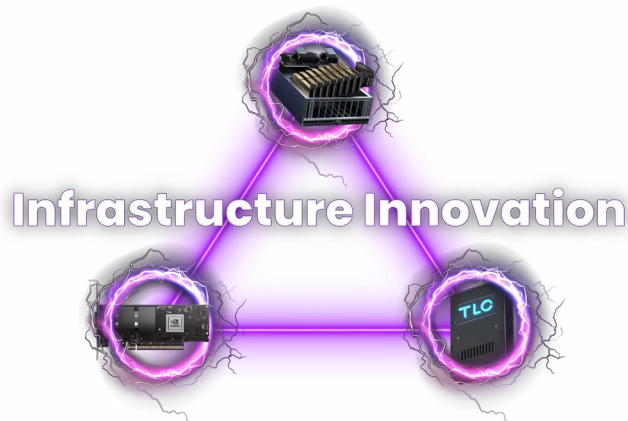


WEKA Data Platform for the Cloud

Cut your cloud storage costs in half and drive 10x faster AI pipelines

The explosion in AI applications has kicked off a new wave of workloads built in the cloud. This new wave is focused on the most performance-intensive applications like LLM model training, high-performance compute, and a host of industry applications such as VFX rendering and post-production; genomics processing and cryogenic electron microscopy (cryo-em); fraud detection and high-frequency trading; electronic design automation; and many more. These next-generation workloads expect the agility and scalability of the cloud while also achieving levels of performance and control that were historically only possible in home-grown data centers.



Cloud infrastructure offerings have rapidly transformed to enable this shift, driven by massive increases in network performance, compute capacity, and the proliferation of GPUs for accelerated computing. However, data infrastructure hasn't kept up - either in the cloud or on-premises. These performance-intensive workloads continue to rely on the same unstructured file systems and legacy data architectures that have been around since the 1990s. Making matters worse, these workloads are often deployed as part of a pipeline, where data is ingested, prepared, analyzed, results captured, and

stored or archived for later use. Every step in the process has a different performance profile across high throughput, IO, handling lots of small files, sometimes heavy on reads and look-ups, other times heavy on writes. Legacy storage systems are optimized for one stage in the pipeline, not all. To mitigate the situation, organizations either over-provision storage resources at a significant cost or simply accept a slow data pipeline.

So, while technology leaders are turning to the cloud as their preferred option for AI and many other performance-intensive workloads, it's not yet a certainty. A recent study from [Enterprise Strategy Group](#) revealed that "performance requirements could not be met" is now the #1 reason holding cloud migrations back (for the first time ahead of "implementing security measures"), with "too costly to migrate" tied for the #2 spot! Most technology leaders are realizing that legacy approaches to data are holding back some of their most strategic projects. In the [2023 Global Trend in AI Report from S&P Global](#), leaders cited "data management" (32% of respondents) as the biggest inhibitor to AI success.

The WEKA Data Platform

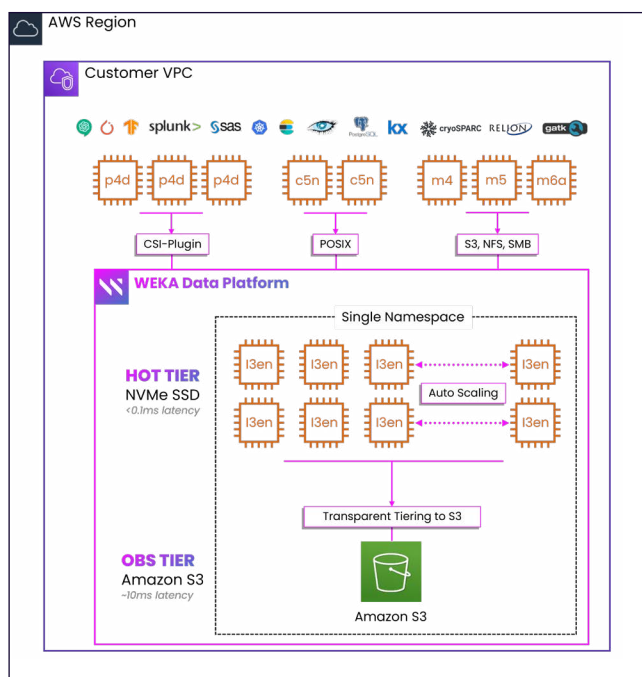
A transformational data platform for AI and HPC workloads in the cloud. The WEKA Data Platform delivers the performance and scalability needed to affordably run performance-intensive applications in the cloud. WEKA provides a software solution purpose-built for high-performance and AI workloads running in the cloud. Used by leaders across every industry vertical, the solution is helping organizations accelerate AI model training, tuning, and inference while reducing costs. WEKA software eliminates the need to over-provision resources in the cloud to meet demanding performance requirements and removes the need to manage multiple copies of data across a pipeline, simplifying and accelerating data pipelines while dramatically reducing costs. Customers who use WEKA for AI, HPC, and other performance-intensive workloads in the cloud reduce infrastructure costs by 50%, accelerate data-intensive workloads by 65%, and accelerate data pipelines by 90%, leading to increased revenue by as much as 20%.

