WEKA for Research

Accelerate insights and ground-breaking research with The WEKA® Data Platform.

CHALLENGES

- Reduce research costs
- Accelerate research insights
- Optimize shared compute
- Scale up and down
- Manage mixed IO
- Improve time to market

WHY WEKA?

- Accelerate discovery with dynamic performance and capacity scaling
- Increase collaboration with integrated cloud connectivity
- Protect critical data without sacrificing performance
- Speed time to build models
- Simplify the environment including backup & DR
- Lower the cost to process and store data
- Scale to meet the demands of modern pipelines
- No need to copy datasets between locations

| **$9B**	**90%**	**UP TO 7X**
expected value of AI in life sciences by 2032	the amount researchers can accelerate project times with the WEKA Data Platform	improved application performance with WEKA

Today's discovery requires conducting more research in less time with more data. Researchers need a flexible infrastructure to affordably handle trillions of files and directories on-premises and in the cloud without slowing down their research. The WEKA Data Platform lets you tame unruly data management and be ready to take on any scientific application, at any time, and in any location.

WEKA eliminates the need to over-provision storage resources to meet performance-intensive research application needs. The bi-directional scaling enables you to rapidly scale up your research workflows to meet your timelines and then scale back down so you never pay for resources you don't use.

Keeping a GPU compute cluster fed with data is critical to overall pipeline performance and efficiency. WEKA powers GPUs up to 20x faster, which means your GPUs and modeling software spend less time idle waiting for data. Our zero-tuning capability also assures that every application in your analytics pipeline gets the performance it needs.

WEKA's zero-copy architecture eliminates the need for multiple data copies and complex data operations, reducing time to insights for your researchers and scientists. Our single namespace for flash and object storage with smart tiering ensures hot data is always ready for analysis, improving application performance by as much as 7x. And because we know flexibility is key, WEKA is designed to be deployed on-premises, supports hybrid and cloud deployments, and integrates seamlessly with AWS.
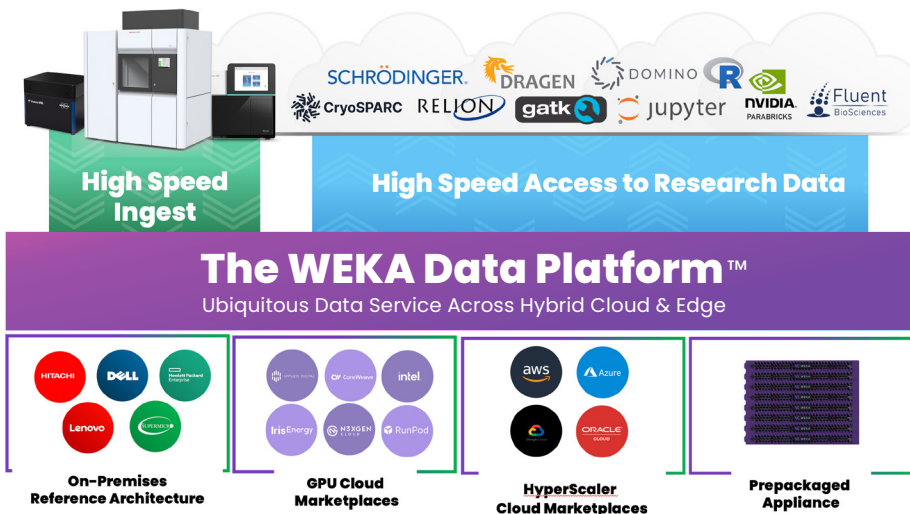
20x lower management overhead

UP TO 75% lower infrastructure costs

UP TO 30% less in wall-clock time

Problem

Needed to streamline data processes in the cloud, achieve performance at scale, and adjust resources based on demand to ensure that researchers could meet publications deadlines.

Solution

Combined a wide range of storage mechanisms and cloud to provide an extendable file system, economics of cloud object stores, exceptional scale, and the speed of SSD and dedicated networking.

Problem

Needed to improve AI epoch times and implement a less costly solutionm that eliminated copying to multiple EBS instances.

Solution

Migrated from native services to reduce epoch time from 3 months to 1 week with 100%+ savings.

Problem

Needed a highly-performant solution that allowed for scaling a shared data set that had scaled beyond NAS capability but OBJ was too slow.

Solution

Scaled capacity by 5x, increased performance by 10x and reduced costs by 75% vs Isilon.

92% reduced epoch time in AWS

40x faster creating and copying 1GB files

30M files being used to train the model

140PB of data in storage

72x faster genome pipeline compared to CPU

75% less storage cost per genome

10x your genomic workflows

WEKA genomics customers have experienced compute workloads being **reduced from 70 days to as low as 7 days.**

"HPC tends to turn CPU problems into I/O problems. WEKA turned those I/O-bound problems back into CPU-bound problems."

"With WEKA, we can expand input/output capabilities to ensure that even the most data-intensive research projects are supported efficiently."

GAVIN BURRIS
SENIOR IT LEADER AT THE WHARTON SCHOOL

"Our scientists can now more easily complete their data analysis without dealing with data infrastructure. Our old workflow was complex and time-consuming, involving staging data to local SSD compute nodes for better application performance. With WEKA, we have so much performance and expandable capacity that we don't have to think about it."

STUART GLENN
ASSISTANT DIRECTOR OF INFRASTRUCTURE & RESEARCH COMPUTING, OMRF

CUSTOMERS WHO TRUST WEKA WITH THEIR RESEARCH