



# GRAID SupremeRAID™

## A Data Protection Solution for NVMe SSDs

NOVEMBER 2020



# Table of Contents

---

Preface.....	3
The Rise of NVMe.....	3
Existing NVMe Data Protection Solutions .....	4
Software RAID .....	4
Hardware RAID.....	5
GRAID SupremeRAID™ .....	6
Solution Advantages.....	7
Test Case & Specifications .....	8
Test Result .....	10
GIGABYTE R282-Z92 All-Flash Server.....	12
Conclusion .....	13

## Preface

---

Due to its popularity within the IT industry, RAID (Redundant Array of Independent Disks) technology is now widely used in various computing / storage systems that employ a large number of hard disks.

During the past few decades however, the development of RAID technology has only focused on mechanical hard disks as the storage medium, and the basic characteristics of these disks such as hardware interface and read / write performance has not changed very much at all. This all changed with the introduction of SSDs (Solid-State Drives).

## The Rise of NVMe

---

Early on, most SSDs used traditional interfaces such as SATA or SAS to connect with a computer's data bus, but due to the characteristics of NAND flash they quickly hit a performance bottleneck, as SATA or SAS had been designed only with mechanical hard disks in mind. Therefore, from 2009 a working group led by Intel began research into a suitable alternative, which resulted in the development of the NVMe (Non-Volatile Memory Express) interface.

In contrast to where multiple hard disks share a single PCIe controller connected via the SATA (Serial) based AHCI (Advanced Host Controller Interface), NVMe drives directly connect to the host system via the high-speed PCIe (Peripheral Component Interconnect Express) interface. In addition, the number and depth of NVMe queues was increased greatly, allowing a system to take full advantage of the high concurrency and low latency of flash memory.

---

In contrast to where multiple hard disks share a single PCIe controller connected via the SATA (Serial) based AHCI (Advanced Host Controller Interface), NVMe drives directly connect to the host system via the high-speed PCIe (Peripheral Component Interconnect Express) interface.

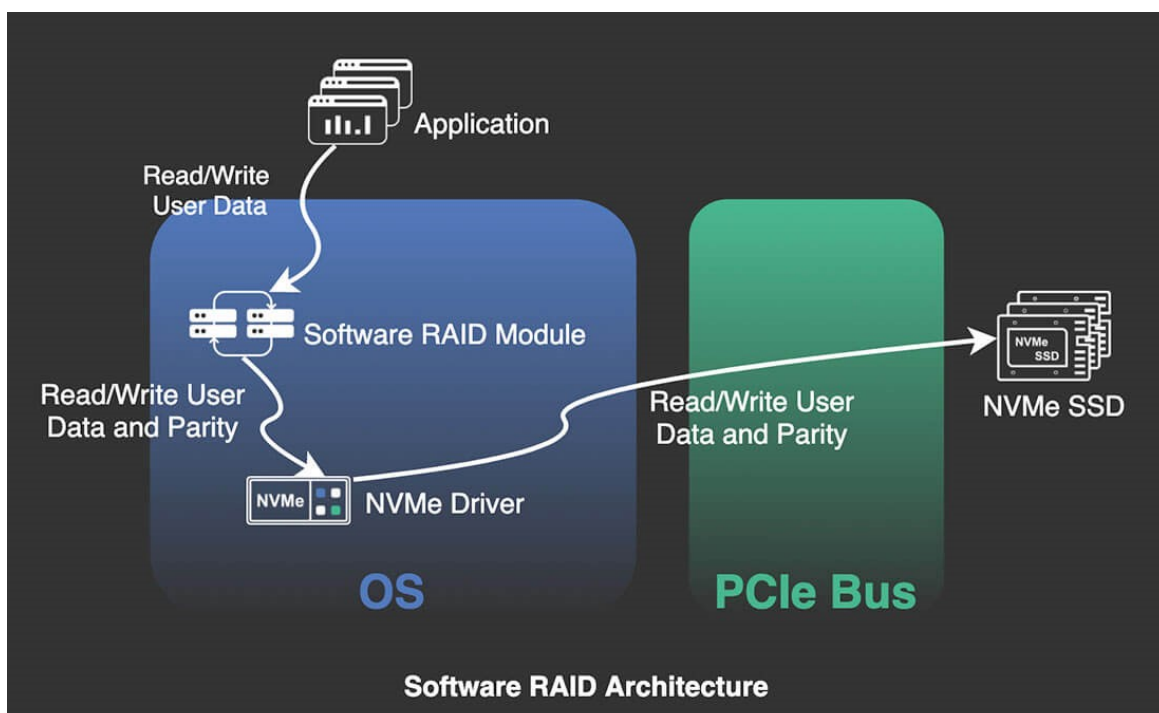
This has prompted more and more computing applications that require high IO performance to adopt NVMe SSDs. However, attempting to maintain such high performance after implementing a RAID data protection system brings new challenges to a technology that was originally designed only for mechanical hard disks. 《Glossary: [What is NVMe?](#)》

