physical protocols. They cannot be used with network-attached devices and don't scale performance fully or efficiently.

In this paper, we discuss and benchmark a third option: **Hardware-Accelerated Software RAID**. This option provides protocol independence and the high performance needed for network-attached Flash storage.

## Summary

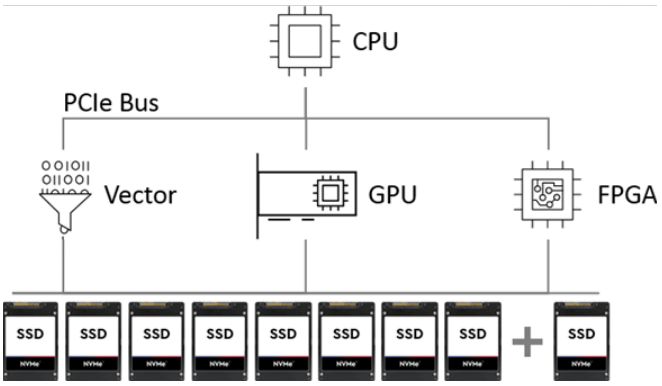
With the rapid adoption of NVMe SSD as a prevailing standard in storage infrastructure, data centers face a crucial challenge: the need for a forward-looking solution that can harness the performance of NVMe SSD while safeguarding data security and ensuring uninterrupted business operations. Simply put, existing storage architectures are struggling to keep up with the swiftly evolving landscape of flash storage performance.

In Software-Composable Infrastructure (SCI), compute, storage, and networking resources are abstracted from their physical locations and are usually managed with software via a web-based interface. With the emergence of NVMe SSD and NVMe-oF™ technologies, SCI can disaggregate storage resources without sacrificing performance and latency. As NVMe SSD technology rapidly evolves, a significant performance bottleneck is introduced — RAID data protection.

## GPU-based Hardware Accelerated Software RAID Solution

Although implementing complex RAID levels like 5 and 6 while maintaining high performance on NVMe drives has posed a significant challenge, primarily due to traditional software parity calculations, the SupremeRAID solution addresses this issue by offloading and parallelizing CPU-intensive parity calculations onto a hardware accelerator.

There are multiple options for executing these calculations using hardware engines. The first option involves utilizing CPU extensions such as Vector/SIMD to offload and parallelize parity calculations, thus enhancing RAID performance. Alternatively, dedicated hardware accelerators like GPUs (DPUs) or FPGAs can also be employed to offload and parallelize these calculations. Graid Technology Inc. presents the GPU-based RAID solution tested in this project, known as SupremeRAID SR-1010, and provides a detailed block diagram showcasing its implementation.



## Benchmarking Infrastructure

- Eight Ultrastar SN650 NVMe drives added into Supermicro® server
- Preconditioning of drives completed
- Power state of all the drives remain at 18W
- mdadm comparison with RAW drive performance, along with SupremeRAID SR-1010 card in RAID 5 (7+1)

## Benchmarking Methodology

- Flexible IO (FIO) is the workload generator
- Fundamentally, the process has two phases – the sequential process using 128K-4M blocks (to measure bandwidth) and the random process using 4K blocks (to measure IOPS). We ran multiple instances with different workloads of the tests and averaged the results.

