Solution Brief

The AI-Native Data Platform with VAST and Cloudera

Unified Data Infrastructure Optimized for Analytics and AI

VAST | CLOUDERA

From Analytics to AI: More Value, Less Silos

VAST Data and Cloudera help enterprises move beyond legacy and fragmented data architectures to AI-ready infrastructure that scales from analytics to training, inference, and agentic AI.

By combining Cloudera's industry-leading data services, AI Workbench, and AI Inference Service with the VAST AI Operating System, organizations can eliminate unnecessary data movement, establish a true single source of truth, and maximize the return on AI infrastructure investments.

The result is a modern AI Factory architecture where data is continuously ingested, governed, refined, and delivered to analytics and AI services without the latency, duplication, and operational complexity of traditional data pipelines.

Unified Data Infrastructure: Store less. Process more

Traditional enterprise architectures separate files, objects, data lakes, warehouses, and AI infrastructure across multiple storage systems and data platforms. As organizations adopt generative and agentic AI, these silos create governance challenges, increase infrastructure costs, and force teams to maintain complex pipelines simply to move data between systems.

The VAST AI Operating System changes this model by providing a scalable, high-performance data foundation that unifies files, objects, structured tables, vectors, and event streams on a single all-flash platform optimized for analytics and AI workloads.

Cloudera analytics, streaming, machine learning, and AI services operate directly against this shared data foundation, allowing organizations to store less and process more while leveraging higher quality and fresh enterprise data.

The Modern Data Foundation for AI

Unify files, objects, tables, vectors, and events on a single all-flash data platform.

AI Performance

Deliver the throughput and low latency required for large-scale training, inference, and real-time analytics.

GPU Efficiency

Reduce storage-induced stalls and maximize utilization across training and inference infrastructure.

Simplicity

Eliminate data copies, reduce ETL complexity, and consolidate infrastructure silos.

Economics

Reduce storage overhead, lower TCO, and improve infrastructure efficiency across analytics, training, and inference workloads.

Why Run Cloudera Data Services with VAST?

Eliminate Data Movement

Replace fragmented architectures and ETL pipelines with a shared data foundation where analytics and AI services operate directly on the same data.

Accelerate AI Infrastructure

VAST's Disaggregated Shared-Everything (DASE) architecture and integration with NVIDIA GPUDirect Storage reduce latency between data and compute, helping maximize accelerator utilization and improve AI infrastructure efficiency.

Scale Inference Without Contention

Every VAST compute node (CNode) can directly access the full VAST data set, eliminating storage hotspots and enabling large-scale inference and KV cache workloads without east-west traffic.

Power Real-Time AI Pipelines

Combine enterprise AI services with a unified data foundation to support streaming ingestion, vector retrieval, RAG, inference, and agentic AI workflows at scale.

Build the Enterprise AI Factory

Together, VAST and Cloudera provide a practical path from legacy, fragmented data architectures to modern, production-scale AI systems.

Organizations gain a unified platform for analytics, training, inference, and agentic AI while improving governance, reducing infrastructure complexity, and maximizing the value of both data and AI investments.

From data ingest to model deployment, VAST Data and Cloudera deliver the foundation for the next generation of enterprise intelligence.

Key Capabilities



Unified Data Infrastructure
Support structured, unstructured, and vector data on a single platform with native support for files, objects, tables, and events.



AI-Optimized Storage Architecture
VAST's DASE architecture provides scalable throughput and low-latency access for modern AI workloads.



Enterprise AI Services
Deploy Cloudera AI Workbench, AI Inference Service, analytics, and streaming services consistently across data centers, private clouds, and public clouds.



Open and Flexible Deployment
Preserve existing investments with open architectures, hybrid deployment models, and support for existing analytics and AI ecosystems.

About VAST Data

VAST Data is the AI Operating System company – powering the next generation of intelligent systems with a unified software infrastructure stack that was purpose-built to unlock the full potential of AI. The VAST AI OS consolidates foundational data and compute services and agentic execution into one scalable platform, enabling organizations to deploy and facilitate communication between AI agents, reason over real-time data, and automate complex workflows at global scale. Built on VAST's breakthrough DASE architecture – the world's first true parallel distributed system architecture that eliminates tradeoffs between performance, scale, simplicity, and resilience – VAST has transformed its modern infrastructure into a global fabric for reasoning AI. Learn more at vastdata.com and follow VAST Data on LinkedIn, YouTube and X.

About Cloudera

Cloudera is the only hybrid data and AI platform company that large organizations trust to bring AI to their data anywhere it lives. Unlike other providers, Cloudera delivers a consistent cloud experience that converges public clouds, on-prem data centers, and the edge, leveraging a proven open-source foundation. As the pioneer in big data, Cloudera empowers businesses to apply AI and assert control over 100% of their data, in all forms, improving security, governance, and real-time and predictive insights. The world's largest brands across all industries rely on Cloudera to transform decision-making and ultimately boost bottom lines, safeguard against threats, and save lives.

To learn more, visit Cloudera.com and follow us on LinkedIn and X.

© 2026 Cloudera and associated marks are trademarks or registered trademarks of Cloudera, Inc. All other company and product names may be trademarks of their respective owners.

©2026 VAST Data, Inc. All rights reserved. All trademarks belong to their respective owners.

7/26 | 12100