## Existing NVMe Data Protection Solutions

Due to the high-performance and low-latency characteristics of NVMe, many companies have begun to adopt NVMe SSDs as the main storage medium in their servers. However, NVMe SSDs connected directly to the host could only be used with traditional RAID technologies for data protection, such as the commonly known Software RAID and Hardware RAID.

### Software RAID

The concept of Software RAID for NVMe is very similar to that already being used for mechanical hard disks, using the CPU of the host system to process NVMe instructions and make checksum calculations. The difference is that since NVMe connects to storage devices via PCIe the bandwidth is higher, the latency is lower and the instruction set design is simpler, so it is highly efficient to directly process RAID via the CPU.



Take a RAID-0 read as an example: when an application reads any 4K block, it will generate an instruction to read NVMe. After receiving this instruction, the software RAID module only needs to interpret and generate new NVMe commands to the specified SSD. The SSD can then send the data directly through DMA to a buffer that can be accessed by the application.

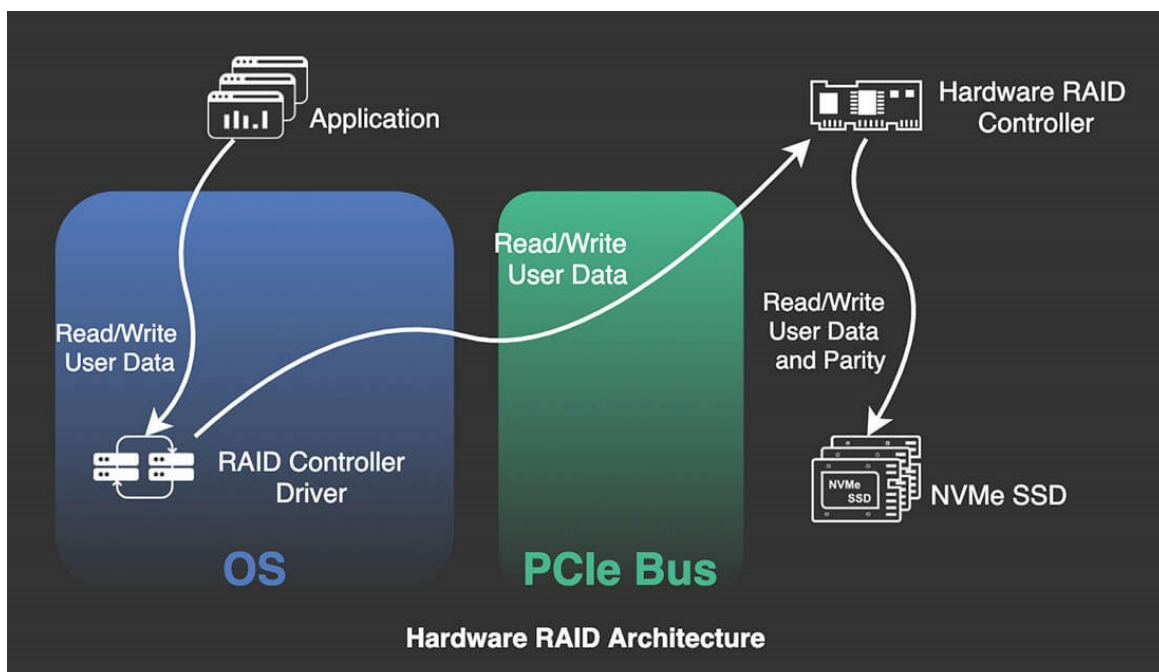
However, a big problem with Software RAID is for RAID modes that require checksum calculations, such as RAID-5 or RAID-6. Take RAID-5 as an example: a 4K random write request will generate two additional read and one additional write commands as well as a checksum calculation.

