High-Availability NVMe Storage with
xiRAID + Lustre for the Helma Supercomputer

About NHR@FAU

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) hosts the Erlangen National High-Performance Computing Center (NHR @ FAU).

NHR@FAU is one of nine national academic centers in Germany focused on HPC and supporting a wide range of AI-focused research, in areas such as:

1. **Deep Learning for scientific applications:**
   Physics simulations, biomedical image analysis, materials science.

2. **Natural Language Processing (NLP):**
   Large language models (LLMs), multilingual translation, text mining.

3. **Computer vision:**
   High-resolution image recognition, microscopy image segmentation, object detection.

4. **Reinforcement learning and robotics:**
   Simulation-heavy environments requiring real-time data feeds.

5. **AI for engineering & Smart Manufacturing:**
   Digital twins, CFD (Computational Fluid Dynamics) with AI-assisted optimization.

These applications often involve massive datasets, frequent I/O operations, and long GPU training runs.

In 2024, NHR@FAU implemented a storage cluster made of seven storage servers, configured with 24 PCIe Gen4 SSDs. These servers used Xinnor's xiRAID to provide data protection services and Lustre Open Source to provide fast storage to their GPU cluster called Alex. The cluster was implemented as scratch storage and is still in production at NHR@FAU. The solution has operated without issues for over a year, reaffirming stability of their software consisting of Xinnor xiRAID and Lustre.

More recently, NHR@FAU has deployed a new GPU cluster named Helma. The system is dedicated to AI / ML, atomistic simulation, and quantum-chemistry research and is co-funded by the Bavarian "BayernKI" initiative, the Federal Government and the State of Bavaria, and investments from some other Bavarian universities. This cluster combines 192 dual-socket AMD EPYC 9554 "Genoa" compute nodes with 768 NVIDIA H100/H200 GPUs, making it the most powerful university-owned AI supercomputer in Germany and #51 on the June 2025 TOP500 list.

Storage Challenge

The typical AI/ML workload on the Helma cluster exhibits the following characteristics:

- TB-scale training datasets (images, videos, text corpora).

- Small file access random reads (e.g., JPEGs, PNGs, text files).

- Multi-GPU jobs (e.g., H1/200s), distributed training using MPI/Horovod/DeepSpeed.

- Many users running concurrent jobs; high metadata and throughput demands.

- Frequent writes of model weights (every few minutes), potentially TBs per checkpoint.

- On-the-fly data augmentation or tokenization requiring fast I/O pipelines.

This workload requires a storage capable of:

- Sustained multi-hundred-GB/s reads and >1 M IOPS metadata were required to keep 768 GPUs busy.

- High availability, the system had to survive server and drive failures with close to zero downtime, a non-negotiable requirement for a national HPC center.

- FAU wanted TOP-tier performance without a multi-rack, bespoke appliance and with the freedom to choose commodity hardware.

Solution

MEGWARE and Xinnor designed a half-rack, all-NVMe Lustre file-system protected by xiRAID software RAID storage systems running on 10x Celestica SC6100 PCIe Gen5 storage controllers. Each controller was configured with:

✓ **CPU / RAM:**
2 x AMD EPYC 9454P, 384 GB DDR5 (1 per server)

✓ **Metadata storage:**
4x Phison Pascari PCIe 5.0 X200E 6.4 TB (3 DWPD)

✓ **Data storage:**
20x Phison Pascari PCIe 5.0 X200P 30.72 TB (1 DWPD)

✓ **Network:**
4x NVIDIA ConnectX-7 NDR 400 IB (2 per server)

✓ **OS:**
AlmaLinux 9 on all servers

✓ **Parallel file system:**
Lustre 2.16.1 integrated with Corosync & Pacemaker

✓ **Data Protection:**
xiRAID Classic 4.2 integrated with Corosync & Pacemaker

High Availability

xiRAID Classic supports the integration of Pacemaker failover software to enable building a highly available cluster. In case of drive failures, xiRAID takes care of drive protection while in case of a complete server node failure, the integration with Pacemaker and Corosync failover software allows to automatically fail-over the RAIDs from the failed server to the surviving one. Once the failed server is restored, the original RAID groups are failed-back.

