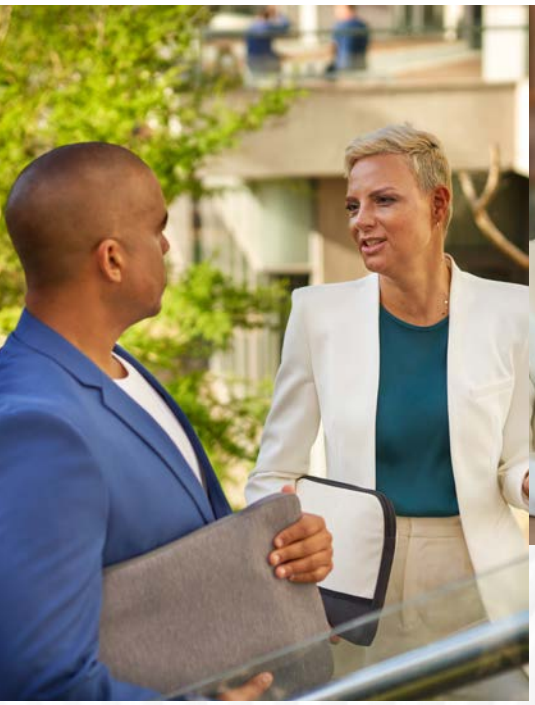
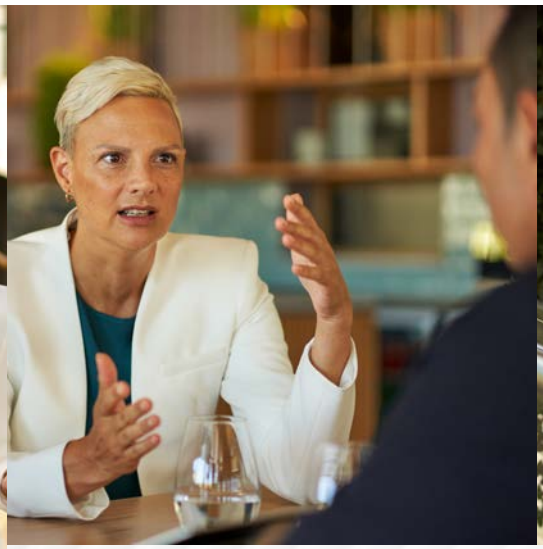
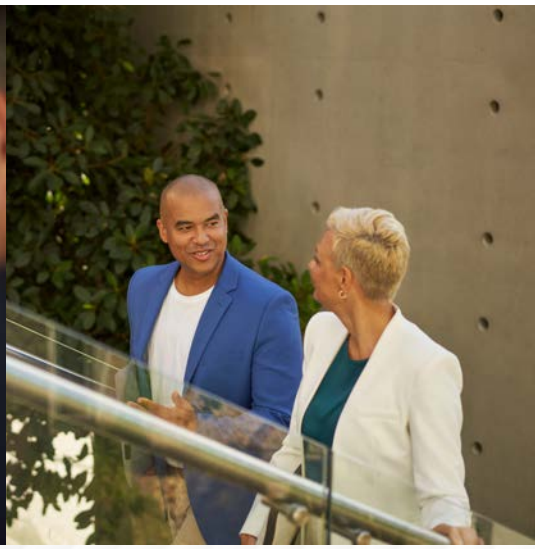
HPE GreenLake for File Storage boosts productivity

File storage with an intuitive cloud experience







Artificial intelligence (AI) and other modern data-intensive applications aim to extract maximum value from data. To deliver insights, advance innovation, and drive breakthrough solutions, these apps process huge amounts of file data at high speed. Processing at this scale requires modern file storage capable of sustained performance — but in practice, even having enterprise performance at AI scale is not enough.

In today's competitive environment, modern file storage must also enhance the productivity and efficiency of the people who run the applications while delivering enhanced efficiencies that yield a compelling ROI. HPE GreenLake for File Storage delivers on both counts.