This process will end up consuming a large portion of the CPU's resources if you wish to fully maximize the performance of all your NVMe SSDs. Therefore, applications that utilize NVMe SSDs as the storage medium will often force users to adopt very high-end CPUs, leading to a substantial increase in the cost of the system.

## Hardware RAID

Hardware RAID is a good solution when employed with traditional hard disks. All RAID logic is completed on a separate hardware controller, which offloads computation from the host CPU. However, it is precisely because of this that all data reads and writes must pass through the RAID controller.

The most common current NVMe SSD transmission interface on the market is PCIe Gen3 x4: if you use a better specification SSD, the RAID controller connected to the host via



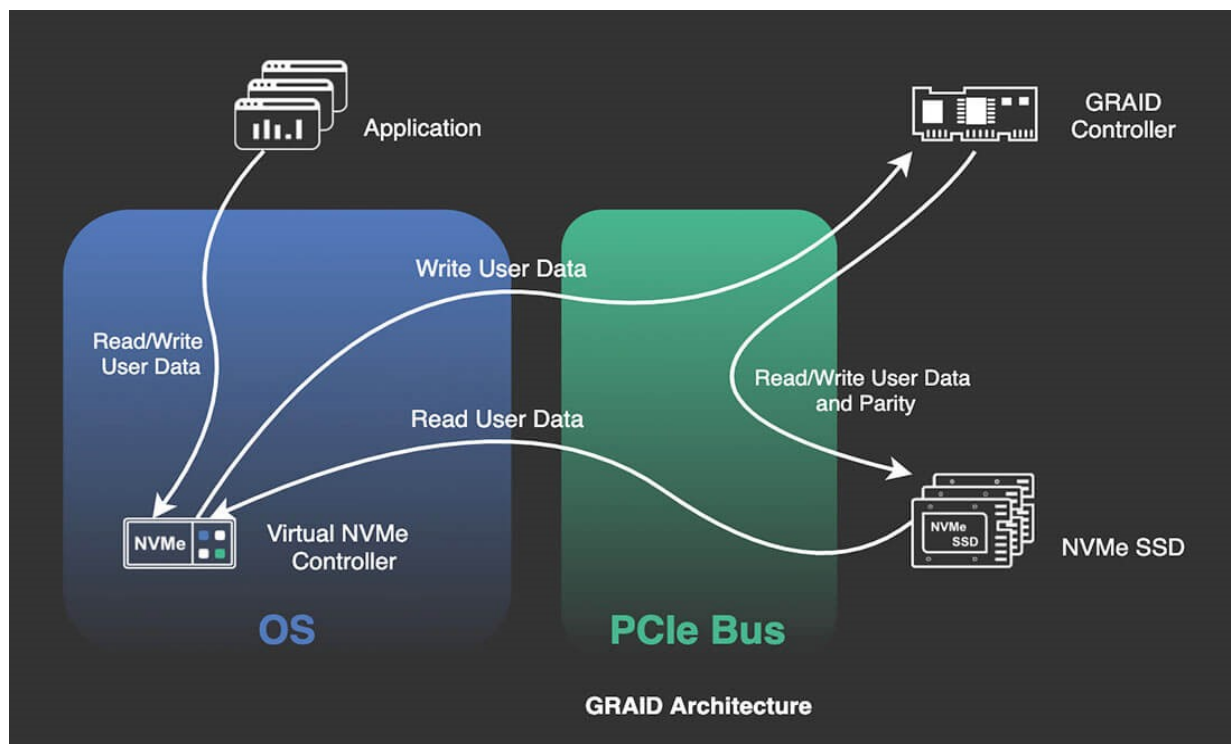


PCIe Gen3 x8 or x16 will easily become a performance bottleneck. In addition, all SSDs must be directly connected to the hardware RAID controller, but since the number of PCIe lanes of the controller itself is very limited, this will directly limit the number of SSDs that a controller can use to set up a RAID unless a PCIe switch is added, which in turn will have a considerable impact on the design and cost of the server system.

