

Solution Brief

Host-Managed SMR at Near-Full Hardware Throughput Ultrastar® Data102 3000 with LeilOS

Challenges

- Unstructured data from AI/ML, analytics, media and backup is growing faster than storage budgets.
- Flash is too costly for capacity tiers; tape is too slow for active data.
- SMR improves cost per TB, but most file systems cannot use host-managed SMR.
- Software overhead can leave hardware throughput unused.
- Power, cooling and floor space grow with every petabyte.
- Data must stay protected for years without costly migrations.

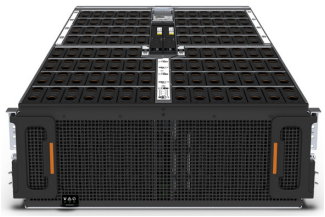
Highlights

- **Up to 3.26 PB raw in one 4U enclosure** with the latest high-capacity Ultrastar SMR drives.
- **18.0 GB/s reads and 13.8 GB/s writes** from one server and one Data102 3000.[1]
- **99.5% of achievable read and 99.7% of achievable write throughput**, measured against fair, real-world ceilings.
- **One 24Gb/s SAS-4 HBA** serves all 102 drives.
- **8+2 erasure coding:** any two drives can fail at once with no data loss, and 80% of raw capacity is usable.
- **Reuse existing CMR hardware** and extend it with new SMR or CMR capacity in one namespace, scaling without downtime.

Together, Ultrastar Data102 3000 and LeilOS deliver petabyte-scale storage at near drive-limit throughput with erasure-coded data protection, helping organizations cost-effectively scale AI data lakes, training datasets, and growing repositories of AI-generated data. The solution scales capacity and performance by adding drives or nodes without downtime, providing an efficient foundation for rapidly growing AI infrastructure.

The solution pairs Ultrastar Data102 3000 high-density storage platforms and host-managed SMR drives with LeilOS, a distributed, POSIX-compliant file system built for HDD-based storage. The Data102 3000 provides dense, cost-efficient capacity, while LeilOS makes host-managed SMR practical for demanding workloads.

Ultrastar Data102 3000 Storage Platforms



The Ultrastar Data102 3000 Storage Platform redefines what's possible in modern storage infrastructure. As a foundational building block for disaggregated and software-defined storage (SDS) architectures, it delivers massive capacity without sacrificing flexibility — empowering organizations to strike the right balance between performance and cost across a variety of workloads. Patented IsoVibe™ vibration isolation technology counteracts drive-induced vibration to protect performance under load, while ArcticFlow™ thermal zoning channels airflow precisely through the center of the chassis for targeted, highly effective cooling. Packed with up to 3.26PB[2] of raw storage in a compact, efficient form factor, the Data102 3000 is purpose-built for the demands of scalable data-intensive environments.

Features

- Up to 102 Ultrastar HDDs (SAS or SATA), including host-managed SMR
- Up to 3.26 PB of raw storage in 4U
- Choose dual-port SAS for high availability or single-port SATA for low cost
- Up to 12 x 24Gb/s SAS-4 host connections

[1] Performance measured on a Data102 3000 populated with 102 × 20TB Ultrastar DC HC650 host-managed SMR drives (2.04 PB raw), one server and two clients; see Performance Details and Configuration Details. Higher-capacity Ultrastar SMR drives are expected to deliver similar per-drive sustained read throughput, so comparable read results are expected with them.

[2] With 102 × 32TB drives. One terabyte (TB) is equal to one trillion bytes and one petabyte (PB) is equal to 1,000TB. Actual user capacity may be less due to operating environment.

Host-Managed SMR at Near-Full Hardware Throughput Ultrastar Data102 3000 with LeilOS

LeilOS

Leil Storage builds LeilOS, HDD-native data storage software that makes high-density SMR drives practical for mainstream enterprises and brings hyperscale storage economics on-premises. LeilOS is a distributed, POSIX-compliant file system with separate metadata and data paths, so clients read and write many drives in parallel. Data is stored across chunkservers coordinated by a master, with a metalogger keeping a copy of the metadata. Every chunk carries CRC checksums, and each file or directory can be protected with Reed-Solomon erasure coding or replication.

Host-managed SMR drives are served natively by the LeilOS SMR chunkserver, with metadata kept apart from the data on NVMe™ SSDs. LeilOS runs in Docker containers or as host packages.

Existing hardware stays in service. Servers and enclosures already running CMR drives can be reused and extended with fresh SMR or CMR capacity in the same namespace, so capacity grows without a migration.

