



A High Performance BeeGFS Storage Solution powered by Western Digital's OpenFlex™ Data24 NVMe-oF™ Storage Platform and Xinnor xiRAID



Executive Summary

In high-performance computing (HPC), designing a well-balanced storage system to achieve optimal performance presents significant challenges. The HPC environment (which can consist of hundreds of nodes), processes high volumes of data to solve complex calculations. The HPC environment is such that it must be able to accommodate both large blocks of data for Big Data analytics and many small files for Machine Learning (ML) and Artificial Intelligence (AI) workloads without performance degradation. Organizations seek low-cost, high-performance, high-availability solutions that are also easy to manage whilst scaling to the growing demands of different types of IO-intensive workload challenges.

Composable Disaggregated Infrastructure (CDI) represents the modern architectural approach to data centre infrastructure, disaggregating compute, storage, and network resources into shared pools that can be composed for on-demand allocation.

The Western Digital OpenFlex Data24 NVMe-oF Storage Platform is a vertically integrated storage architecture that leverages an Open Composable Infrastructure (OCI) approach in the form of disaggregated data storage using NVMe-over-Fabrics (NVMe-oF).

Xinnor xiRAID is a lightweight software RAID offering which complements CDI solutions whilst performing very closely to raw device capabilities and exceeding that of traditional hardware RAID solutions.

BeeGFS is an easily deployable hardware-independent POSIX parallel file system developed with a strong focus on performance whilst designed for ease of use, simple installation, and easy management.

The purpose of this document is to showcase a combined solution of BeeGFS cluster deployment with Xinnor xiRAID volumes deployed on OpenFlex Data24 namespaces. This document is not an endorsement of Xinnor xiRAID or BeeGFS by Western Digital, and no warranty of either product is either expressed or implied.

Highlights

- Western Digital's OpenFlex Data24 NVMe-oF Storage Platform. It provides low latency sharing of NVMe™ SSDs over a high-performance Ethernet fabric to deliver similar performance to locally attached NVMe SSDs.
- Xinnor xiRAID is a high-performance software RAID developed specifically for NVMe storage devices to utilize up to 97% of hardware performance capabilities.
- BeeGFS is a high-performance parallel file system designed for performance-oriented environments like HPC, AI, and deep learning workloads. BeeGFS includes a distributed metadata architecture for scalability and flexibility reasons. Its most important aspect is data throughput.

Benefits

- High Performance
- High Availability
- Scalability
- Robustness
- Easy to deploy and integrate with existing infrastructure
- Easy data management
- Optimized for highly concurrent access
- BeeGFS - open-source software with enterprise features

Problem Statement

HPC environments have been served by magnetic disks as a storage backbone for decades. Many applications in the HPC environment now must accommodate both large blocks of data and many small files. For instance, training an ML model might use millions of small files, while a big data analytics workload runs on one massive dataset. There is a need for increased metadata performance in many workloads today. Key components of the HPC cluster, such as compute, storage, and network, need to be evolved to serve these modern workloads. Many distributed parallel file systems have started to support the latest compute and NVMe-oF storage to utilize high performance NVMe SSDs effectively in HPC clusters. BeeGFS combines multiple storage servers to provide a highly scalable shared network file system using striped file contents, with which the high throughput demands of large numbers of clients can easily be satisfied.

A High Performance BeeGFS Storage Solution powered by Western Digital's OpenFlex™ Data24 NVMe-oF™ Storage Platform and Xinnor xiRAID

Modern IO-intensive workloads like Big Data, ML, AI, and Internet of Things (IoT) workloads require data infrastructure designed to scale storage and compute independently to ensure both are provisioned efficiently and effectively. CDI ensures compute, storage, and network resources are placed into shared pools that can be composed for on-demand allocation. This enables compute to become stateless, elastic, and scalable independent of storage. Data durability and availability need to be ensured for an organization to provide continued operations or services. Xinnor xiRAID is a high-performance software RAID that can be used to provide continued availability and fault tolerance for the storage.

In this document we demonstrate a solution that addresses the growing demand for high-performance parallel file systems for HPC clusters by deploying a high-performance distributed parallel file system on a composable infrastructure. Western Digital's validated design for HPC BeeGFS High-Capacity Storage solution with Xinnor xiRAID is a fully supported, easy-to-use, high-throughput, scale-out, parallel file system storage solution with well-described performance characteristics.

OpenFlex Data24 Overview

OpenFlex is Western Digital's architecture that supports OCI through storage disaggregation. The OpenFlex Data24 is a 2U rack-mounted data storage enclosure, built on the OpenFlex platform. This Just-a-Bunch-Of-Flash (JBOF) platform leverages the OCI approach in the form of disaggregated data storage using NVMe-oF. NVMe-oF is a networked storage protocol that allows storage to be disaggregated from compute, in turn makes that storage widely available to multiple applications and hosts. For more details, refer to OpenFlex Data24 NVMe-oF Storage Platform.

Composable Infrastructure seeks to disaggregate compute, storage, and networking fabric resources into shared resource pools that can be available for on-demand allocation (i.e., "composable"). Composability occurs at the software level, while disaggregation occurs at the hardware level using NVMe-oF. NVMe-oF will vastly improve compute and storage utilization, performance, and agility in the data center.