Both Software and Hardware RAID have their own advantages and disadvantages. However, since applications that use NVMe SSD storage will not only consume a high amount of CPU resources but also cannot compromise IO performance, the industry urgently needs a new type of RAID technology, especially when we are trending towards a NVMe SSD that can reach up to 1 million IOPS. This new technology should be able to provide RAID level data protection while fully utilizing the performance of NVMe SSDs.

## GRAID SupremeRAID™ – The Next Generation of NVMe RAID Technology

The concept of a hardware-assisted Software RAID solution already exists, which previously used hardware such as a HBA with a RAID BIOS, or a motherboard that is integrated with a RAID BIOS. However, these solutions still depend on the CPU to



process the RAID logic and could not solve the main problems faced by Software RAID in an environment that used NVMe.

Now that a single NVMe SSD will start reaching 1 million IOPS, it will be extremely difficult to design such a high-speed hardware accelerator card that can meet this performance – this development cycle simply cannot keep up with the growth rate of SSD performance. Therefore, Software RAID technology combined with programmable AI-chips – GRAID – has come into being.

GRAID SupremeRAID™ works by installing a virtual NVMe controller onto the operating system and integrating a PCIe device into the system equipped with a high-performance AI processor to handle all RAID operations of the virtual NVMe controller.

### **This setup offers many advantages:**

- Takes full advantage of NVMe performance – 6 million random IOPS which is currently the industry leading performance benchmark
- Unlike Software RAID it does not consume a large amount of CPU resources
- Overcomes many limitations of Hardware RAID cards, such as computing performance, PCIe bandwidth etc.
- Plug and play, and can be used even for systems without PCIe switches that used SSDs directly connected to the CPU via PCIe without needing to change the hardware design
- SCI (Software Composable Infrastructure) compatible and can be used for external SSDs connected via NVMeoF.
- Highly scalable, and new software functions such as compression and encryption can easily be added.

---

GRAID SupremeRAID™ works by installing a virtual NVMe controller onto the operating system and integrating a PCIe device into the system equipped with a high-performance AI processor to handle all RAID operations of the virtual NVMe controller.

## Test Case

---

The following test case of a GRAID system used GIGABYTE's R282-Z92 server with dual AMD EPYC™ 7282 processors and 10 Intel® Optane™ 905P SSDs. Since AMD's EPYC processor platform provides a high number of PCIe lanes, it can be used without a PCIe switch to connect to a large number of NVMe SSDs, and Intel's Optane™ 905P SSDs provides extremely high and stable write performance. This combination delivers an extremely streamlined and effective system. We used fio as our testing tool and tested both RAID-5 and RAID-10, the two most commonly used RAID modes in real life situations.