This architecture allows the solution to survive multiple drive failures as well as a complete server failure in each pair of servers.

To implement this solution, MEGWARE selected Celestica's SC6100, a high availability storage controller. The SC6100 is comprised of two independent server nodes sharing 24 PCIe Gen 5 NVMe solid state drives. The redundant architecture of the SC6100 enables the system to continue operation in case one node fails.

The system allocates four PCIe lanes to each drive, two lanes dedicated to each server node, thereby providing a redundant path for each drive. To assure performance optimization, each drive has been divided into two namespaces, allowing concurrent utilization of each drive by both SBB nodes. Therefore, a total of 48 virtual drives (namespaces) are created per server.

On each Celestica SC6100 system, the following RAID topology has been created:

- **4 drives (8 namespaces)** with 2 groups of RAID10 (2+2) for Metadata using the write intensive 6.4TB drives.

- **20 drives (40 namespaces)** with 4 groups of RAID6 (8+2) for Object Storage using the read-intensive 30.72TB drives.

Celestica SC6100
Pacemaker-based Dual-node Cluster 1

Celestica SC6100
Pacemaker-based Dual-node Cluster 10
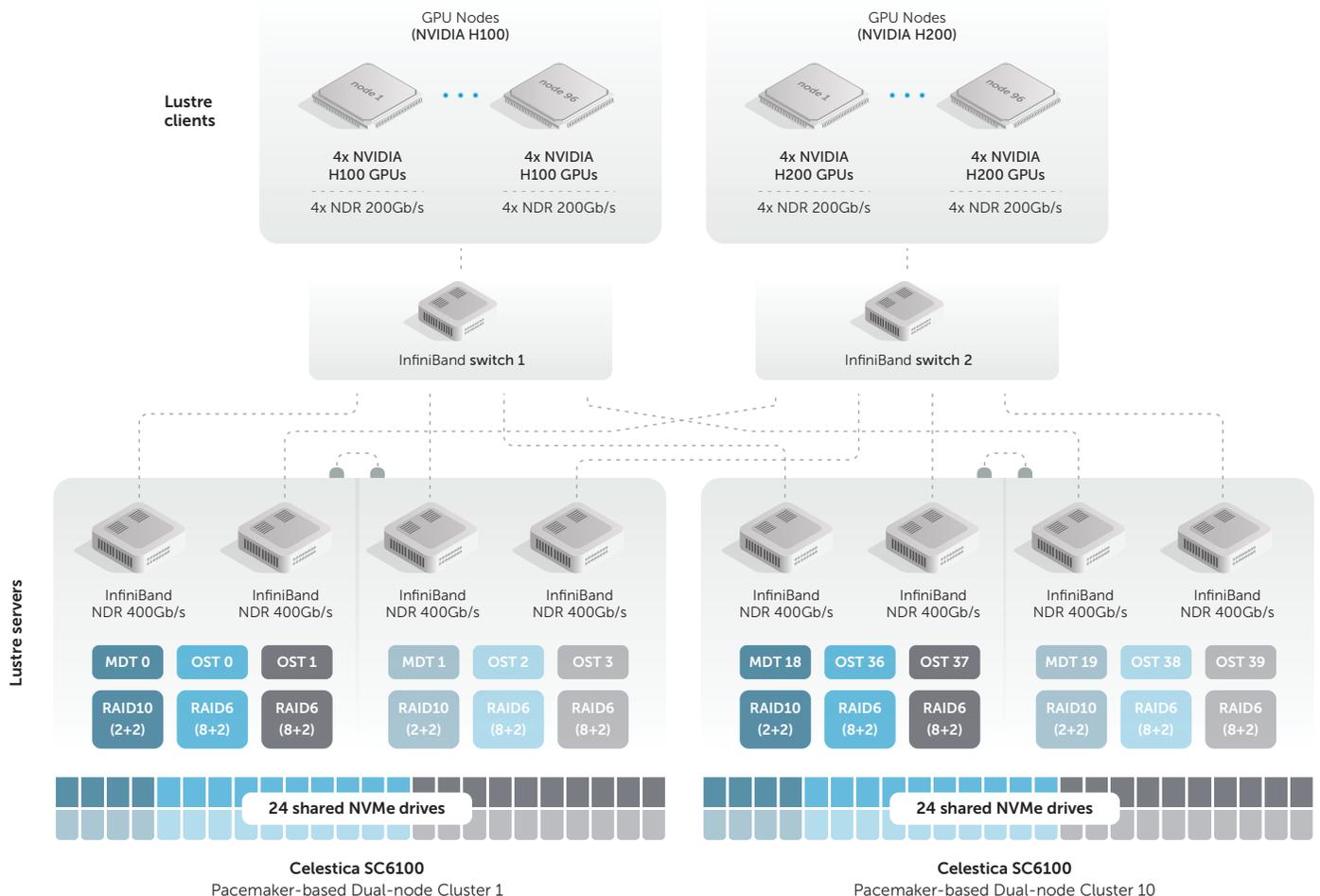
IO500 Results