BeeGFS Overview

BeeGFS is a hardware-independent POSIX parallel file system (a.k.a., software-defined parallel storage) developed with a strong focus on performance and designed for ease of use, simple installation and easy management. It is designed for all performance-oriented environments, including HPC, AI, deep learning, life sciences, oil, gas, media, and entertainment.

Learn more about downloading and using BeeGFS along with its end user license agreement at www.beegfs.io.

Xinnor xiRAID Overview

Xinnor xiRAID ensures fast and effective access to data by allowing for the creation of high-performance RAID from NVMe and SAS/SATA SSD's. Designed for the most demanding enterprise-grade tasks, xiRAID is easy to maintain and suited for operating in large server infrastructures. For more details, refer to @Xinnor. Xinnor xiRAID is a software RAID presented by Linux kernel module and management utility (CLI).

- Adjusted for the most popular Linux® distribution (Ubuntu, RHEL, Oracle® Linux, Rocky Linux, Alma Linux)
- Works with local and remote drives.
- Provides RAID as a standard Linux block device.
- POSIX API support.

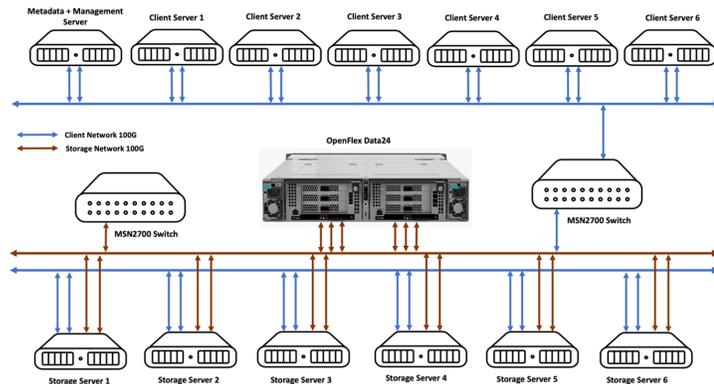
Advantages of the Xinnor Software RAID

Due to its fast coding and decoding ability, xiRAID provides the stable performance levels needed for smooth and uninterrupted business operations. Fast RAID rebuild protects storage from extensive system downtime and mitigates the impact on workflows. This is crucial for data-intensive systems and high-density storage infrastructures where even a single drive failure can cause checksum recalculations for a vast amount of data.

A High Performance BeeGFS Storage Solution powered by Western Digital's OpenFlex™ Data24 NVMe-oF™ Storage Platform and Xinnor xiRAID

BeeGFS on OpenFlex Data24 Deployment Topology

Configuration: a single Data24 with 6 storage BeeGFS storage instances, a single metadata server instance and 6 clients.



Conclusion

BeeGFS on the Western Digital's OpenFlex Data24 Storage with xiRAID addresses the need of IT/HPC with a well-designed solution that is easy to manage and fully supported. The solution includes the added benefit of the open composable infrastructure environment and enables disaggregated compute on the Nvme-oF storage platform. BeeGFS on the Western Digital's OpenFlex Data24 systems shows excellent benchmark results for storage data streaming with both RAID and Non-RAID configuration. This highlights the superiority of Nvme-oF devices over traditional technology for high-end services with different workloads.

In this document we have explained the performance achieved with BeeGFS using xiRAID on OpenFlex Data24. The test results clearly shows that BeeGFS, Xinnor and OpenFlex Data24 has significant advantages of performance and ease of deployment. Additionally, we have shown that using xiRAID user will be able to achieve significant performance with data protection and high availability. This new software technology is well suited to accelerate storage hardware systems to new levels of high performance.

By leveraging the solution offered by Western Digital's enterprises and service providers can

- Minimize the cost of the installation
- Implement high-performance HA file system solution
- Accelerate time to market for new application services

Using this solution based on the latest NVMe-oF and CDI technologies, organizations can quickly deploy a proven, self-service, composable infrastructure solution, helping customers move to a more flexible, variable cost model.

This document is not an endorsement of xiRAID or BeeGFS by Western Digital, and no warranty of the product is either expressed or implied.

For full setup description and step by step implementation, please contact your Western Digital sales team and request the White Paper titled *A High Performance BeeGFS Storage Solution powered by Western Digital's OpenFlex™ Data24 NVMe-oF™ Storage Platform and Xinnor xiRAID*.



IOR Performance Results

The maximum performance achieved with IOR without buddy mirroring on Xinnor RAID5 volumes. This configuration lacks redundancy and presents a risk of data-loss:

BeeGFS performance on Xinnor RAID5 volumes (without Buddy Mirror)

IOR Test Results	GBps
Read	42 GBps
Write	24 GBps

The maximum performance achieved with IOR with buddy mirroring on NON-RAID volumes:

BeeGFS performance on Non-RAID volumes with Buddy Mirror

IOR Test Results	GBps
Read	45 GBps
Write	7.7 GBps

The maximum performance achieved with IOR on without buddy mirroring on NON-RAID volumes:

BeeGFS performance on Non-RAID volumes (without Buddy Mirror)

IOR Test Results	GBps
Read	45 GBps
Write	30 GBps