

Western Digital OpenFlex® Data24-4200 and IBM Storage Scale High-Performance, Disaggregated File Storage for AI and HPC



OpenFlex Data24 4000 Series Storage Platforms

The OpenFlex Data24 4000 series NVMe-oF™ storage platform extends the high performance of NVMe flash to shared storage. The 4000 series provide low-latency sharing of NVMe™ SSDs over a high-performance Ethernet fabric to deliver similar performance to locally attached NVMe SSDs. Western Digital RapidFlex™ NVMe-oF controllers, allows up to six dual pathed hosts to be attached without a switch. The OpenFlex Data24 4000 series uses Western Digital's RapidFlex C2000 Fabric Bridge Adapters to provide up to 12 ports of 100GbE which can connect to RDMA and/or TCP configured host ports.

Challenges

- Storage performance lags compute growth, leaving GPUs clusters underutilized due to I/O bottlenecks.
- Traditional shared storage architectures fail to scale, limiting bandwidth and IOPS as workloads and client counts increase.
- Proprietary interconnects and closed platforms restrict flexibility, increasing cost and slowing infrastructure evolution.
- Complex, high-risk deployments delay time-to-production, requiring extensive tuning and validation before delivering results.

Highlights

- High performance, shared NVMe-oF storage combining OpenFlex Data24 4200 and IBM Storage Scale for AI, HPC, and analytics.
- OCCL validated performance delivering up to 92 GB/s writes and 1.5M+ 4K read IOPS.
- Open, Ethernet based architecture supporting vendor agnostic, dual ported NVMe SSDs.
- Single, POSIX compliant namespace with parallel access across multiple clients.
- Independent scaling of capacity, performance, and clients without disruptive redesign.

IBM Storage Scale

IBM Storage Scale is a high-performance, software-defined parallel file system designed to provide shared access to large-scale datasets across distributed compute environments. Storage Scale enables multiple storage and client nodes to operate in parallel, delivering high aggregate throughput and low-latency access to data through a single, global POSIX-compliant namespace. The platform aggregates block storage devices into Network Shared Disks (NSDs) and distributes I/O across storage servers to maximize concurrency and scalability. Designed for AI, HPC, analytics, and data-intensive enterprise workloads, IBM Storage Scale allows storage and compute to scale independently while maintaining consistent performance under load. Support for high-speed Ethernet networking and NVMe-based storage enables Storage Scale to efficiently exploit modern NVMe-oF architectures without relying on proprietary interconnects.

Open, Composable Infrastructure

At the core of the solution is a composable design that cleanly separates storage, compute, and networking. Storage is provided as a shared NVMe-oF resource, allowing each layer to scale independently as workload requirements evolve. Standards-based Ethernet connectivity ensures broad ecosystem compatibility while preserving long-term flexibility.

The storage layer is built on the Western Digital OpenFlex Data24-4200, a high-density NVMe JBOF platform engineered for extreme aggregate throughput. The platform features a 24-drive NVMe chassis, dual I/O modules for parallel access, and embedded RapidFlex ASICs that enable NVMe-oF and optimize fabric connectivity. Multiple 100 and 200 GbE fabric connections enable very high aggregate bandwidth while maintaining predictable latency. Support for dual-ported, vendor-agnostic NVMe SSDs reinforces the solution's open design.

Solution

The Western Digital Data24-4200 and IBM Storage Scale solution delivers a high-performance, disaggregated storage solution designed for AI, HPC, and data-intensive enterprise workloads. By combining a dense NVMe-over-Fabrics (NVMe-oF) storage platform with a proven, massively parallel file system, the solution provides extreme bandwidth, scalable IOPS, and architectural flexibility without proprietary lock-in.

Western Digital OpenFlex Data24-4200 and IBM Storage Scale High-Performance, Disaggregated File Storage for AI and HPC

Parallel File System Integration

IBM Storage Scale provides the shared file system layer, aggregating NVMe-oF block devices into a single POSIX-compliant namespace. Dedicated Storage Scale storage nodes connect directly to the Data24-4200 over the NVMe-oF fabric, with NVMe-oF traffic isolated from client and internal file system networks. NVMe devices are configured as Network Shared Disks (NSDs), allowing Storage Scale to fully exploit NVMe-level parallelism while maintaining advanced capabilities such as metadata services, policy-based data placement, and concurrent multi-client access.

From the client perspective, applications access the shared file system over standard Ethernet. I/O requests are distributed across multiple storage nodes, which perform parallel NVMe-oF operations directly against Data24-4200 namespaces. This streamlined, end-to-end data path minimizes overhead and enables high sustained bandwidth and scalable IOPS

Validated Performance

The solution is validated through the Western Digital Open Composable Compatibility Lab (OCCL), which provides structured interoperability testing, guided proofs of concept, and documented best practices. OCCL validation ensures that the integrated solution behaves predictably under real-world, data-intensive workloads and reduces deployment risk for customers.

Initial testing using gpfspcrf, with IBM Storage Scale accessing the NVMe-oF file system directly, demonstrates the performance potential of the solution:

- Sequential performance per storage server of approximately 65 GB/s writes and 55 GB/s reads.
- Sustained 92 GB/s aggregate sequential write bandwidth across two storage servers.
- Strong random I/O scalability, exceeding 1.5 million 4 KB random read IOPS and 420K 4 KB random write IOPS with two servers.

These results highlight strong write-side scaling and excellent random I/O performance, with ongoing optimization focused on improving multi-node sequential read scaling.

Scalable and Future-Ready

The solution is intentionally modular. Capacity scales by adding NVMe drives or additional Data24 enclosures, while performance scales through additional storage nodes, network bandwidth, or client concurrency. IBM Storage Scale's mature management and policy framework supports operational growth as environments expand.

By combining an open NVMe-oF storage platform, a proven parallel file system, and OCCL-validated reference designs, this solution provides a clear, low-risk path for organizations seeking to modernize shared storage infrastructure while preserving performance, openness, and future extensibility.

