HPE GreenLake

Solution overview

HPE Solutions With WEKA

Data storage for artificial intelligence in medical imaging

The Tsunami of Medical Imaging Data in Medical Research and Clinical Settings

Medical imaging institutions collect and store millions of digital images each year and the volume of those images is growing. IDC estimates that on average, approximately 270 GB of healthcare and life science data was created for every person in the world in 2020.[1] The majority of that data is image data. These images contain valuable insights, which are generally hidden or obscure and as a result, are frequently left unutilized.

There are thousands of high-resolution medical imaging studies that include computed tomography, magnetic resonance imaging, mammography, which will eventually also include digital pathology. All of these are overwhelming diagnostic clinicians' ability to keep ahead of the data onslaught and case volume. Automation is needed to assist clinicians in analyzing complex images, identifying hidden lesions, reducing false positives, and recognizing patterns of change over time. Legacy NAS systems can't keep up with these new demanding imaging volumes.

Detecting lesions and abnormalities in this rising sea of higher resolution medical image data is becoming increasingly tedious and challenging for medical professionals to keep pace. Medical imaging is the ideal environment to deploy the transformational power of artificial intelligence (AI) to increase clinician efficiencies and improve clinical outcomes.

HPE Solutions with WEKA in Health and Life Sciences

HPE Solutions with Weka provide an answer to the medical imaging challenges, serving as a digital assistant to streamline and automate clinicians' diagnostic workflow. HPE Solutions with Weka provide the large throughput needed to keep NVIDIA DGX A100 servers fully fed with training data, helping maximize their capabilities. The platform's **no-copy architecture** reduces data travel and copying times between resources in the workflow, offering resilient, nonstop operation while future-proofing the platform with unlimited performance, scalability, and capacity to support the ever-expanding medical imaging resolutions and procedures.

Deep learning has revolutionized the ability to augment the diagnostic imaging process, providing automation assistance to the clinician where needed. By training a model using thousands of images, HPE Solutions with Weka deliver innovation that can instantly detect and evaluate abnormalities in a medical image, speeding the diagnostic process.

Why do Legacy Systems Fall Short?

Legacy NAS storage systems can't keep up with these new demanding performance requirements, and traditional parallel file systems are generally too complex for most commercial enterprises. The required copying of data from one resource to another in the typical medical imaging workflow only results in wasted time. HPE Solutions with Weka delivers the primary design objective with deep learning model training to deliver compute, networking, and storage resources that are capable of constantly saturating the NVIDIA DGX A100 GPUs doing the training processing. This is done by providing high throughput at low latency through the Weka file system, where the image data for learning is stored. The more data a deep learning model can learn from, the faster it can identify abnormalities and the better its accuracy.

[1] IDC Blog—The Data Dilemma and its Impact on AI in Healthcare and Life Sciences, June 2021.

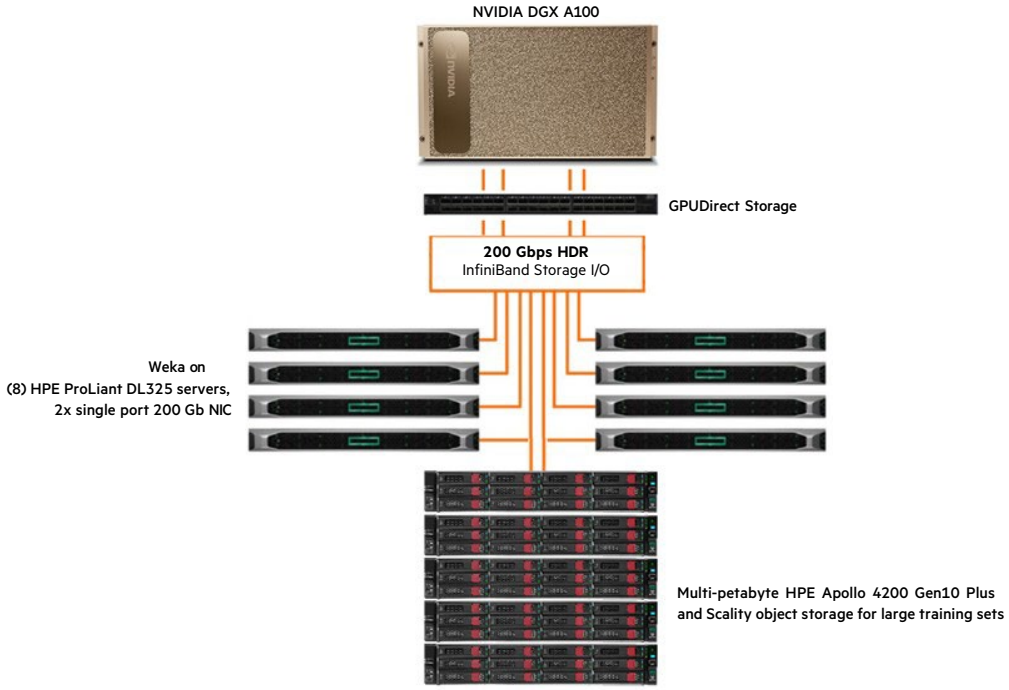

FIGURE 1. HPE Solutions with Weka—AI innovation in medical imaging

WEKA on HPE Servers is at the Core of an Accelerated Medical Imaging Deep Learning Data Flow