The WEKA Data Platform in the Cloud

The WEKA Data Platform combines high-performance flash and low-cost object storage in a single namespace. WEKA software automatically manages the tiering of data between flash and object storage so that every workload gets the performance it needs without the need to over-provision storage resources. WEKA auto-scaling enables customers to dynamically add resources to meet either performance or capacity requirements and then scale back down when those resources are no longer needed. In this way, customers never pay for capacity or performance they don't need. WEKA software automatically handles tuning of the storage system to deliver the performance profile for every workload - high IO, high throughput, handling lots of small files, and any mixture of reads and writes. This approach eliminates the need to maintain multiple copies of data across a data pipeline.

Simplify Cloud Deployments and Reduce Costs in the Cloud

Cloud Native Deployments



The WEKA Data Platform runs in the customers' VPC across all four major cloud providers - AWS, Azure, Google Cloud, or Oracle Cloud. WEKA software resides on a cluster of dense storage VMs in the cloud (such as Amazon EC2 i3en instances), utilizing the local NVMe storage available on each instance to form a high-performance storage layer. The single WEKA namespace extends to an object store bucket (such as Amazon S3 or Azure Blob Storage) to add massive capacity at an affordable cost. The entire data set is available to the applications without the need to move or copy data. Applications residing either in traditional instances or containerized in Kubernetes deployments (Amazon EKS or Microsoft AKS) can access data in the WEKA namespace through multiple storage protocols - S3, NFS, SMB, POSIX, and CSI.

WEKA stores hundreds of petabytes of data in cloud object storage while retaining hot data needed for data-hungry applications in NVMe storage. It leverages proprietary [tiering technology](#) to efficiently and optimally place data in NVMe storage to drive high performance when applications need it.

WEKA auto-scaling enables customers to scale up either the performance or capacity of the WEKA system independently. To scale performance up, WEKA software leverages compute autoscaling to add new instances to the cluster as needed to deliver near-linear increases in performance. Autoscaling also enables customers to scale the cluster back down (or completely away) when projects are not running.

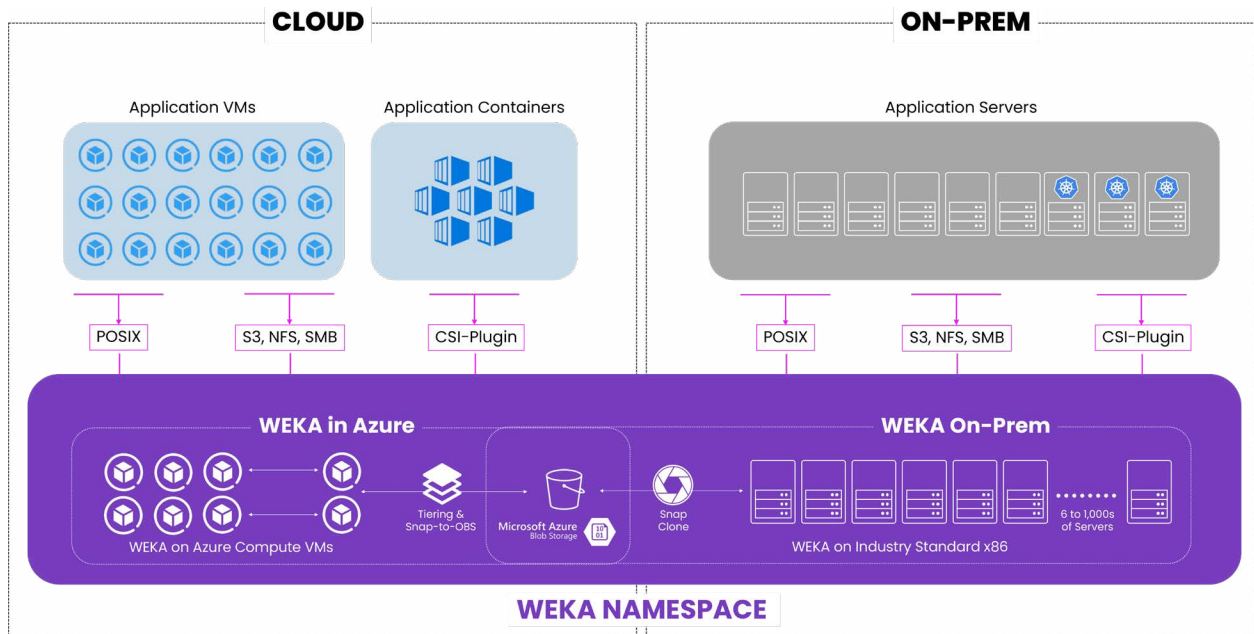
Hybrid and Multi Cloud Deployments

The same WEKA Data Platform software that is available for the cloud also runs in the customer’s data centers and any cloud, including Microsoft Azure, Google Cloud, and Oracle Cloud. WEKA Snap to Object enables simplified data operations by eliminating the need to create, move, and maintain multiple copies of data across data silos located on-premises and in the cloud. Snap to Object creates a self-describing snapshot of the entire WEKA environment - including data and

This ensures customers don’t have to over-provision storage resources to meet a performance requirement, and they don’t pay for resources they don’t use. Scaling the single WEKA namespace capacity through cloud object stores like Azure Blob is simple and scales to hundreds of Petabytes or even Exabytes of capacity.

