The WEKA® Data Platform: The High-Performance Solution Redefining HPC and AI Workloads

The WEKA® Data Platform is setting a new benchmark in high-performance data solutions, delivering record-breaking speed, unmatched efficiency, and optimal resource utilization for HPC, AI, and other data-driven workloads. With top industry benchmarks and powerful real-world results, WEKA showcases its unique ability to handle the most demanding applications with high levels of efficiency.

As HPC, AI and machine learning reach new levels, the need for a data platform that can keep up is essential. WEKA rises to this challenge by eliminating traditional data bottlenecks, maximizing GPU utilization, and accelerating every stage of the data pipeline. This not only reduces costs but also enables faster insights. Dive in to discover how WEKA empowers organizations to unlock their infrastructure's full potential, transforming data into a decisive competitive advantage. For those looking to boost performance, streamline total costs, and prepare for the future, WEKA is the purpose-built solution that delivers.

Exceptional Performance

WEKA has consistently proven its high-performance capabilities, leading in various industry benchmarks, such as SPECStorage™ and MLPerf®. These achievements underscore its prowess in managing extensive workloads with speed and efficiency. For instance, WEKA demonstrated its performance dominance by achieving a 6x higher load count than competing benchmarks driving greater efficiency. By leveraging a flexible architecture that performs equally well across hyperscale clouds and on-premises environments, WEKA allows customers to confidently deploy applications in the optimal infrastructure without performance trade-offs.

Performance Efficiency and Cost-Effectiveness

In addition to raw performance, WEKA excels in performance efficiency. By minimizing latency and maximizing throughput, WEKA achieves significant efficiency savings for users. For example, in one benchmark WEKA achieved 175% of the competitor's job count at only 64% of the infrastructure cost, resulting in a 2.7x lower cost per job. This efficiency provides customers with the dual benefits of speed and lower expenditure, which is especially advantageous for hyperscale cloud deployments where budgetary constraints are common.

Moreover, WEKA's zero-copy architecture, which simplifies data access paths, has been shown to reduce time-intensive processes drastically. This approach also optimizes resource allocation, allowing users to gain the highest throughput with minimal hardware investment. This enables a leaner, more efficient use of IT resources and translates into a smaller data center footprint while maintaining top-tier performance.

Enhanced GPU Utilization for HPC Workloads

In HPC and AI workloads, maximizing GPU utilization is crucial, as these resources are among the most expensive in a data center. WEKA's Data Platform optimizes GPU utilization by consistently delivering high-performance data to GPUs, preventing data bottlenecks that would otherwise lead to idle GPUs and wasted resources. For example, Stability AI, a WEKA client, saw GPU utilization jump from 30% to 93% after moving their data pipeline to WEKA. This efficiency is especially critical for deep learning workloads, where keeping GPUs running near full capacity can dramatically shorten training times and reduce costs.

WEKA's performance in the MLPerf benchmarks further demonstrates its ability to handle massive data workloads while keeping GPU utilization high. The platform achieved the #1 spot in A100 tests for both 3D-Unet and ResNet50, underscoring its capacity to maximize throughput and accelerate training times without any need for tuning changes between tests. This consistency is particularly valuable for organizations with extensive AI pipelines that require handling both large and small file IO profiles efficiently.

Superior Scalability and Data Management for Diverse Workloads

WEKA's Data Platform was designed from the ground up to handle the mixed IO patterns commonly found in modern data-intensive applications. It efficiently manages diverse workloads, including read, write, metadata-intensive, IOP-driven, and throughput-driven tasks, all without requiring configuration changes between workloads. By eliminating bottlenecks that typically arise in data pipelines, WEKA allows for substantial reductions in time-to-completion for data processing tasks, which is essential in industries where time is critical, such as drug discovery and AI-driven analytics.