Your most valuable assets are your people and your data. By people, we mean both users of IT services — your subject-matter experts (SMEs), data scientists, and line-of-business (LOB) application owners — and the IT staff that provides those services. You must optimize the interaction between your data managers/innovators and your data resources, making them highly productive and efficient to achieve your digital transformation goals and seize competitive advantage.

HPE GreenLake

Data stakeholders perceive productivity and efficiency differently

Users of IT services and IT staff can have very different perspectives on productivity and efficiency. Consider these three groups:

1. SMEs and data scientists want to set up and run their jobs easily and drive faster time to insight. They don't want to be hindered by file storage, but sometimes, they are compelled to use complicated file systems to get the performance they need. This group often spends time tuning and doing setup to run their jobs, which inhibits productivity. They don't care who provides the needed file storage, be it the LOB application owner, IT staff, or public cloud, as long as it's fast and simple to use.
2. LOB application owners want to get their apps up and running. They don't care if there are storage silos as long as goals are achieved. Sometimes, LOB application owners will obtain storage resources without going through IT staff or even use the public cloud despite its trade-offs to speed up progress, so your organization risks getting stuck with storage silos from various application areas. From an IT point of view, this is a huge management problem.
3. IT staff members want to keep users happy but also want to minimize storage silos. They want to orchestrate file data simply and efficiently across the board — including storing, managing, protecting, and tracking data — with low overhead. Their worst nightmare is data stored across silos with no one knowing what and where all the data is. This can happen when key data managers and innovators who generate, accumulate, and use data leave the organization without cataloging that data. In this case, IT must figure out what data is stored in the various silos that remain.

Boosting the productivity and efficiency of your teams

One of the key strengths of HPE GreenLake for File Storage is that it enhances productivity and efficiency for your data managers and innovators. Application owners and users get simple setup and fast job completion, with easy file share creation and data sharing across a global namespace for each storage cluster. They no longer need to manage file storage infrastructure as they did previously to get the file storage performance their applications need. And they aren't burdened by cumbersome, legacy file systems that require technical expertise, involve intricate setup processes, and result in storage silos. Consequently, you get more productivity and efficiency from your highly compensated SMEs as they save valuable time and effort and focus on their work priorities.

Meanwhile, the IT team can leverage radically simple, unified file and storage management via a single cloud console and automated, nondisruptive upgrades that keep users happy. With the HPE GreenLake edge-to-cloud platform, IT can help eliminate complex management, storage infrastructure silos, poor performance scaling, and low productivity from edge to cloud — while modernizing data management via a comprehensive suite of over 50 cloud data services and a unified operating experience. Both IT staff and data managers/consumers can leverage a self-service cloud experience, accessible from anywhere, on any device, empowering them to work on strategic initiatives rather than being weighed down by day-to-day operations.

Optimizing your capital investments

In addition to highly productive data teams, you want maximum productivity, efficiency, and ROI from your capital equipment investments. With HPE GreenLake for File Storage, you can flexibly scale performance and capacity independently as your needs grow, making your storage infrastructure future-proof. You can efficiently scale up and out by adding controller and storage nodes as needed. This is better than having to get another box or appliance that sacrifices flexible and efficient scaling while potentially leaving you with underutilized or imbalanced compute and storage resources.

Data-intensive AI models have an insatiable appetite for storage capacity, rack space, and power consumption. But legacy NAS solutions with shared-nothing architectures are unable to efficiently scale out to keep pace with the capacity density, cost per TB, and power efficiency demands of AI workloads. In contrast, HPE GreenLake for File Storage can slash storage costs with high-density compute and storage — as well as lower power usage and carbon footprint with industry-leading data reduction, non-disruptive upgrades, and an AI storage-as-a-service consumption model that helps eliminate overprovisioning.