### Test Case Specifications

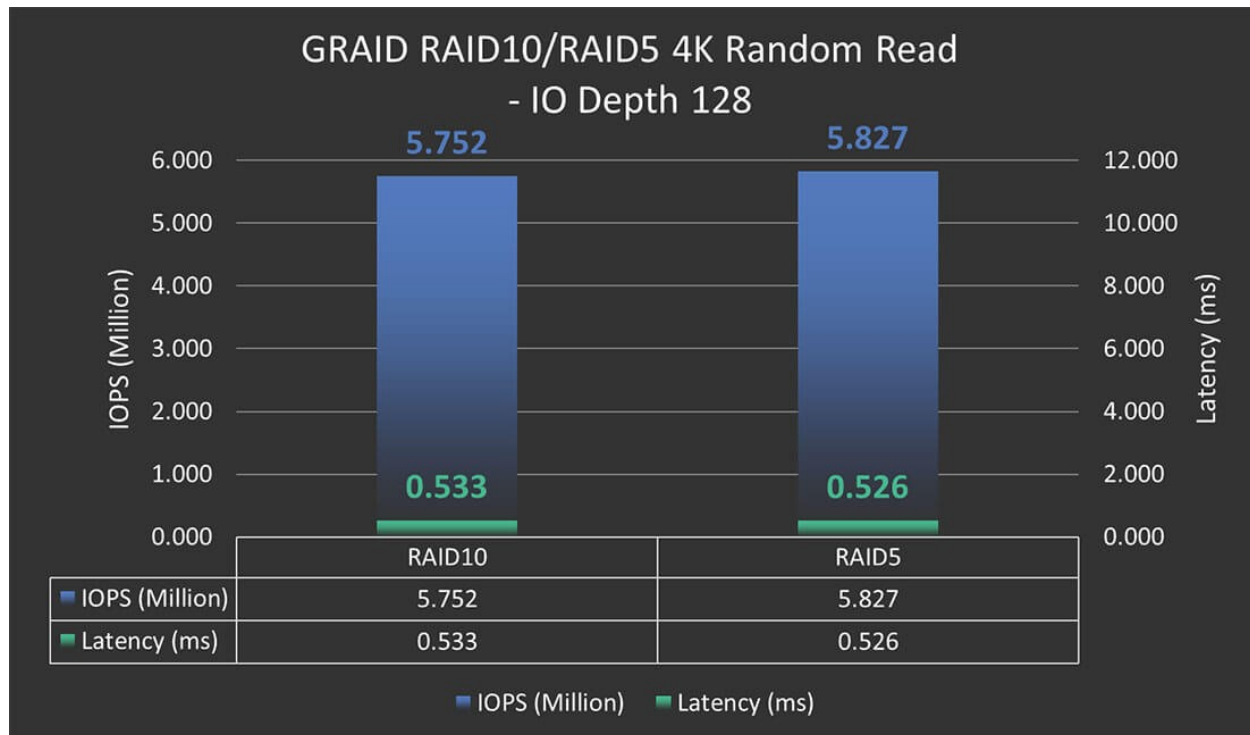
<b>Test Server Specifications</b>	<ul style="list-style-type: none"><li>• GIGABYTE R282-Z92 + 2 x AMD EPYC™ 7282 16 cores processor at 2.8GHz</li><li>• 1 x GRAID NVMe RAID Controller</li><li>• 10 x 480G Intel® Optane™ SSD 905P NVM Express* (NVMe*) drives</li><li>• 1 x NVIDIA Mellanox MCX515A-CCAT ConnectX-5 EN Network Interface Card 100GbE</li><li>• 128 GB RAM</li></ul>
<b>Operating System</b>	Centos 8
<b>Testing Tool</b>	fio-3.7
<b>RAID Modes Tested</b>	RAID-10, RAID-5