WEKA deployments are automated using AWS CloudFormation or Hashicorp Terraform scripting. CloudFormation templates create a private subnet with the customer’s VPC, set up Auto Scaling groups, register and start up a predefined set of instances in the cluster, create the requisite S3 bucket, and deploy WEKA software to the cluster.

file system metadata and replicates that snapshot to another object store either in a customer’s data center or in any cloud. The WEKA Zero Copy Architecture ensures customers need only maintain a single copy of data to deliver their performance requirements for every stage in their data pipeline. With WEKA, customers can maintain a single copy of data, maintain consistency across any location, and ensure high-performance, low-latency access to data from anywhere in the world.



Customers use WEKA for various hybrid cloud use cases, including cloud bursting, cloud archive and backup, disaster recovery, and migrations. Bursting to the cloud has become common in the life sciences and financial services industries, where data governance is important, but resources in data centers are limited. For this use case, WEKA customers typically maintain their primary environment in their data center. As needed, they use Snap-to-Object to create a full copy of their WEKA environment in their cloud object store. Using WEKA deployment automation and autoscaling, customers can spin up a new WEKA environment using on-demand cloud instances and run analytics to support their application. Once complete, the customer can archive the analysis results to their data center and scale the entire WEKA environment in the cloud back down. In this way, no resources are left running in the cloud, and all data is retained in the customer's data center.

To accelerate and simplify disaster recovery operations, customers can use WEKA in a similar fashion to create a fully usable copy of their data center environment in any cloud and spin up the needed cloud compute instances and WEKA environment to run disaster

At the Edge

Bursting to the cloud is a rapidly growing use case for edge deployments that rely on capturing a lot of data quickly and then processing or analyzing that data to drive business results. A wide variety of industries see edge data analytics as strategic to their business. For example, scientists in a lab capture thousands of image files from electron microscopes and send that data to cloud-based resources to analyze protein structures or DNA molecules. In a hospital, medical technicians capture large image files from MRI, CT, X-ray, and diagnostic tools and increasingly rely on cloud resources to analyze images. Artists in a creative agency access and edit video files and render at 120 frames per second using Autodesk Flame, all entirely in the cloud.

recovery tests and actual recovery operations as needed. Once the test is complete, or the data center event is resolved, the entire WEKA environment in the cloud can shift back on-premises and scale down.

Many customer migrations from on-premises to the cloud actually happen as an offshoot of their disaster recovery test operations. In this case, the customer uses WEKA to shift data to the cloud of their choice and simply does not shift it back. The WEKA environment can start at a modest size to support planned, incremental migrations, and scale as additional applications move to the cloud.

Customers can also rely on the cloud-to-cloud data portability enabled through WEKA Snap to Object to migrate their WEKA environment running in one cloud and move it to the object store in another cloud. This provides customers with a level of flexibility and risk management that is not possible with other approaches. Customers running in AWS can use Snap to Object to move a fully usable snapshot of their data to Microsoft Azure, Google Cloud or Oracle Cloud - or vice versa. Best of all, cloud migration operations using WEKA are fast and easy.

WEKA enables these scenarios by enabling rapid data capture at the edge - at the doctor's office, in the lab, or the artist's studio. Once captured, WEKA Snap to Object manages the data movement into the cloud of their choice for analysis. The WEKA environment can scale up to process data analytics and support the data pipeline. Once complete, analysis results can be captured and copied back to the customer's data center, and the entire WEKA deployment in the cloud can scale back down.

Maximum Flexibility, Minimum Hassle

In the cloud, at the edge, or in hybrid deployments, WEKA helps make your cloud journey seamless and painless. WEKA makes transitioning to cloud-based workflows easier by accelerating performance-intensive applications, reducing costs, and simplifying data operations. According to a recent study conducted by the Enterprise Strategy Group, WEKA customers:

- Accelerate performance-sensitive and data-intensive applications by up to 10x
- Complete data pipeline analysis for HPC and AI in the cloud by 90%
- Drive faster project velocity leading to a 1% to 20% surge in revenue
- Streamline data operations leading to a 64% reduction in staff hours dedicated to data pipeline management orchestration
- Reduce cloud costs by an average of 38% annually
- Reduce environmental impact with 260 tons of CO2 saved per PB

Next Steps/CTA

Learn more about how WEKA for the Cloud at www.weka.io/cloud



[weka.io](https://www.weka.io)

844.392.0665

