Western Digital.

Solution Brief

SupremeRAID™ with OpenFlex™ Data24

SupremeRAID™

Highlights

- Software RAID
- Hardware RAID
- GPU-based Hardware Accelerated Software RAID

Introduction

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. SCI makes data center resources as readily available as cloud services and is the foundation for private and hybrid cloud solutions. 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.

RAID Computations

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

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

Software RAID

Operating System (OS) Software RAID provides an independent solution that can work with multiple media types (HDD or SSD) and protocols (SATA, SAS, NVMe). The challenge with OS Software RAID is generally poor performance with a high cost for CPU resources. Sequential bandwidth especially Read bandwidth, can achieve high-performance levels, but sequential writes require protection computations. Small block I/O patterns generally have even lower RAID performance levels to render this option generally usable. In summary, this option has the protocol independence needed on network-attached storage devices but lacks the required performance.

Hardware RAID

Hardware RAID was convenient because the SAS adapter card could provide it to the client who was in line with the storage housed in an external enclosure. In the HDD era, a simple ASIC on a RAID card was capable enough to handle all I/O – after all, even with SAS HDD, maximum performance was only around 200 IOPS and 150MB/s of throughput. However, a single NVMe SSD can now deliver around 1 M IOPS and 7Gb/s of throughput.

The hardware RAID Cards were slow to adapt from slower HDDs to higher performing NVMe SSDs. That transition has primarily occurred and can provide higher performance levels when using SSDs. 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 summary, these adapters can potentially have the needed local performance but do not offer protocol independence to work on network-attached devices, severely limiting their usefulness in modern Software-Composable Infrastructures or high-performance applications. These considerations also prevented their testing in these benchmarks.

GPU-based Hardware Accelerated Software RAID

The challenge of implementing complex RAID levels such as 5 and 6 while maintaining high performance on NVMe drives is usually parity calculations. Hardware RAID parity calculations use a hardware engine within the ASIC, while software RAID can only use the CPU's instruction set, whose performance is often limited.

Graid Technology Inc.[1] provides the GPU-based RAID solution tested in this project, the SupremeRAID™ SR-1000.

[1] G-RAID is owned by and is a registered trademark of Western Digital. Graid Technology, Inc. is not affiliated with, related to, or endorsed by Western Digital. Western Digital does not endorse or authorize use of Graid by GRAID Technology, Inc.

SupremeRAID™ with OpenFlex Data24

OpenFlex Data24 NVMe-oF Storage Platform

Western Digital's OpenFlex Data24 NVMe-oF Storage Platform is similar to a 2.5" SAS JBOD Enclosure. It provides 24 slots for NVMe drives and a maximum capacity of 368TB[1] when using Western Digital Ultrastar® DC SN840 15.36 TB devices. Unlike a SAS enclosure, the Data24's dual IO modules use Western Digital RapidFlex™ C1000 NVMe-oF Controllers. These controllers allow full access to all 24 NVMe drives over up to six ports of 100 Gb Ethernet.

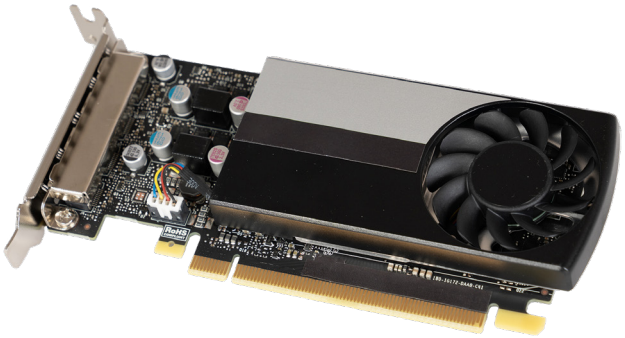
The Data24 is a close replacement for the traditional SAS enclosures. However, the Data24 offers a significant benefit over these enclosures: the ability to integrate directly into Ethernet fabric, allowing for an Any-to-Any mapping of Object Storage Targets to Object Storage Servers.

The OpenFlex Data24 design exposes the full performance of the NVMe SSDs to the network. With 24 Western Digital Ultrastar DC SN840 3.2 TB devices, the enclosure can achieve up to 71 GB/s of bandwidth and over 15 MIOPS at a 4K block size.

SupremeRAID™ SR-1000 NVMe-oF RAID

The SupremeRAID™ SR-1000 (for PCIe 3, 4, and 5 servers) delivers SSD performance in AI-accelerated compute, All Flash Array (AFA), and High Performance Computing (HPC) applications. Designed for both Linux and Windows® operating systems, it supports RAID levels 0/1/10/5/6/JBOD, while the core software license supports up to 32 native NVMe drives.

The SupremeRAID™ SR-1000 enables NVMe/NVMe-oF, SAS, and SATA performance while increasing scalability, improving flexibility, and lowering TCO. This solution eliminates the traditional RAID bottleneck in mass storage to deliver maximum SSD performance for high-intensity workloads.



SupremeRAID™ SR-1000 for PCIe 3, 4, and 5 servers

Conclusion

An NVMe-oF storage enclosure such as the OpenFlex Data24 allows for a broader degree of performance, flexibility, and cost savings not found with traditional hardware or OS-based software RAID.

This GPU architecture outperformed the Advanced Software RAID solution in all areas except large block sequential reads in these tests.

Consider the following:

- SupremeRAID™ SR-1000 adapter is essentially a plug-and-play solution using a commercially available GPU.
- SupremeRAID™ allows competitive pricing as the silicon architecture is not proprietary for this use.
- The ability to separate the data path from the logic path adds value and flexibility.
- A GPU upgrade or a GPU firmware upgrade could provide new features and performance improvements, possibly with low operational impacts
- Traditionally, the data path has presented itself as the bottleneck via an AISC-based RAID controller or CPU computation. Direct IO between the CPU and GPU is efficient and allows the GPU's massive computational capability to manage RAID calculations in the data path.

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
www.westerndigital.com

May 2023 / Revision 01