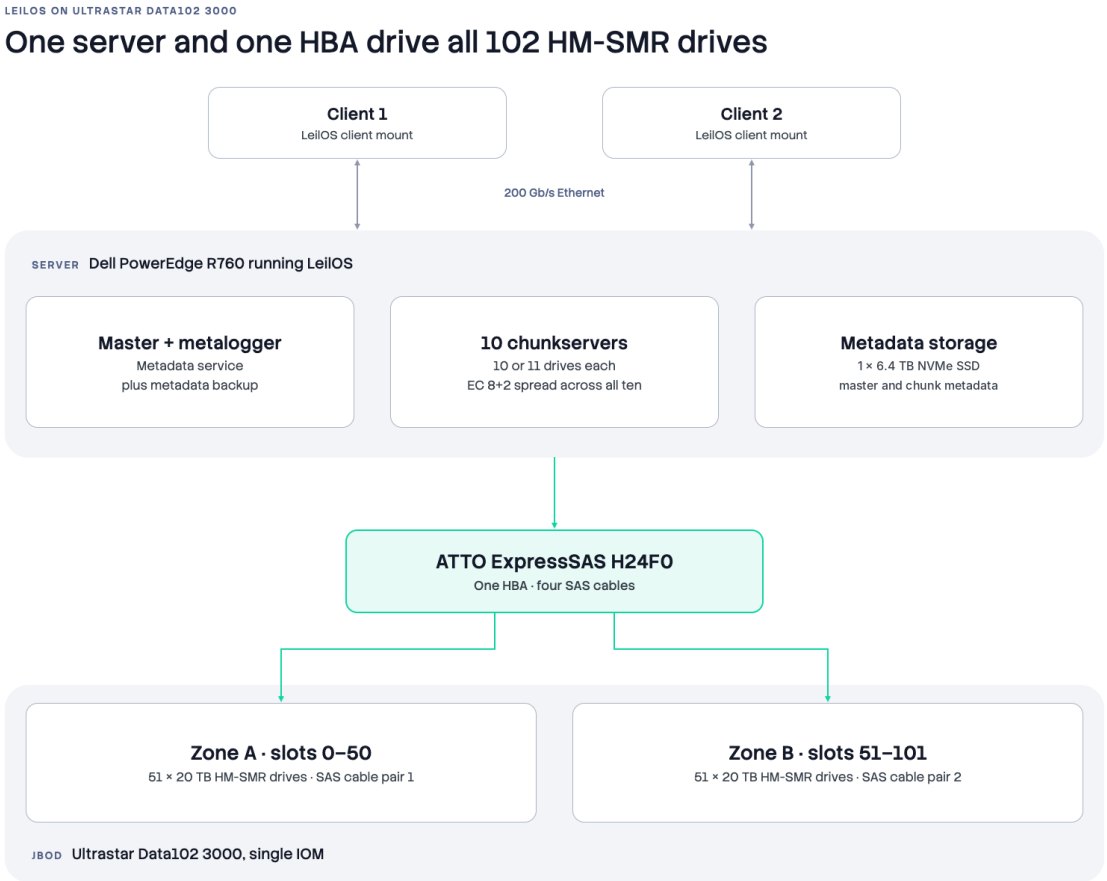
Key Capabilities

- Host-managed SMR, CMR, and mixed capacities in one namespace
- Erasure coding and replication per file or directory
- CRC checks and background scrubbing
- Instant copy-on-write snapshots
- Scale out without downtime
- S3, NFS and SMB; native Linux and Windows clients
- Petabyte to exabyte-scale

Reference Architecture

One Dell® PowerEdge® R760 server attaches a fully populated Ultrastar Data102 3000 through a single ATTO ExpressSAS H24F0 24Gb/s SAS-4 HBA. Four SAS cables connect the HBA to the enclosure, which is zoned into two halves of 51 drives, each reached by its own cable pair. This single-HBA design matched a two-HBA configuration to within 0.3%, removing a card and a PCIe® slot from every system.

On the server, LeilOS runs a master, a metalogger and ten chunkserver instances of 10 or 11 drives each. Ten is what 8+2 erasure coding needs: the eight data parts and two parity parts of every stripe land on ten different chunkservers, so 80% of what the drives write is user data. Here master and chunk metadata share one mirrored NVMe pair; the OS drives or a dedicated pair both work. Clients mount LeilOS over 200 Gb/s Ethernet and read and write the chunkservers in parallel.



