DELL Technologies

Top 10 Reasons

Top 10 reasons to choose Dell PowerScale for AI workloads

Dell PowerScale delivers parallel performance, flexible architecture, and enterprise-grade cyber resilience from edge to exascale for AI, HPC, and analytics workloads.

1 | Parallel performance architecture keeps GPUs fully fed with line-rate data delivery.

PowerScale delivers parallel file performance with NVIDIA GPUDirect Storage, NFS over RDMA, and pNFS, enabling near line-rate data delivery from multiple nodes. This reduces I/O bottlenecks, keeps GPUs fully utilized, and maximizes AI training and inference throughput.

2 | Optimized RAG and LLM inference with 19x faster time to first token performance[i].

PowerScale and ObjectScale accelerate RAG and LLM inference with KV cache offload, delivering faster time to first token versus standard vLLM configurations. RDMA-accelerated NFS/S3 access and MetadataIQ with NVIDIA NeMo/NIM connectors enable real-time metadata extraction and low-latency retrieval for highly responsive RAG applications

3 | Certified by NVIDIA for DGX SuperPOD, delivers validated performance for AI/ML.

PowerScale F710 is certified for NVIDIA DGX SuperPOD and validated in NVIDIA Cloud Partner and NVIDIA Certified Enterprise Storage programs, with reference architectures for DGX H100 and DGX B200. This delivers benchmarked storage performance and compatibility, eliminating storage bottlenecks and enabling confident GPU scaling for AI/ML.

4 | Best-in-class efficiency delivers up to 5x less rack space, 72% lower power consumption[ii].

PowerScale F710 NVIDIA-certified performance reduces infrastructure overhead, requiring up to 88% fewer backend switches[ii]. This minimizes network complexity and maximizes data center space for GPUs, making AI storage investments viable in power- and space-constrained environments.

5 | Flexible architecture scales from 3 nodes to hundreds without forklift upgrades.

PowerScale clusters scale from 3 to hundreds of nodes in a single global namespace, supporting tens of petabytes and multi TB/s throughput. Organizations can mix all flash, hybrid, and HDD nodes to optimize performance and capacity across the data lifecycle without disruptive replatforming or upgrades.

6 | Unified multi-protocol access across data center, edge, and cloud.

PowerScale supports simultaneous NFS, SMB, S3, HDFS, HTTP/FTP, and REST access in a single namespace, on-premises and in public cloud. Applications access the same data via multiple protocols without copy, movement, or format conversion, simplifying integration and enabling protocol agnostic access across data center, edge, and cloud.

7 | Open, modular AI data platform foundation for the full AI lifecycle.

PowerScale, a storage engine for the Dell AI Data Platform, underpins an open, modular stack supporting all AI lifecycle stages—ingest, preparation, training, fine-tuning, and inference. Open formats like Iceberg and Delta, plus integrations with Elastic, Starburst, and NVIDIA, ensure flexibility and avoid vendor lock-in.

8 | Industry-leading 2:1 data reduction guarantee program lowers capacity and cost risk.

PowerScale delivers at least 200% effective capacity via inline compression and deduplication across NVMe, hybrid, and HDD tiers. By tying efficiency to contract backed infrastructure outcomes, it provides predictable capacity planning and cost control as unstructured data volumes grow.

9 | Secure by design with integrated cyber resilience, real-time detection, air-gap recovery.

PowerScale combines Ransomware Defender, operationally air gapped cyber vaults, and the PowerScale Cybersecurity Suite for real time threat detection and recovery. Integrations with Superna, SIEM, and SOC tools automate detection, response, and airgap workflows, delivering end-to-end cyber resilience for unstructured data.

10 | Proven enterprise platform with market leadership and tens of thousands of customers.

PowerScale, a trusted enterprise platform, boasts years of maturity in Exabyte+ environments and thousands of customers – including 1,500+ running GPU-intensive workloads. With market leadership, it ensures reliable performance for file consolidation, research workloads, and AI factories

[i]Based on internal Dell Technologies testing using the LLaMA-3.3-70B Instruct model with Tensor Parallelism=4. Tests measured Time to First Token (TTFT) performance with a 100% KV Cache hit rate, comparing Dell's vLLM + LMCache + NVIDIA NIXL stack on PowerScale and ObjectScale storage to a baseline standard vLLM configuration. Actual results may vary. Nov, 2025.
[ii]Based on Dell internal analysis of NVIDIA-validated reference designs for 64 SU configurations that adhere to the NVIDIA Cloud Platform Reference Architecture specification for high-performance storage, August 2025.

Learn more about Dell PowerScale storage solutions.

Contact a Dell Technologies Expert

Join the conversation

DELL Technologies