The IO500 is the premier global ranking system for high-performance storage solutions, evaluating systems across multiple dimensions including bandwidth, IOPS, and metadata operations.

After deploying the system at FAU, MEGWARE ran the IO500 benchmark on the Helma cluster and submitted the results to the IO500 organization.

The Helma cluster achieved extraordinary results, securing the #1 position among Lustre-based solutions and the #3 ranking in the global IO500 benchmark.

IOR Easy Read:
1,798.77 GiB/s

IOR Easy Write:
811.33 GiB/s

MDtest Easy Stat:
8,221.83 kIOPS

Find Operations:
3,016.99 kIOPS

Overall Bandwidth:
438.62 GiB/s

Overall IOPS:
1,604.84 kIOPS

Total IO500 score: 838.99

Business & Operational Benefits

No single point of failure:
Dual-port NVMe, mirrored MDTs, and distributed OST pools tolerate simultaneous server and drive failures.

Performance at ½-rack scale:
Delivers Top-3 IO500 performance with one tenth the footprint of many competitors.

Commodity economics:
100% off-the-shelf hardware; no vendor lock-in.

Green efficiency:
Maximizing GPU utilization with fewer servers → lower power & cooling OPEX.

Future-proof:
PCIe 5.0 NVMe and NDR 400 InfiniBand leave headroom for the next Helma expansion.

Conclusion

The implementation of **xiRAID combined with Lustre in a high-availability configuration** brought substantial benefits to the **NHR@FAU Helma project,** directly enhancing the performance, reliability, and scalability of its AI workloads. This storage solution eliminated I/O bottlenecks that previously limited GPU utilization, enabling researchers to fully leverage Helma's advanced GPU capabilities for large-scale deep learning, data-intensive scientific computing, and distributed AI training.

With **xiRAID's high-throughput, fault-tolerant architecture and Lustre's parallel I/O performance,** Helma now supports concurrent access to petabyte-scale datasets with low latency and robust fault recovery. High availability ensured seamless operation even during hardware or node failures—minimizing downtime and maintaining productivity across multi-user environments. As a result, the storage infrastructure now matches the demands of cutting-edge AI research, helping NHR@FAU accelerate innovation across disciplines while optimizing resource usage and system uptime.

Having reached the podium of the IO500 list is a further validation of the ability of xiRAID to maximize the performance of the underlying hardware without any compromise on reliability and at the same time minimizing the cost of the deployment.

022026