HPE Solutions with Weka are specifically designed to address the rise of medical imaging data rates and processing requirements. This platform is uniquely able to accelerate medical imaging deep learning, increasing training accuracy and clinical outcomes. The design philosophy behind the platform's Weka file system is to create a single storage architecture that runs on-premises or in the public cloud with the performance of all-flash arrays, with simplicity and feature set of NAS and the scalability and economics of the cloud. Weka is a software-only, high-performance, file-based storage solution that is highly scalable and easy to deploy, configure, manage, and expand.

Incorporating HPE Solutions with Weka into the medical imaging deep learning workflow saturates data transfer rates to NVIDIA DGX A100 systems and helps eliminate wasteful data copying and transfer times between storage silos to geometrically increase the number of medical images that can be analyzed per day.

HPE Solutions with Weka enables AI innovation in medical imaging, delivering extremely high I/O performance (IOPS) and a scalable AI data platform that delivers the resilient and consistent performance needed to keep up with the data output of medical imaging deep learning systems. HPE Solutions with Weka offers unlimited linear performance scalability headroom to support the ever-increasing medical imaging resolutions in the future.

Reduce Data Travel and Copying Times with the WEKA No-Copy Architecture

Legacy medical imaging processing and storage architectures are constantly moving data from one processing system to another, slowing results with expanding data transfer times. This results in increasing infrastructure costs tied to increasing storage capacity, complexity, and management. HPE Solutions with Weka provides a **no-copy** architecture, which reduces data copying and travel time, shortening inspection processing duration. With the no-copy architecture, data is simply written once and transparently accessed by resources in medical imaging deep learning data flow.

Weka participates in the HPE Complete program, which provides:

- **Completeness**—Delivers and manages best-in-class third-party products that augment, enhance, and complete HPE solutions.
- **Confidence**—Validates third-party product compatibility and interoperability with HPE products and IT ecosystems. Develops purpose-built systems along with product integration, configuration, and sizing guides. Integrates third-party software interfaces with HPE IT management software such as HPE InfoSight and HPE OneView.
- **Convenience**—Simplifies acquisition of third-party solutions by making them orderable on one HPE purchase order with options for delivery within HPE GreenLake edge-to-cloud platform solutions.

Seamless Namespace Expansion into Object Storage

This integrated system from Hewlett Packard Enterprise and Weka can also integrate object storage, providing for economically high capacity and easy access to storage, protecting large numbers of medical images for the duration of the deep learning training process. HPE Solutions with Weka includes the ability to seamlessly expand its namespace to and from object storage, using Scality RING or ARTESCA, with the data residing in a single Weka namespace where metadata stays on flash tier for fast, easy access and management. Large files are reduced to small objects and tiny files are packed into larger objects to help maximize parallel performance access and space efficiency.

GPUDirect Storage

HPE Solutions with Weka are architected to deliver low latency and higher bandwidth to speed I/O to the GPU, helping ensure it's saturated with data while also supporting NVIDIA®'s GPUDirect Storage protocol that bypasses the CPU and memory. This enables GPUs to communicate directly with Weka storage, accelerating throughput to the fastest possible performance.

An Architecture Designed for the Lowest Latency Possible

Medical AI workflows consist of intense random reads across image datasets where the low latency can accelerate training and inference performance. Weka's small 4K block size matches that of NVMe SSD media block sizing for optimum performance and efficiency. Weka evenly distributes metadata processing and direct data access across storage servers (with no back-end network), further lowering latency and increasing performance. What's more, Weka designed low-latency, performance-enhanced networking: Weka does not use standard TCP/IP services but a purpose-built infrastructure that uses the Data Plane Development Kit (DPDK) to accelerate packet processing workloads without any context switches and zero-copy access. Weka bypasses the standard network kernel stack helping eliminate the consumption of kernel resources for networking operations. Weka implements a custom protocol over User Datagram Protocol (UDP) for even lower latency. This performance is available on standard Ethernet and InfiniBand networks. The DPDK consists of libraries to accelerate packet processing workloads. For customers implementing stateful containers, Weka's CSI plug-in has been validated with HPE Ezmeral Runtime Enterprise and is available in the HPE Ezmeral Marketplace.

Learn More At

hpe.com/storage/hpe-solutions-for-weka

hpe.com/us/en/solutions/healthcare.html

Make the right purchase decision. Contact our presales specialists.

Chat now (sales)

Call now

Visit GreenLake.HPE.com

Get updates

Hewlett Packard Enterprise

a50005578ENW, Rev.1