HPE Storage

Solution brief

Getting your storage AI-Ready with HPE GreenLake for File Storage

Scale-out file storage with enterprise performance, simplicity, and enhanced efficiency at AI scale

Unleash the power of AI

Everywhere you look, AI is driving transformation in ways never experienced before. Organizations across industries are now leveraging AI to achieve faster innovation, drive faster times to market, and gain a competitive advantage. At the heart of AI is the data that fuels your organization. As you prepare to fully leverage the vast potential of AI, start by making your data and infrastructure AI-ready.

But given the exponential growth of file data, enterprises face a huge challenge in managing, processing, and extracting value from that data. And as file data growth continues to accelerate, IT organizations have difficulty scaling their data infrastructures—to keep up and achieve data-first modernization across the hybrid cloud for AI.

To fully leverage the vast potential of AI, enterprises must resolve the IT problems slowing them down: complex management, storage infrastructure silos, tiered data pipelines, poor performance scaling, and low productivity across edge to cloud.

For file-based and data-intensive applications, IT leaders need a different approach to storage—one that delivers a cloud experience everywhere, runs all workloads on a single infrastructure, maximizes performance at AI scale, and expedites setup and job completion.

HPE GreenLake

A different approach to file storage for AI

HPE GreenLake for File Storage offers enterprise performance at AI scale to span all the stages of AI and accelerate the most data-intensive AI applications, unlocking more value from your data; radical simplicity at AI scale for an intuitive cloud experience that boosts productivity; and enhanced efficiency at AI scale for increased ROI and sustainability.

The solution is managed through the HPE GreenLake cloud and built on HPE Alletra Storage MP hardware, with standard configurations as well as higher compute and storage density options for AI. HPE GreenLake for File Storage is a key product in HPE's rich, comprehensive portfolio of unstructured data and AI storage products. It offers highly differentiated file storage by avoiding trade-offs. It makes storage simple while delivering scalable performance with a unique modular, flexible architecture.

Key features

1. **Enterprise performance at AI scale to unlock more value from all your data—**All-NVMe speed for fast, predictable performance that spans the scale of your data for all the stages of AI; no front-end caching, data movement between media, or tiered data pipelines; multiprotocol NFS and SMB access to the same data; store all data for AI and data-intensive workloads
2. **Simplicity at AI scale for increased productivity—**Streamlined deployment with HPE GreenLake edge-to-cloud platform; increased productivity of data scientists and line of business application owners; single cloud console, simple setup, and fast job completion; enhanced collaboration with a single global namespace
3. **Enhanced efficiency at AI scale for high ROI and sustainability—**Superior data reduction and protection; high flash efficiency and compute and storage densities; reduced data center footprint, power, and cooling requirements for lower carbon footprint

HPE demonstrates its leadership in storage for AI with NVIDIA DGX BasePOD™ certification and NVIDIA OVX™ systems storage validation for HPE GreenLake for File Storage. With the certification, HPE GreenLake for File Storage can be deployed as the storage component of the NVIDIA DGX BasePOD reference architecture. This provides customers with a proven enterprise file storage solution for accelerating GenAI, AI, and GPU-intensive workloads at scale with NVIDIA AI Computing by HPE.

HPE customers can confidently use HPE GreenLake for File Storage to run AI applications with the quick implementation enabled by the DGX BasePOD architecture. With an integrated AI infrastructure, customers can now implement the most impactful, mission-critical AI use cases that DGX BasePOD is designed to address—Health and Life Sciences, Financial Services, Energy, Telecommunications, and Large Language Models (LLMs).

Enterprise performance at AI scale

Performance at scale limitations can be an issue for AI storage—particularly with legacy scale-out NAS solutions with shared-nothing architectures that can scale capacity but can't scale performance linearly to match that capacity.

HPE GreenLake for File Storage accelerates your most data-intensive applications with enterprise performance at AI scale. It's performance that spans all the stages of AI, from data aggregation, data preparation, and training and tuning to inferencing, not just performance that reaches a peak at some point in time for a small dataset. Rather, it's fast, sustained performance that spans the entire scale of your data for the most demanding data-intensive AI applications. Enterprise performance at AI scale helps you to unleash more value from all your aggregated data delivering faster time to insights that give you a competitive advantage.

HPE GreenLake for File Storage has a disaggregated, shared everything, highly resilient modular architecture that allows you to scale performance and capacity independently—and it's designed for exabyte scale. With all-NVMe speed for fast, predictable performance and no front-end caching, data movement between media, or tiered data pipelines, you can supercharge your most data-intensive AI applications. Now with new higher compute and storage density options, you get 2x the throughput and 2x the performance density of the currently shipping release of HPE GreenLake for File Storage,[1] helping you keep pace with your large-scale AI workloads as they grow. In fact, there is a competitive advantage of up to 2.3x the capacity density of Dell PowerScale and Pure Storage FlashBlade.[2] In addition, you get six-9s availability[3] with no rebuild times for controller failures. And you get the flexibility of multiprotocol NFS and SMB access to the same data.

[1] These improvements are based on comparisons to the most recent shipping product featuring the two-node, 2U controller boxes and 2U storage shelves using 30.72 TB NVMe SSDs.

[2] Comparisons are based on raw capacity for capacity density with Dell PowerScale F900 and Pure Storage FlashBlade.

