

WEKA for Amazon SageMaker HyperPod

Reduce epoch times from weeks to hours and optimize infrastructure utilization

Accelerate Distributed Model Training and Tuning

WEKA provides a high-performance data platform for SageMaker HyperPod optimized for every phase of FM model training across data loading, pre-processing, model training, checkpointing, verification, tuning, and data set archiving.

Increase HyperPod Cluster Utilization

High-performance storage from WEKA eliminates data bottlenecks driving up GPU infrastructure utilization above 90% (from an avg 30%) and ensuring your HyperPod infrastructure is never starved for training data.

Increase HyperPod Cluster Resilience

WEKA reduces FM model checkpoint times by 90%, enabling faster training times and increasing the resilience of HyperPod deployments.

Increase Developer Productivity

WEKA software accelerates model data load times by 50%, accelerating model training times and improving data scientist and developer productivity.

WEKA for Amazon SageMaker HyperPod

Deploy a high-performance data platform for AI model development that scales massively, drives GPU infrastructure utilization, and reduces costs.

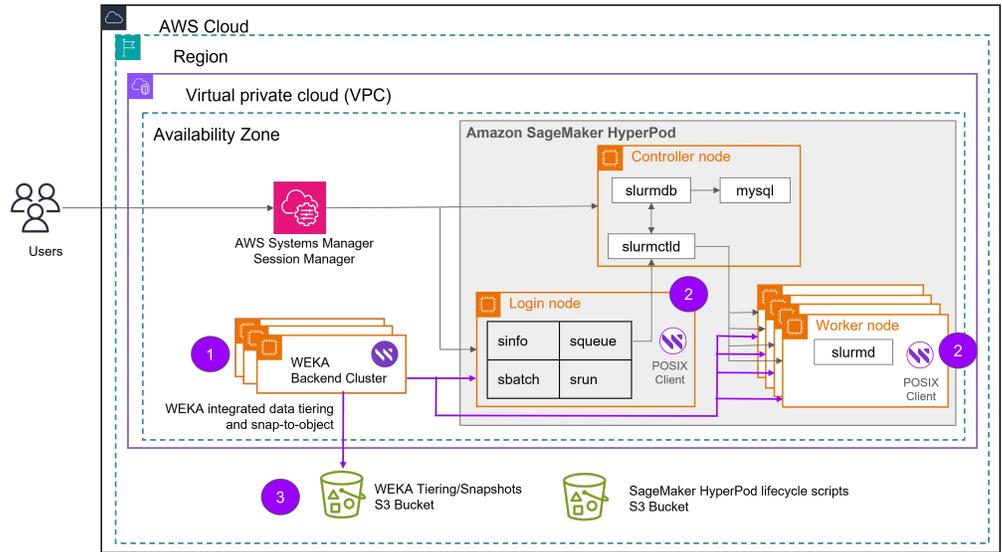
- ✓ Accelerate AI model training by 10x
- ✓ Increase GPU Utilization by 90%
- ✓ Accelerate model load times by 50%
- ✓ Reduce model checkpoint times by 90%

WEKA For Amazon SageMaker HyperPod Sample Architecture

1. WEKA cluster deployed on i3en EC2 instances using WEKA Terraform modules for AWS to serve as a single high-performance storage

2. WEKA POSIX Client software deployed on the SageMaker HyperPod uses Linux Kernel bypass enabling high throughput and extremely low latency

3. WEKA seamlessly tiers data between NVMe SSD volumes on i3en EC2 instances and Amazon S3 object storage to optimize storage efficiency



WEKA For AWS Specs

Virtual Machines	Amazon EC2 i3en	No. of files or directories	Up to 6.4 trillion files
Cluster Size (VMs)	6 to Hundreds	No. of files in a single directory	Up to 6.4 billion
Object Storage	Amazon S3	Capacity with object store	Up to 14 Exabytes
Protocols	POSIX, NFS, SMB, S3, CSI	No. of File Systems	Up to 1,024 per Cluster
Automation	Hashicorp Terraform AWS CloudFormation AWS Marketplace	Snapshots	Up to 24,576 Readable Snapshots (14,336 IF ALL WRITABLE)

Get started with [WEKA in the AWS Marketplace](#) today!



weka.io

844.392.0665

