

# Cloudian Delivers Integrated AI Inferencing and Data Storage Solution to Simplify AI Infrastructure

*HyperStore now features integrated vector database, streamlining enterprise AI inferencing workflows and supporting the AI Data Platform vision*

SAN MATEO, CA, UNITED STATES, July 8, 2025 /EINPresswire.com/ -- [Cloudian](#), the leader in enterprise object storage solutions, today announced a breakthrough integration that combines high-performance data storage with AI inferencing capabilities in a single, unified platform that

simplifies the infrastructure required to support AI workflows. Cloudian HyperStore now features integrated vector database support on the industry's highest-performing object storage platform, delivering a streamlined solution for enterprise AI inferencing workflows while advancing towards a comprehensive AI Data Platform (AIDP) architecture.



“

The integration of data storage and AI inferencing into a single, efficient platform represents a fundamental shift in how enterprises approach AI infrastructure.”

*Neil Stobart, CTO at Cloudian*

This integration addresses the critical challenge enterprises face as AI inferencing workloads scale to production environments. Modern AI applications require massive storage capacity for vector datasets that can reach petabytes in size, along with supporting index files and operational logs, while simultaneously demanding ultra-low latency access for real-time inferencing operations. By unifying data storage and AI inferencing infrastructure, organizations can eliminate data movement bottlenecks and reduce the complexity of deploying enterprise-scale AI

solutions.

"The integration of data storage and AI inferencing into a single, efficient platform represents a fundamental shift in how enterprises approach AI infrastructure," said Neil Stobart, CTO at

Cloudian. "This solution directly supports the AI Data Platform vision by providing the foundational storage and inferencing capabilities that modern AI workloads demand, while dramatically simplifying deployment and improving performance."

The solution leverages [Milvus](#), the world's most advanced open-source vector database, to power similarity search and AI inferencing applications including recommendation systems, computer vision, natural language processing, and retrieval-augmented generation (RAG). Milvus efficiently stores, indexes, and queries high-dimensional vector embeddings generated by machine learning models, enabling millisecond-level query response times for billion-scale vector datasets.

Key Benefits of Cloudian's Integrated AI Inferencing Solution:

- Unified Architecture: Eliminates the complexity of managing separate storage and inferencing systems, reducing operational overhead and accelerating time-to-production for AI initiatives.
- Performance Leadership: Industry-leading object storage read performance of 35GB/s per node enables faster AI model inference and improved application responsiveness.
- Enterprise Scalability: Proven exabyte-scale object storage seamlessly supports massive vector datasets while maintaining high-performance access for real-time inferencing workloads.
- Cost Efficiency: Integrated solution reduces total cost of ownership compared to deploying separate storage and inferencing platforms, with simplified management and reduced data movement costs.

This advancement also represents a significant step toward realizing the AI Data Platform vision, which envisions unified, accelerated infrastructure that seamlessly integrates data processing, storage, and AI computation. By providing both the storage foundation and inferencing capabilities in a single platform, Cloudian enables enterprises to build comprehensive AI infrastructure that can scale from pilot projects to production workloads.

The integrated AI inferencing and storage solution supports both on-premises and hybrid cloud deployments, giving organizations maximum flexibility in their AI infrastructure strategy. Development teams can leverage existing S3-compatible tools and workflows while benefiting from performance optimizations specifically designed for AI inferencing operations.

Cloudian's integrated AI inferencing solution is available immediately for evaluation.

## About Cloudian

Cloudian provides an exabyte-scalable data platform optimized for AI workloads, data protection, and advanced analytics. The company's HyperStore solution delivers native S3 API compatibility, government-verified security, and industry-leading performance, enabling organizations to efficiently manage massive datasets across distributed environments. Through

partnerships with leading technology providers like NVIDIA, Cloudian delivers the ideal foundation for data-intensive applications that drive business innovation and competitive advantage. Learn more at [www.cloudian.com](http://www.cloudian.com).

Jon Toor  
Cloudian  
[email us here](#)

---

This press release can be viewed online at: <https://www.einpresswire.com/article/829095491>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2025 Newsmatics Inc. All Right Reserved.