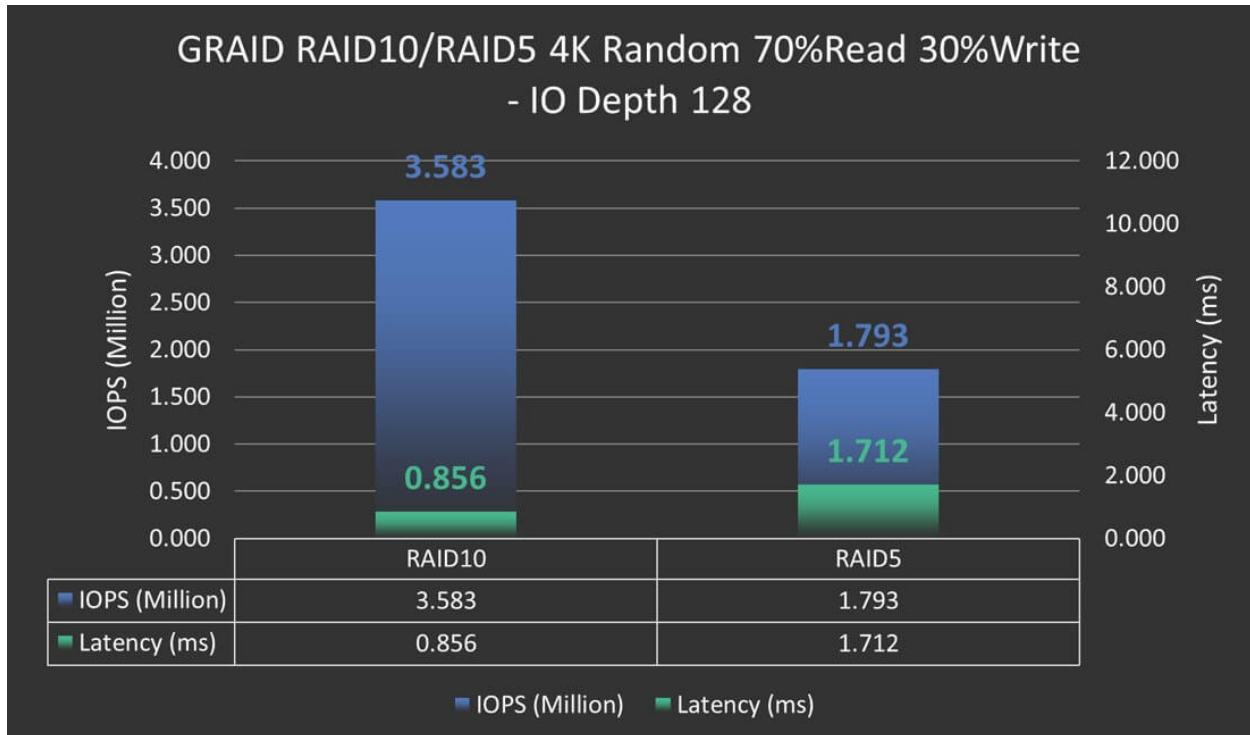
<b>Random Read &amp; Write Test Parameters</b>	[global] ioengine=libaio direct=1 iodepth=128 group_reporting=1 time_based=1 runtime=300 randrepeat=1 bs=4K numjobs=32 cpus_allowed=0-31 cpus_allowed_policy=split rw= [randread, randrw] rwmixread=70
<b>Sequential Read &amp; Write Test Parameters</b>	[global] ioengine=libaio direct=1 iodepth=64 group_reporting=1 time_based=1 runtime=300 randrepeat=1 bs=1M numjobs=7 cpus_allowed=0-6 cpus_allowed_policy=split rw=[read, write] offset_increment=200G size=200G loops=128

## Test Result

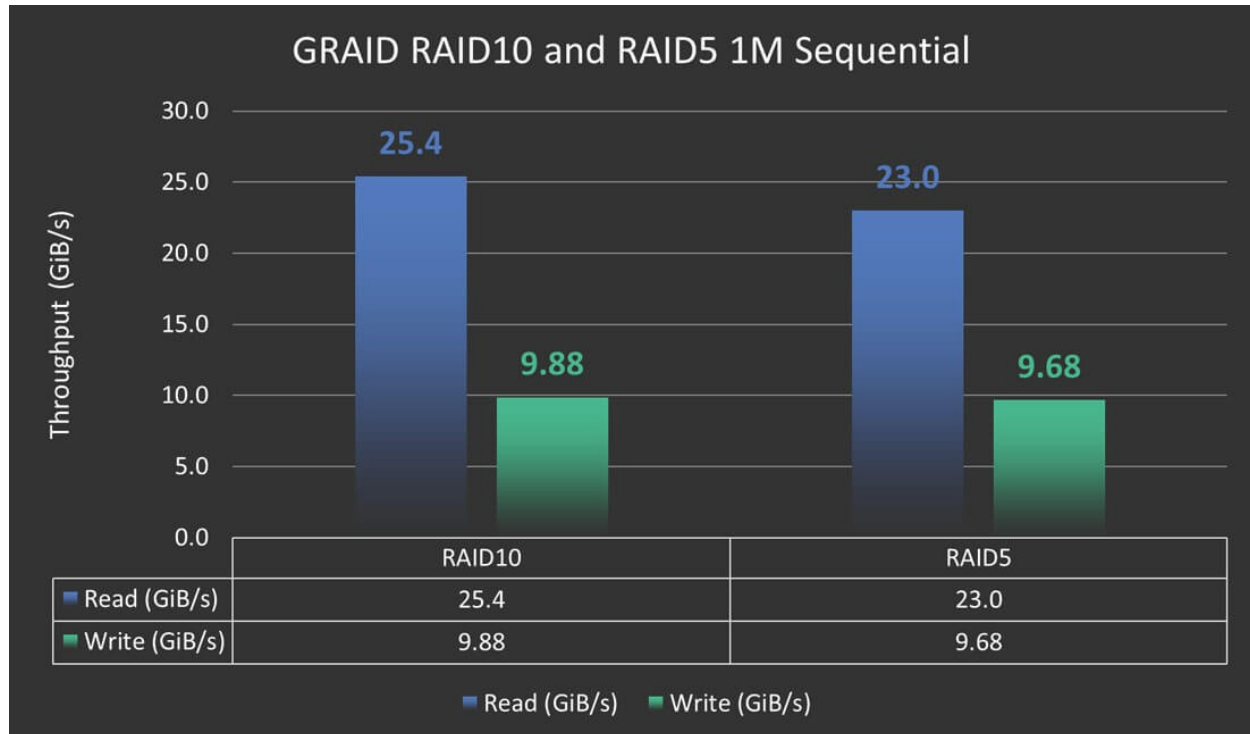
The test result includes IOPS and corresponding latency and throughput.