Host-Managed SMR at Near-Full Hardware Throughput Ultrastar Data102 3000 with LeilOS

Performance Details

Two client machines ran fio simultaneously against the mounted file system, reading and writing large files sequentially; their results were summed. All 102 drives were in use at every point, and every test phase started from an empty, freshly formatted file system. The raw drives were also measured with no file system, to establish what the hardware itself can deliver.

A Fair Ceiling for HDD Performance

With 24Gb/s SAS-4 connectivity, the path between server and enclosure no longer holds the drives back: raw throughput scales across all 102 drives, so the drives themselves become the limit. That makes it possible to state precisely how much of the hardware's real-world capability a file system delivers — measured against two fair ceilings:

Fair write — 80% of raw write. With 8+2 erasure coding, a fifth of every write is parity, so at most 80% of what the drives write can reach clients as data.

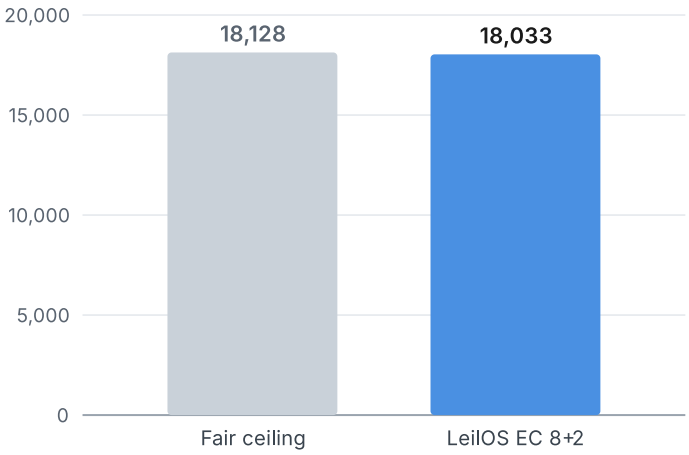
	Read	Write
Raw drives, no file system	24.30 GB/s	17.30 GB/s
Fair ceiling	18.13 GB/s	13.84 GB/s
LeilOS, EC 8+2, at 80 jobs	18.03 GB/s	13.79 GB/s
vs fair ceiling	99.5%	99.7%

Fair read — raw read with twenty concurrent streams per drive. One sequential stream gives a drive its top read rate. A file system reads many files in parallel, each split across drives by erasure coding, so at full load every drive serves twenty streams at once.

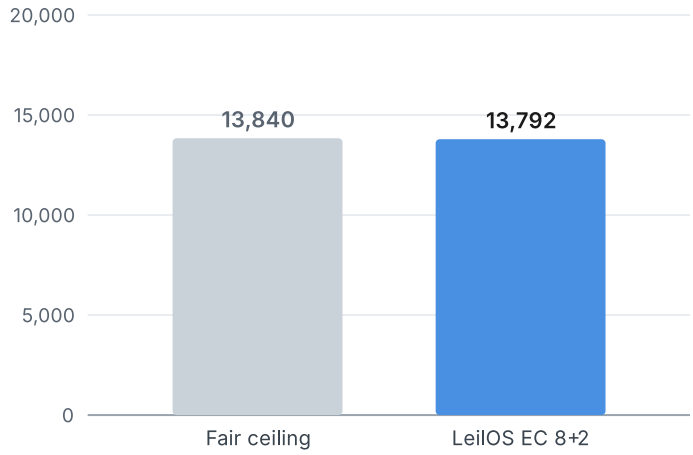
Sequential throughput against the fair ceiling, LeilOS EC 8+2 (MB/s)

102 host-managed SMR drives, 8+2 erasure coding, two clients with 80 files open on each.[3] The fair ceiling is the most any file system can deliver to clients on this hardware.

Reading
of the fair ceiling
99.5%



Writing
of the fair ceiling
99.7%



Reads climb as more files are served in parallel and reach 18.0 GB/s at 80 jobs per client. That is 177 MB/s per drive, against 178 MB/s from the bare drives with twenty concurrent read streams on each.

Writes reach the fair ceiling at four to eight jobs per client and hold it: from 16 to 102 jobs per client they stay between 13.73 and 13.86 GB/s. That is 169 MB/s per drive written to the media, the same rate as the bare drives with no file system in the way.

Headroom. At peak, LeilOS ran at the drives' fair ceiling with CPU and network to spare: 16% of 64 threads and 144 of 200 Gb/s. For large sequential files, throughput grows with the number of drives and servers, given enough files in parallel.