With support for optimized GPU utilization through InfiniBand, NVIDIA® GPUDirect, and RDMA, HPE GreenLake for File Storage accelerates AI workloads by boosting performance for model training and tuning via faster checkpointing. This can maximize your GPU utilization — and therefore your GPU ROI — with enterprise performance at AI scale.

Let's look at a few more advantages of HPE GreenLake for File Storage:

Investment protection

An architecture designed for exabyte scale ensures you have plenty of headroom, which is critical for large data lakes, as migrating tens of PBs of data every 4–5 years is cost-prohibitive. Built on HPE Alletra Storage MP modular hardware, the architecture of HPE GreenLake for File Storage reduces hardware silos, islands of data, and migration issues, providing an infinite data lifecycle that is transparent to applications. Easily maintain storage infrastructure and make nondisruptive upgrades while benefitting from investment protection and high ROI with data-in-place upgrades.

SSD longevity

With data foresight technology for all-NVMe storage, HPE GreenLake for File Storage delivers efficiencies for flash longevity. The system learns how data is laid down during ingestion into the metadata layer and takes advantage of being able to see the full data estate. The foresight engine intelligently classifies data as it flows through the system and places it in stripes according to life expectancy. Data ages out of stripes simultaneously, helping eliminate write amplification and enabling years of SSD longevity.

Efficient data protection

HPE GreenLake for File Storage features a new erasure code algorithm that rebuilds a failed SSD without reading the entire very large stripe that the system uses. The algorithm aims to maximize capacity efficiency without sacrificing fast rebuild times and includes a Storage Class Memory buffer for large stripes. The new erasure code uses parity drives as a force multiplier in rebuild times, so only a quarter of the drives must be read to rebuild a stripe. This results in a very low overhead of around 2.7%.[1] Along with high efficiency in how data is laid down, you also benefit from full system performance. It's the best of both worlds.

Superior data reduction

HPE GreenLake for File Storage further increases ROI with highly efficient data reduction that uses the unique Similarity algorithm. Similarity is important when you're looking for data reduction in file data. Whereas many all-flash block systems rely on compression and deduplication and work well in environments for virtual machines (where you have many exact copies of operating systems or multiple copies of the same databases for QA and testing), in the unstructured world of files, there is a lot of data that will be similar but not identical.

Compression is fine-grained but local: you are reducing redundancy over a small piece of data. It's also very computationally intensive. Deduplication is global over a large amount of data but very coarse. Data is broken up into chunks of the same block size, and the system looks for exact matches. Similarity looks for similar but not identical data on both a global and fine-grained basis.

Take the example of DNA data. A DNA strand will have mostly the same information, but there will be differences at various points of uniqueness. If an algorithm looks for identical matches, there would be a mismatch. However, similarity will spot blocks of data that are mostly the same, compress them together, track the changes between them, and store only what is common. This gives you the best of both worlds from compression, which is fine-grained but local, and deduplication, which is global but coarse.

Figure 1 illustrates Similarity's superior data reduction. Here is a before and after comparison of a data footprint using Similarity vs. compression and deduplication:

