

High-Performance NVMe Storage Solution for VMware™

XINNOR

This case study details a high-performance storage solution designed and implemented for a media production company for their performance-intensive virtualization environment.

This solution was developed for a highly demanding media production environment, requiring maximum sequential read and write throughput. The system is built on VMware vSphere, where xiRAID Opus virtual machines with a configured NVMe-oF target are used as storage nodes for VMware Data Store.

Challenge

The media production environment requires a unique combination of speed and reliability that traditional SAN or NAS solutions often fail to deliver, especially in virtual environments. Additionally, storage subsystems specialized for VMware are often too expensive and complex and may not provide the required level of performance.

The main challenges addressed were:

1. High-performance requirements. The need to provide maximum sequential read and write speeds for seamless editing and transcoding of large video files (4K/8K).
2. Simple VMware infrastructure integration. Providing high-speed storage capabilities for virtual machines running across VMware hosts via a simple and easily supported NVMe-oF target.

About xiRAID Opus, the NVMe composer with high-speed data protection

xiRAID Opus is a Linux user-space software solution that unifies local and network-attached NVMe drives into a high-performance, energy-efficient storage platform. It maximizes speed and reliability for demanding applications while minimizing hardware overhead, reducing power costs, and simplifying infrastructure management.

xiRAID Opus extends data protection functionality with:

- Integrated network storage for seamless scaling
- Native NVMe-oF initiator and target support across drives, volumes, and RAID arrays
- Built-in vhost virtualization
- End-to-end QoS controls for predictable performance in shared environments

At its foundation, xiRAID Opus employs a Linux user-space datapath engine that bypasses the kernel I/O stack using polling mode to reduce latency and eliminate OS dependencies. This not only improves performance but also ensures frictionless Linux distribution updates: no kernel tuning required or compatibility issues expected.

Environment

ESXi Hypervisor Host Server: ASRock Rack SIENAD8UD3 motherboard with AMD Epyc 8124p CPU

RAM: 320GB DDR5 ECC Registered RAM

NVMe Drives: 4x KIOXIA 15.36TB CM7-R NVMe PCIe Gen5 U.3

Network Adapter: 100Gb Intel E810 NIC

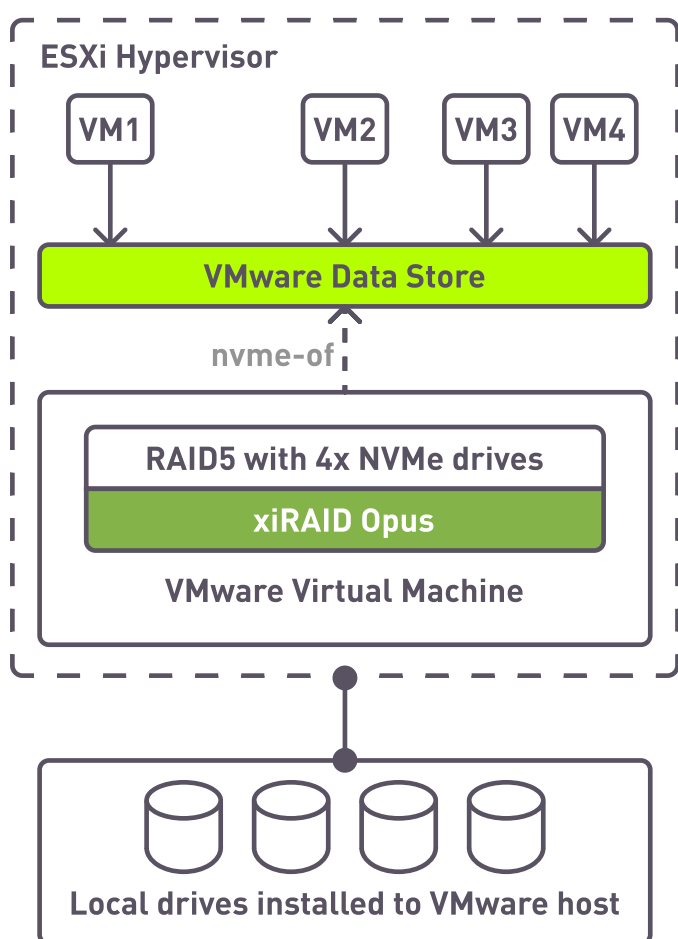
VMware virtual machine with xiRAID Opus:

- **CPU:** 4 cores
- **RAM:** 16GB
- **OS:** Ubuntu 24.04
- **RAID:** xiRAID Opus (RAID 5) across 4x KIOXIA CM7-R drives

Solution Architecture

To achieve these goals, the Xinnor xiRAID Opus solution was selected as the native NVMe-oF target with RDMA support and ultrafast software RAID. Operating in userspace with minimal resource consumption, this solution made it possible to fully leverage the performance potential of KIOXIA CM7-R drives, limited only by network bandwidth. It delivered high performance with ultra-low latency while maintaining flexibility and keeping resource usage to a minimum.

Key architectural details of the solution based on xiRAID Opus can be found in the diagram below:



1. The physical NVMe drives, located in one of the ESXi hypervisor hosts, are directly passed through (Passthrough) to a Linux virtual machine (Ubuntu 24) as PCIe devices.
2. The xiRAID Opus software RAID engine is installed on the virtual machine with NVMe drives. This engine aggregates the physical NVMe drives into a high-speed RAID5 array.
3. The created RAID array is used to form the NVMe-oF subsystem. This allows the block storage to be exported over the network, using the NVMe over RDMA protocol (RoCE) on the 100Gb Intel E810 NIC.
4. The resulting block device, exported via NVMe-oF, is connected to a VMware datastore and formatted as VMFS 6.