An example of WEKA's impact on real-world data pipelines comes from Atomwise, a pharmaceutical company using AI for drug discovery. Atomwise reduced its training epoch times from 80 hours to 4 hours—a 20x improvement—by switching to WEKA's Data Platform. This enabled them to run complex AI model experiments in a fraction of the time, accelerating their time-to-market and enhancing innovation potential in life sciences.

Robust Support for Modern Data Ecosystems

WEKA's architecture is purpose-built for high-throughput, low-latency access to data. Its scalability across both cloud and on-premises environments, combined with optimized GPU utilization, positions WEKA as an ideal solution for organizations that require a dependable, high-performance data platform. WEKA's support for modern data ecosystems enables it to drive substantial improvements in productivity and resource efficiency, essential for meeting the challenges of the AI and GenAI boom.

Furthermore, WEKA's ability to saturate multiple accelerators with minimal hardware resources is a game-changer for resource-constrained organizations. This not only improves performance and increases efficiency, but also reduces the hardware footprint, making WEKA a sustainable and cost-effective option for long-term growth in HPC and AI applications.

KEY INSIGHTS

WEKA vs GPFS for HPC and AI Workloads

When it comes to AI and high-performance computing (HPC), the choice of a data platform can significantly impact performance, scalability, and ease of management. While IBM Spectrum Scale (formerly GPFS) has a long-standing presence in the data storage market, in the realm of AI and next-generation workloads, WEKA surpasses IBM Spectrum Scale by providing an AI-native data platform that excels in GPU utilization, ease of management, and flexible deployment options. For organizations looking to optimize their infrastructure for the future of AI, WEKA offers a purpose-built solution that combines performance and simplicity, giving it a decisive edge over legacy approaches like IBM Spectrum Scale.

Designed for Modern AI Needs

GPFS, developed in the early 1990s, was originally designed for multimedia and scientific applications. As such, it often requires high levels of expertise to configure and maintain, which can be a drawback in dynamic environments where rapid deployment and scalability are essential. In contrast, WEKA was purpose-built to handle the demanding IO profiles of HPC, AI, and machine learning, such as random reads and mixed IO workloads, with minimal configuration. WEKA's user-friendly architecture enables seamless scalability across cloud, on-premises, or hybrid environments, allowing organizations to optimize resources for AI without the steep learning curve and configuration complexity often associated with Spectrum Scale.

Greater Ease of Use and Cost Efficiency

WEKA's platform offers a simplified, easy-to-manage solution compared to Spectrum Scale. IBM's Spectrum Scale often requires specialized hardware and expertise to achieve high performance, and its management can be labor-intensive, especially in multi-node configurations. WEKA, however, provides a software-defined, hardware-agnostic solution that is deployable on standard, commodity hardware, in the cloud, or as a hybrid solution. This approach reduces both infrastructure costs and operational complexity, allowing organizations to scale efficiently and focus more on innovation rather than infrastructure management.

Multi-Cloud and Hybrid Capabilities

As more HPC workloads transition to cloud and hybrid environments, WEKA's platform excels by enabling seamless data mobility across different infrastructures, including AWS, Azure, and Google Cloud. While IBM Spectrum Scale has cloud-compatible options, it lacks the agility and flexibility of WEKA, particularly in handling diverse workload requirements across different environments. WEKA's native cloud integration and support for hybrid deployments empower organizations to dynamically scale workloads wherever needed without sacrificing performance.
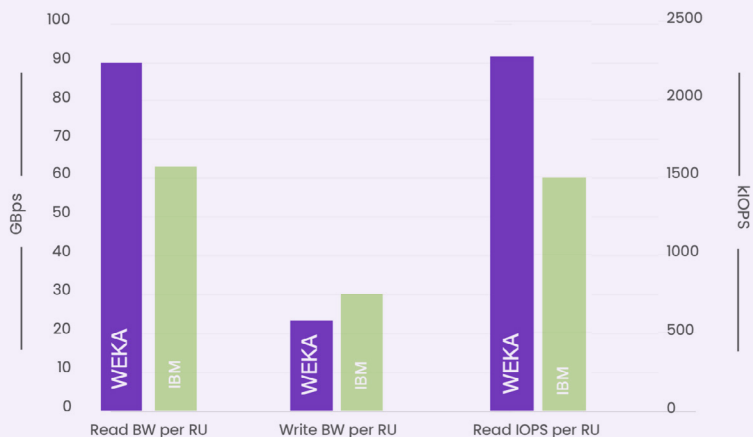
	WEKA	IBM
Model (Rack Units)	WEKApod ™ (1U)	ESS 3500 (2U)
Read BW (per RU)	90 GBps	63 GBps
Write BW (per RU)	23.3 GBps	30 GBps
Read IOPS (per RU)	2,280 kIOPS	1,500 kIOPS
Write IOPS (per RU)	535 kIOPS	no data