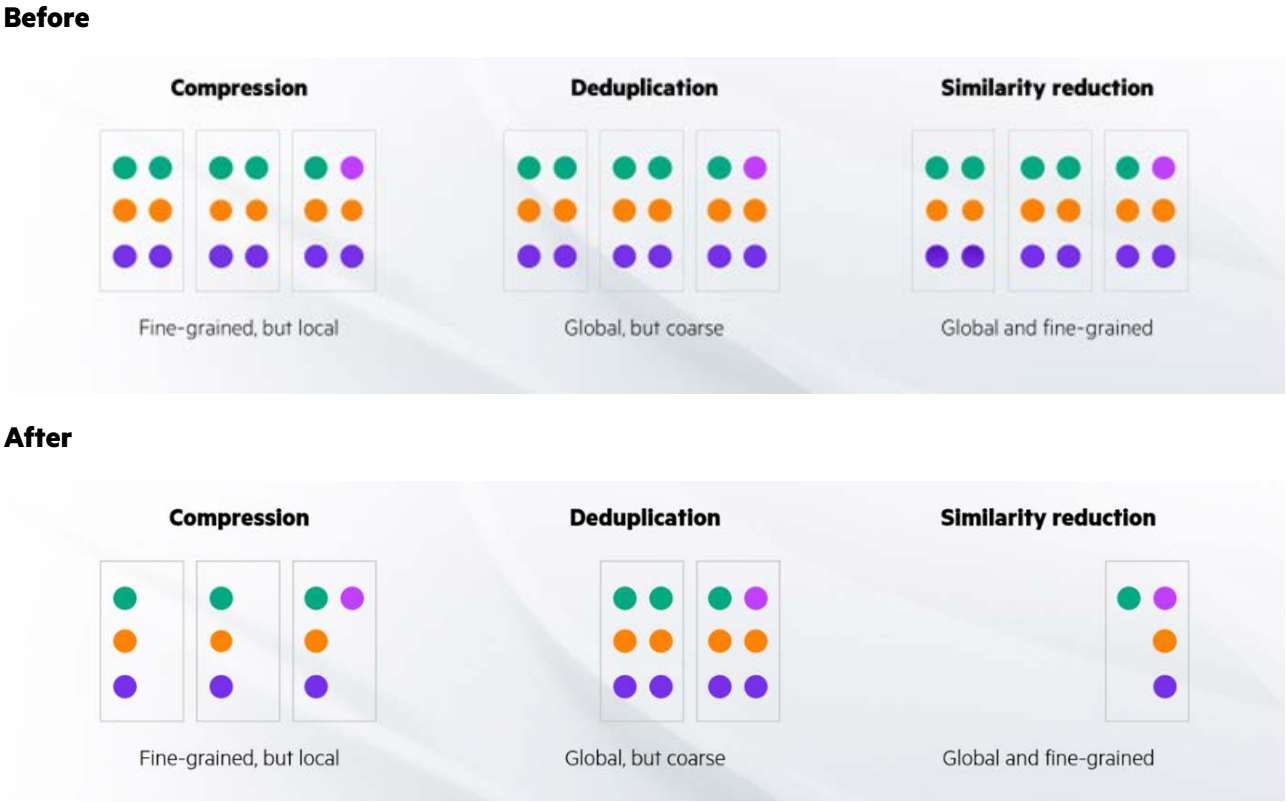


Figure 1. Similarity is game-changing storage efficiency

[1] VAST Data white paper content on Locally-Decodable Error Correction Codes

For example, savings from the Similarity algorithm is up to 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.[2]

With investment protection, SSD longevity, efficient data protection, and superior data reduction, HPE GreenLake for File Storage delivers impressive efficiency, productivity, and high ROI from your capital investments.

Enhanced productivity and efficiency from human and capital resources

When considering file storage IT infrastructure, you must secure performance at AI scale for today's data-intensive workloads. And think about two additional factors: First, how will your solution enhance the productivity and efficiency of your valuable human resources across the board, from LOB application owners and SMEs to IT staff? For that, you need efficient file storage infrastructure and simplified management. Second, you must ensure that your capital investments will deliver a high ROI. In summary, you need enterprise performance, simplicity, and enhanced efficiency, all at AI scale.

As you evaluate and decide on your file storage strategy and investments, it's important to have a comprehensive view that not only accounts for performance but also the effective utilization of both your human and infrastructure resources. HPE GreenLake for File Storage is designed to deliver enhanced productivity and efficiency on both fronts.

Learn more at

HPE.com/us/en/HPE-GreenLake-file-storage.html

[2] HPE storage substantiation

Explore **HPE GreenLake**

Chat now (sales)

Hewlett Packard Enterprise

a50009539ENW, Rev. 2