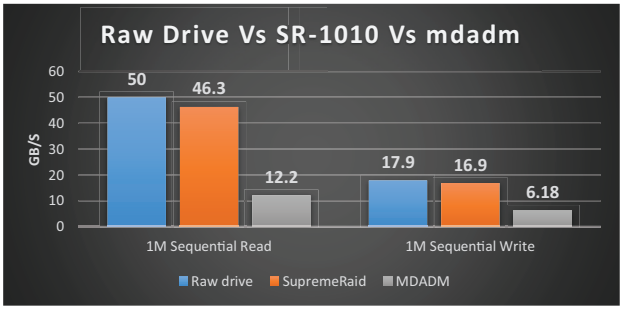
## Results

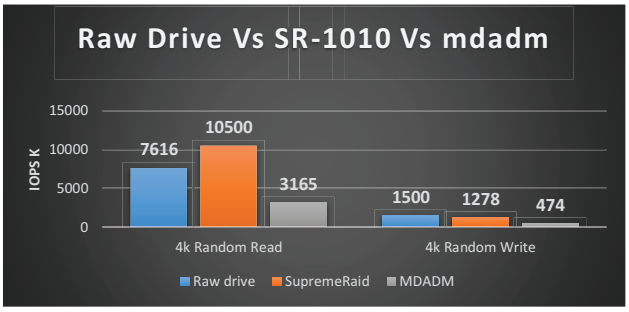
1. Raw Drive Vs SupremeRAID Vs mdadm performance
2. RAID 5 (7+1) operation Sequential/Random read & write performance
3. Performance with one drive failure & Performance while rebuilding
4. CPU utilization in RAID 5 operation

**Raw Drive vs SupremeRAID vs mdadm performance**

mdadm is one of the most widely used Linux utility tools across the IT industry to manage and monitor software RAID devices. It is used in modern Linux distributions to replace older software RAID utilities, such as raidtools or raidtools2. But, for the latest generations of NVMe SSDs, mdadm appears to lack previous performance and rebuild efficiency seen previously. When tested against the SupremeRAID SR-1010, SupremeRAID achieved both superior performance and rebuild speeds. Testing suggests that software RAID (e.g., Linux mdadm/mdraid) may be too slow in newer workload combinations.

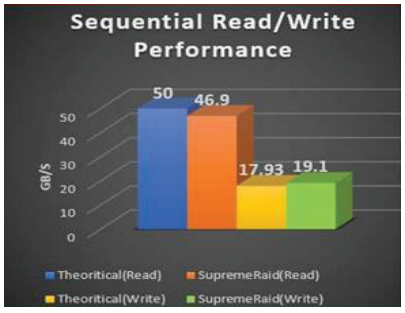
Seen below, results show that during sequential workloads SupremeRAID is closer to raw drive performance, achieving 94% of the raw drive performance, compared to mdadm achieving only ~20%. When it comes to random workloads, SupremeRAID achieves 100% of raw drive performance, with mdadm achieving ~45%.

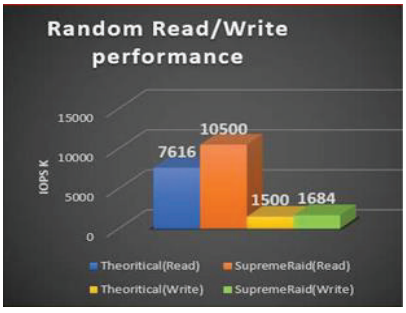



**RAID 5 (7+1) operation Sequential/Random read & write performance**

Results also show that SR-1010 RAID 5 provides almost twice the performance of open-source software RAID such as mdadm RAID 5. SR-1010 RAID 5 performance is competitive with Linux MD RAID 10 while providing more usable capacity and better data security. Similarly, results are touching the theoretical value even though it is present in RAID group which is providing the data security.

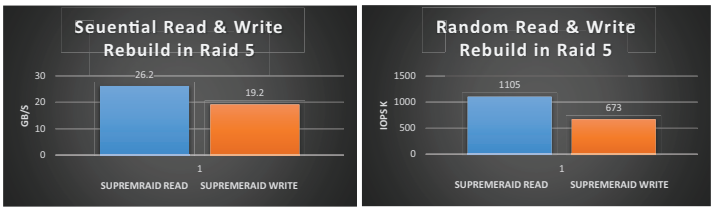



**Performance with one drive failure & performance while rebuilding**

During the rebuilding process, the performance of sequential read and random read experiences a reduction of approximately 50%, which can be attributed to the increased traffic caused by the rebuilding activity. However, with the advantage of leveraging the computational power of GPUs, SupremeRAID exhibits a smaller drop in performance compared to other solutions. SupremeRAID RAID 5 efficiently maintains the rebuild throughput while simultaneously achieving equal or even higher levels of sequential write and random write performance.