Solution Capacity Scale

WEKA Data Platform: 10 TB – 14 EB
IBM Spectrum Scale: 48 TB – 8 EB
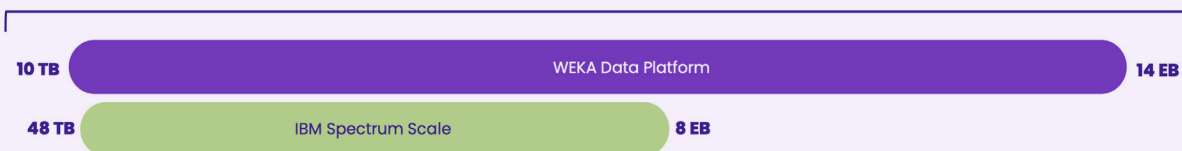
	WEKA	IBM SPECTRUM SCALE
Commodity Hardware	Yes, Industry standard COTS	No
SW Only	Yes	Yes
Tiering to Object	Yes, Global Namespace	Yes
Tiering/Hybrid Disk Storage	Yes , Tier to object storage	Yes
Ease of Use	Easy	Hard
PROTOCOLS		
NFS	Yes	Yes
LDAP	Via additional gateway	
SMB	Yes	Yes
GPUDirect	Yes, SMB2/3.1 including Multichannel and SMB Direct	Yes, SMB2/3 or via additional gateway
POSIX	Yes	Yes
S3	Yes	Limited or via additional gateway

	WEKA	IBM SPECTRUM SCALE
SECURITY & RECOVERY		
Data Encryption	Yes, at-rest & in-flight	Yes, at-rest & in-flight
Data Protection	N+2, N+4	N+3, Reed Solomon only ESS Appliance
Data Efficiency	Yes	Limited
Snapshots	Yes	Yes, With perf impact
Max Snapshots per file system	24,000	256
Snapshot to S3	Yes	No
CLOUD		
Cloud Supported	AWS, Google, Oracle, Azure	AWS, Azure, IBM Private Cloud, Oracle
Tier to Cloud	Yes	Yes
Backup to Cloud	Yes	Yes
Burst to Cloud	Yes	No

Conclusion

WEKA stands out for delivering high performance, exceptional efficiency, and robust support for AI and data-intensive workloads, distinguishing itself even against legacy solutions like IBM Spectrum Scale. Its achievements in benchmark tests underscore its capability to meet the demands of modern applications, while maximizing the value of costly GPUs and infrastructure investments. With WEKA's platform, organizations can effectively accelerate their data pipelines, reduce costs, and gain faster time-to-insight, positioning them for success in a competitive, data-driven landscape.

Compared to IBM Spectrum Scale, WEKA offers unmatched performance and simplicity, empowering organizations to maximize the ROI of their HPC and AI investments. By eliminating infrastructure limitations, WEKA enables data teams to focus on innovation, solidifying its position as the preferred choice for companies seeking to lead in HPC, AI, and machine learning.

WEKA   weka.io | 844.392.0665 |