

Solution Brief

Reliable AI Training with QSAN Storage and Altos GB10

Challenges

- Training jobs running on local NVMe storage are vulnerable to hardware failure, system crashes, or unexpected interruptions
- Without a proper checkpoint mechanism, any failure can result in losing hours or days of computation, wasting GPU resources and time
- Local storage lacks enterprise-grade features such as RAID protection, versioning, and centralized backup

Solution

- Training jobs periodically save checkpoints and ensures continuous progress capture without manual intervention
- Checkpoints are stored on QSAN X1 with RAID protection, enterprise backup features to eliminate risk associated with local disk failures
- High-speed data transfer, RDMA mode enables low-latency checkpoint writing and reduces CPU overhead, thus providing stable data throughput for continuous workloads

Risk in Long-Running AI Workloads

In the era of LLMs (Large Language Models), AI training and fine-tuning are intensive, long-running processes that can span days or weeks. Enterprises operating in regulated industries face these primary hurdles when deploying Generative AI:

- **Vulnerability of Local Storage:** Modern AI training often relies on local NVMe drives. If a drive fails or a system crashes during a long-running job, the entire computational progress and the associated electricity and labor costs is lost.
- **Manual Backup Inefficiency:** Relying on manual data backups is prone to human error and often results in significant "data gaps" between the last backup and the point of failure.
- **Scale vs. Stability:** As enterprises scale to multi-node or multi-GPU environments, the statistical probability of a hardware interruption increases, making "stability over quantity" a critical requirement.

Altos GB10 F1 + QSAN XN1 Integration

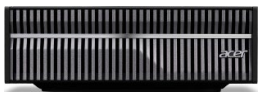
Altos provides a "Blueprint for Resilience" by integrating the high-performance Altos GB10 F1 compute node with an enterprise-grade QSAN XN1 unified storage via high-speed RDMA connectivity.

- **Automated Checkpointing Loop:** The system is configured to automatically write LLM fine-tuning "checkpoints" to the QSAN XN1 at defined intervals. This shifts the risk from volatile local storage to a protected, RAID-enabled environment.
- **High-Speed Data Path:** Utilizing NVIDA ConnectX-7 and RDMA mode, the architecture ensures that large model weights can be backed up rapidly without throttling the GPU's training performance.
- **Rapid Recovery and Continuity:** In the event of a hardware failure, the training job can be instantly resumed from the most recent checkpoint stored on the QSAN XN1 storage, reducing potential downtime from days to minutes.

Resilient GPU Training with Automated Checkpoints

Automated checkpointing with QSAN X1 unified storage ensures resilient GPU training on Altos GB10, enabling fast recovery and zero-loss progress.

Recommended ALTOS Models



Altos BrainSphere™
GB10 F1

- NVIDA® GB10 Grace Blackwell Superchip
- NVIDA® Blackwell GPU 128 GB Unified Memory
- 1 PetaFLOP FP4 AI Computing Performance
- Up to 4 TB NVMe Storage Capacity
- NVIDA® ConnectX®-7 SmartNIC
- One-click AI Developmnet and Deployment Platform

Recommended QSAN Models



QSAN XN1205-4A

- Desktop Tower NAS
- AMD® Ryzen™ 4-core
- 8 GB DDR5 DIMM up to 64 GB
- 3.5” Slot x 4 + 2.5” Slot x 1 + M.2 NVMe x 2
- Maximum Raw Capacity up to 153 TB



QSAN XN1209-4A

- Desktop Tower NAS
- AMD® Ryzen™ 4-core
- 8 GB DDR5 DIMM up to 64 GB
- 3.5” Slot x 8 + 2.5” Slot x 1 + M.2 NVMe x 2
- Maximum Raw Capacity up to 276 TB

QSAN ALTOS an Acer Group Company

© 2026 QSAN Inc. All Rights Reserved. QSAN, the QSAN logo, and the marks listed at www.qsan.com are trademarks of QSAN Inc. Other company and product names may be trademarks of their respective owners.
April 2026