[3] "HPE Storage substantiation," HPE, 2024

Simplicity at AI scale

Legacy file storage for AI can require specialized domain expertise to manage, maintain, and scale—keeping IT and data scientists tied up in routine day-to-day operational tasks rather than focused on faster time to value for new AI models.

HPE GreenLake for File Storage provides an intuitive cloud experience with radically simple file data management that reduces operational and management overhead for your IT staff and increases productivity for your data scientist and line of business app owners. Powered by the HPE GreenLake platform, you get streamlined deployment, easy file share creation, unified storage management with a single cloud console accessible from anywhere, on any device, and automated, nondisruptive upgrades. In addition, you can modernize your data management with a comprehensive suite of over 50 cloud data services and a unified cloud operating experience on the HPE GreenLake platform.

Simple setup and fast job completion also increase the productivity of your data scientists and line-of-business application owners. They get faster results as they're not burdened by cumbersome, legacy file systems that require technical expertise and intricate setup processes just to run. Highly compensated subject-matter experts no longer need to expend valuable time and effort performing IT tasks but can instead focus on their key initiatives.

With its single global namespace, HPE GreenLake for File Storage enhances collaboration for faster time to discovery and innovation.

With the NVIDIA® certification, HPE GreenLake for File Storage customers can leverage the NVIDIA DGX BasePOD reference architecture on which businesses can build and scale AI initiatives. It incorporates a proven design that integrates compute, networking, storage, and software. NVIDIA DGX BasePOD provides a prescriptive AI infrastructure for enterprises that eliminates the design challenges, lengthy deployment cycles, and management complexity to accelerate AI adoption and increase ROI, enabling customers to build their own AI Center of Excellence.

In addition, HPE GreenLake for File Storage can now be used in validated AI configurations with NVIDIA OVX systems in NVIDIA BasePOD reference architecture implementations. NVIDIA OVX servers combine high-performance, GPU-accelerated compute with high-speed storage access and low-latency networking to address a range of complex AI and graphics intensive workloads. This includes the building of 3D virtual worlds using leading 3D software applications, enabling HPE customers to leverage NVIDIA Omniverse Enterprise™—a scalable, end-to-end platform that provides real-time graphics, AI, and immersive digital twin simulation capabilities. NVIDIA OVX provides the compute power required to create the most complex digital twins of factories, buildings, and even entire cities.

Enhanced efficiency at AI scale

Data-intensive AI models have an insatiable appetite for storage capacity, rack space, and power consumption. And legacy NAS solutions with shared-nothing architectures are unable to efficiently scale out to keep pace with the capacity density, cost/TB, and power efficiency demands of AI workloads.

HPE GreenLake for File Storage can slash your AI storage costs with 4x the capacity per RU density, half your power usage—as well as lowering your carbon footprint with industry-leading data reduction, nondisruptive upgrades, and an AI storage as-a-service consumption model. You can also scale performance and capacity independently for higher efficiency at lower cost. This helps eliminate overprovisioning and any performance imbalance between compute and storage resources. And you can also help maximize GPU utilization and therefore GPU ROI with the enterprise performance at AI scale.

The solution provides no-overhead snapshots and native replication, superior flash efficiency, and data reduction with the similarity algorithm, which, unlike compression and deduplication, reduces data with both a global and fine-grained approach. Savings are 2:1 reduction for life sciences data; 3:1 for pre-reduced backups, pre-compressed log files, and HPC and animation data; and 8:1 for uncompressed time-series data.[4] Support for optimized GPU utilization with InfiniBand, NVIDIA GPUDirect®, and RDMA accelerates AI workloads by boosting performance for model training and model tuning with faster checkpointing. Front-end host InfiniBand connectivity to networks, including the NVIDIA Quantum-2–based InfiniBand network platform, provides flexibility. HPE GreenLake for File Storage scales up to 720 PB of effective capacity (with 3:1 data reduction) for large enterprise-scale AI file data.

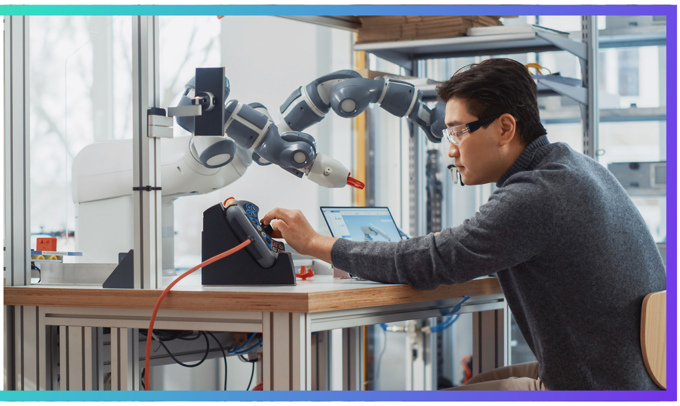
Unlock the value of your data with AI

Imagine supercharging your data-intensive workloads with enterprise performance, simplicity, and enhanced efficiency, all at AI scale, for a competitive advantage with more value from your data, increased productivity, and higher ROI and sustainability. HPE GreenLake for File Storage makes it all possible.

Are you ready to unlock the value of your data?

Spanning all the stages of AI

- Data aggregation
- Data preparation
- Training and tuning
- Inference



[4] "HPE Storage substantiation," HPE, 2024

Learn more at

HPE.com/us/en/HPE-GreenLake-File-Storage.html

Visit **HPE.com**

Chat now (sales)

Hewlett Packard Enterprise

a50010294ENW, Rev. 2