- The datastore with an NVMe-connected xiRAID Opus block device is used as storage for VMware virtual machines, providing them with high bandwidth for multimedia-related tasks.

Performance Results

Performance was measured using one of the virtual machines to which a disk was mounted from VMware Data Store.

The following FIO parameters were used:

```
fio --filename=/dev/nvme1n1
--direct=1
--rw=write/read
--bs=192k
--ioengine=libaio
--iodepth=64
--runtime=60
--numjobs=4
--time_based
--group_reporting -
--offset_increment=20% -
--name=throughput-test-job
```

The performance results are shown in the table below:

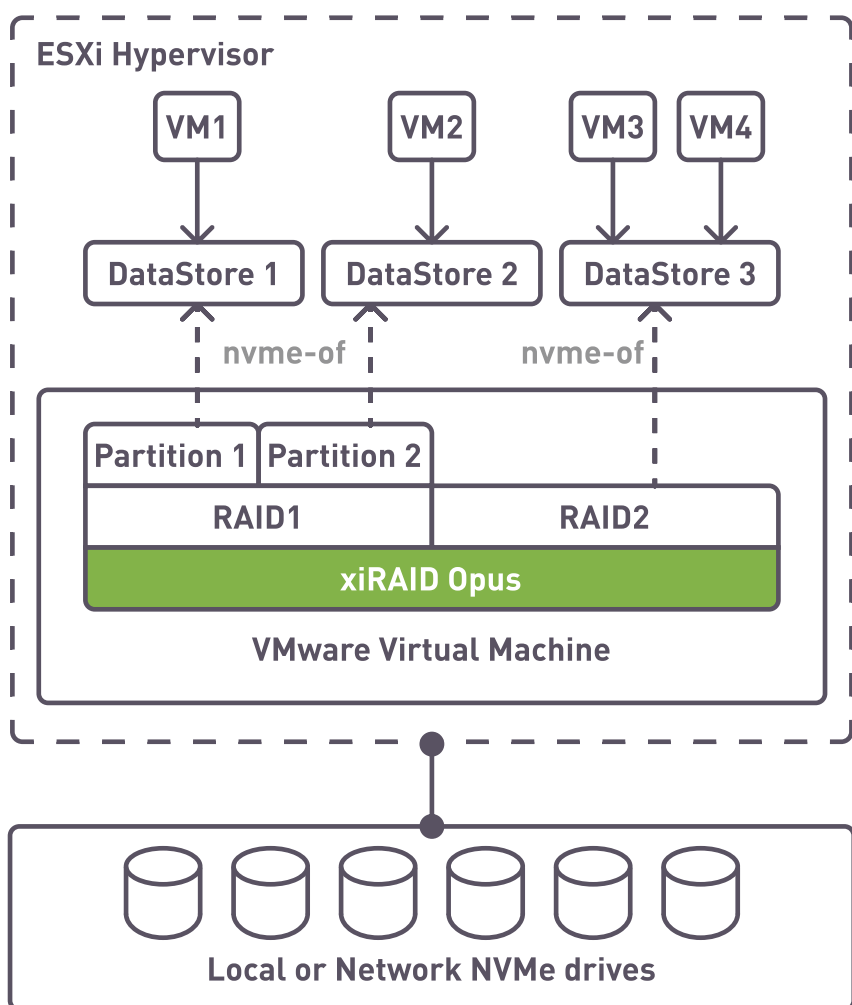
Metric (Tested with FIO)	Sequential Bandwidth (BW)	Average Latency (Read/Write)
Sequential Reads	11.4 GiB/s (12.2 GB/s)	8.23 ms
Sequential Writes	8.5GiB/s (9.1 GB/s)	11.80 ms

The results show that it was possible to achieve the maximum performance, given that the network bandwidth is limited to 100 Gb/s (approximately 12.5 GB/s), which is crucial for real-time 4K/8K video editing and transcoding.

Possibilities for expanding the existing architecture

This architecture can be expanded both at the VMware level (increasing the number of datastores and virtual machines) and at the xiRAID Opus level, increasing the number of RAIDs.

Given the capabilities of xiRAID Opus, administrators can flexibly configure partitions on each RAID device and set up an NVMe-oF subsystem for each partition. Each partition or an entire RAID can utilize QoS, which allows setting IOPS and throughput limits, ensuring predictable performance.



In addition, both local NVMe drives and network drives from EBOF or Linux NVMe-of target can be used as a data storage subsystem. An example of such use is described in a joint publication with Western Digital™: [“Next-Generation Storage for VMware Environments: Virtual RAID Appliance Powered by Western Digital OpenFlex™ Data24 and xiRAID Opus”](#).

Conclusions

The resulting solution based on xiRAID Opus for VMware ESXi provides an easy-to-implement and easy-to-support alternative to dedicated storage systems with no hardware limitations, that can be flexibly scaled as additional drives are added or network bandwidth expands. Comprehensive performance tests conducted on the virtual machine with xiRAID Opus validate its ability to efficiently manage and scale storage resources in VMware virtualized environments while maintaining high performance. In addition, the solution features minimal resource consumption, ensuring seamless integration with existing virtualized infrastructures.