AI checkpointing. The same system ran the MLPerf® Storage v3.0 checkpointing workload (closed division parameters): eight processes over the two clients wrote and read a 105 GB checkpoint through the mounted file system at 8.07 GB/s save and 8.79 GB/s load.

[3] Measured with 102 × 20TB Ultrastar DC HC650 host-managed SMR drives; see Configuration Details.

Host-Managed SMR at Near-Full Hardware Throughput Ultrastar Data102 3000 with LeilOS

Configuration Details			
Storage platform	Ultrastar Data102 3000 (UD102-3200), single IOM, IOM firmware 1000-116	Server	Dell PowerEdge R760
Drives (as tested)	102 × Ultrastar DC HC650 20TB host-managed SMR, SATA (WSH722020ALN601), fw PCGSX421 — 2.04 PB raw	CPU/Memory	2 × Intel® Xeon® Gold 6444Y, 16C/32T each - 256 GB DDR5-5600
HBA	1 × ATTO ExpressSAS H24F0, fw 8.17.1.0 (21), esas6hba 1.04.4-trc	Metadata	Master: 2 × 6.4TB NVMe SSD, RAID 1 - Chunk: separate NVMe SSDs — model, count, capacity
Cabling / zoning	4 SAS cables to one HBA; two zones of 51 drives	Network	NVIDIA ConnectX®-7 200 Gb/s; Spectrum SN4600
LeilOS	5.12.0 - erasure coding 8+2	Clients	2 × server, each 1 × AMD EPYC™ 9124 (16C), 768GB, ConnectX-7
Host OS	Ubuntu 26.04	Tool	fio 3.41, sequential, direct I/O; 1 MiB requests through LeilOS, 8 MiB on the bare drives - MLPerf Storage v3.0

Advantages

- **Near full hardware performance:** 99.7% of achievable write and 99.5% of achievable read throughput for an 8+2 object storage file system on a fully populated Data102 3000, with CPU and network headroom to spare.
- **Host-managed SMR without a penalty:** SMR drives served natively by LeilOS, writing at the same per-drive rate as the bare drives. Existing CMR hardware can be reused in the same namespace.
- **One HBA per enclosure:** a single ATTO ExpressSAS H24F0 carries all 102 drives and matched two cards to within 0.3% — one card, PCIe slot and NUMA boundary fewer in every system.
- **Efficient data protection:** 8+2 erasure coding keeps 80% of raw capacity usable, with the ten parts of every stripe on ten different chunkservers; any two drives can fail at once without data loss. With 32TB drives, that is up to 2.61 PB usable per enclosure, against 1.63 PB with mirroring.
- **Flexible deployment and data services:** Docker containers or host packages CRC checksums and scrubbing, snapshots, POSIX, S3, NFS and SMB, Linux® and Windows® clients, scale-out without disruption
- **Hardware resilience:** IsoVibe, ArcticFlow and hot-swap; lower power, cooling and space per PB for AI/ML, active archive, backup, media and analytics.

Conclusion

LeilOS on the Ultrastar Data102 3000 delivers near-full hardware throughput for host-managed SMR. From one server and one HBA, it reached 18.0 GB/s reads and 13.8 GB/s writes[4] across 102 SMR drives with 8+2 erasure coding protecting every write — 99.5% and 99.7% of the fair ceilings of the hardware. With the latest high-capacity Ultrastar SMR drives, the same 4U enclosure scales to 3.26 PB raw, providing a scalable foundation for AI infrastructure that combines dense, cost-efficient capacity with near drive-limit performance.

[4] Measured with 102 × 20TB Ultrastar DC HC650 host-managed SMR drives. Unverified MLPerf® Storage v3.0 Closed Checkpointing, llama3-8b. Result not verified by MLCommons Association. The MLPerf name and logo are registered and unregistered trademarks of MLCommons Association in the United States and other countries.

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

© 2026 All rights reserved. WD, the WD Design and Ultrastar are registered trademarks or trademarks of Western Digital Corporation or its affiliates in the U.S. and/or other countries. Broadcom is among the trademarks of Broadcom. Dell and PowerEdge are trademarks of Dell Inc. Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries. The NVMe is a trademark of NVM Express, Inc. PCIe is a registered trademark and/or service mark of PCI-SIG in the U.S. and/or other countries. All other marks are the property of their respective owners. Product specifications subject to change without notice. Pictures shown may vary from actual products. Not all products are available in all regions of the world.

November 2026 / Revision 01