It is worth noting that open-source software-based RAID 5 configurations exhibit significantly lower performance due to the degraded reading and resource-intensive parity calculation tasks, which consume a substantial amount of CPU resources.




**Note:**

- Sequential access performance is compared based on the performance of 7 drives, considering the presence of parity data.
- Random access performance is compared based on the performance of 8 drives, as random operations are evenly distributed across the entire drive and are not affected by parity data.

**CPU utilization in RAID 5 operation**

The below results show the CPU utilization under RAID 5.

| Workload | SupremeRAID SR-1010 (RAID 5) | Avg CPU Utilization |
|---|---|---|
| 4k Random Read | 10.5 M IOPS | 15.83% |
| 4k Random Write | 1.7 M IOPS | 7.77% |
| 1k Sequential Read | 46.9 GB/s | 1.67% |
| 1k Sequential Write | 19.1 GB/s | 1.97% |

## Conclusion

To ensure optimal performance for databases like MySQL™ and other applications, leveraging the fastest and most reliable storage is essential. It has become standard practice to utilize multiple NVMe SSDs with RAID configurations to enhance data protection. By selecting SupremeRAID and Western Digital SN650 NVMe drives to unlock the full potential of your storage infrastructure, you gain access to a highly effective and efficient RAID 5 solution which consistently delivers superior performance, data protection, and scalability for your critical applications.

**Additional Benefits:**

- **Data Loss Prevention:** SupremeRAID safeguards against data loss even in the event of the simultaneous failure of two SSDs, ensuring data integrity and business continuity.
- **Faster Rebuild Rates:** Enjoy up to 66% faster rebuild rates while maintaining minimal impact on overall system performance. This allows for quicker recovery and reduces potential downtime.
- **Maximized Performance:** SupremeRAID optimizes the performance of all drives within the RAID, including the parity drive. This is particularly advantageous for sequential workloads, where it ensures high throughput and efficiency.
- **GPU-based Solution:** By utilizing a GPU-based RAID solution like SupremeRAID, you can potentially mitigate or eliminate performance-related issues. Moreover, offloading RAID calculations to the GPU frees up valuable CPU cycles, enabling their allocation to other tasks. This also opens the possibility of considering lower-cost CPUs without compromising system performance.
- **PCIe® Generation 4.0:** The Ultrastar DC SN650, built on the latest PCIe generation 4.0, offers exceptional scalability and performance. This empowers platforms to handle increasingly larger application workloads without compromising efficiency or speed.

**About the SupremeRAID SR-1010**

The SupremeRAID SR-1010 is a groundbreaking solution designed to revolutionize SSD performance in the enterprise realm. By prioritizing scalability, flexibility, and cost reduction, this state-of-the-art technology takes data center performance to new heights.

In today's data center world, speed and throughput are everything. Graid Technology, Inc. recognized the limitations and bottlenecks caused by traditional RAID and developed a GPU-based storage solution to launch RAID technology into the future. With proven performance tests and partnerships with global industry leaders, SupremeRAID delivers maximum SSD performance, comprehensive enterprise data protection, unmatched flexibility, and unbeatable ROI.

**About the Ultrastar DC SN650**

Specifically for cloud and scale-out workloads, the Ultrastar DC SN650 NVMe SSD offers optimized performance, enhanced Quality of Service (QoS) consistency, and improved storage utilization for object, and file storage. As the size and complexity of emerging workloads, such as big data, artificial intelligence, and machine learning, continue to grow, they often require distributed, tiered, or disaggregated architectures for storage. The SN650 excels at efficiently transferring these massive datasets and delivering consistent performance across multiple hosts. This makes it an ideal solution for scaling capacity and maximizing GB/watt efficiency in the face of these increasingly demanding workloads.

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
**US (Toll-Free):** 800.801.4618
**International:** 408.717.6000

www.westerndigital.com