In the random read test, both RAID-5 and RAID-10 reached the maximum performance limit of 10 NVMe SSDs combined while maintaining very low latency.



In the random read and write test, RAID-10 could still fully utilize the performance of the NVMe SSDs, and RAID-5 could even reach 1.8 million IOPS, the highest result currently achieved in the industry.



Finally, in the sequential read and write test, the throughput of RAID-10 reads and writes can reach 25 GB/s and 10 GB/s respectively, which equals the total throughput of 10 NVMe SSDs, while RAID-5 and RAID-10 read performance is similar.

And even with write penalties and checksum calculations, RAID-5 write throughput can still reach 9.68 GB/s, which is very close to RAID-10

## GIGABYTE All-Flash Server

GIGABYTE's R282-Z92 is an all-flash server built for the 2nd Generation AMD EPYC™ processor. The 2nd Gen. EPYC processor is based on 7nm advanced process technology and features up to 64 cores and 128 PCIe lanes, while also supporting the new PCIe 4.0 high speed transmission interface.

Based on these technical advantages, the R282-Z92 can deliver powerful computing performance to process a large amount of data in real time; in addition, it fully utilizes the abundant number of PCIe lanes available to provide numerous PCIe expansion slots for excellent setup flexibility, as well as support for up to 24 2.5-inch U.2 storage drives at the front of the server chassis to meet the needs of applications using large amounts of real-time read / write data.

GIGABYTE's R282-Z92 is an ideal high-density computing server, with a design optimized for storage and a two-fold increase in I/O performance, that can meet the increasingly demanding workload requirements of software-defined and virtualized infrastructure, big data analytics or all-flash high-performance storage services.

## R282-Z92 Features

- Dual AMD EPYC 7002 Series processors
- Up to 32 x DDR4 memory DIMM slots
- 2 x 1Gb/s Ethernet ports
- 24 x 2.5" NVMe SSD drive bays (front)
- 2 x 2.5" hot-swap SATA/SAS drive bays (rear)
- 1 x PCIe 3.0 M.2 slot
- 2 x PCIe 4.0 expansion slots
- 1600W 80 PLUS Platinum redundant power supply



## Conclusion

This white paper has investigated the impact of NVMe SSDs on traditional RAID technology, and which RAID architecture is more suitable for this storage medium. Through the test results, we can see that GRAID implements data protection while fully utilizing the performance of NVMe SSDs in a highly streamlined and efficient platform. It also frees up the CPU's computing resources so they can be used instead for other applications to meet various workload needs in 5G, IoT and AI computing.

GIGABYTE is planning to launch a GRAID solution soon – for more information, please contact us by email at [server.grp@gigabyte.com](mailto:server.grp@gigabyte.com)

GRAID Technology Inc. is headquartered in Silicon Valley, California with an R&D center in Taipei, Taiwan, and is composed of a dedicated team of experts with decades of experience in the SDS, ASIC and storage industries. Contact us to learn more: call 1 (866) GRAID-10, email [info@graidtech.com](mailto:info@graidtech.com), or visit [graidtech.com](http://graidtech.com) today.