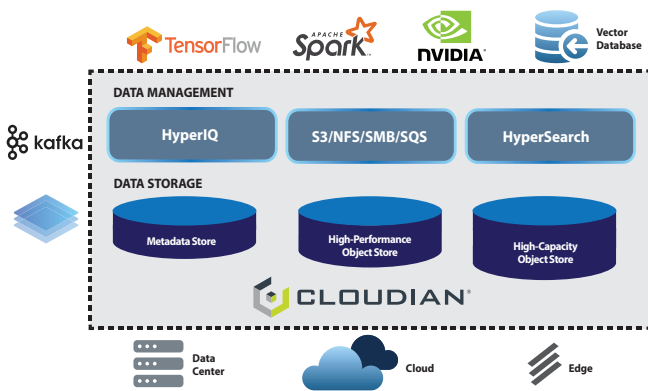


Empowering AI Systems with Scalable and Resilient Data Storage

Artificial intelligence (AI) and machine learning (ML) are revolutionizing industries, driving the next wave of business automation. These technologies heavily rely on vast amounts of data, putting significant demands on IT systems, especially storage infrastructure. Cloudian HyperStore offers the ideal solution for AI systems, providing a scalable, affordable, and fault-tolerant object storage platform that meets the unique requirements of AI workloads.

Addressing Unstructured Data Challenges

AI, ML, and analytics environments require large volumes of unstructured data for training and processing. Cloudian HyperStore excels in managing unstructured data, offering an exabyte-scalable and highly concurrent object storage system. It serves as an optimal foundation for building both stream and batch AI pipelines, enabling organizations to efficiently leverage unstructured data for valuable insights.



- Exabyte Scalable
- Secure / IAM
- Fault Tolerant
- Rich Metadata / Tags
- Hybrid Cloud
- Data Lakehouse
- Multi-Tenancy
- Geo-Distributed

Solution Benefits

- Exabyte Scalability
- Multi-Tenancy
- Cost Efficiency
- Hybrid Cloud
- Data Durability
- Compatible with Popular ML Frameworks

Use Cases

- Address Unstructured Data Challenges
- Optimize AI Workflows
- Streamline Data Management
- Protect Valuable Data
- Enhance Performance and Compatibility

Enabling Seamless Data Pipelines

Cloudian HyperStore serves as a central storage system for various components of AI systems. It acts as a feature store, facilitating the registration and discovery of curated features. For vector databases like Milvus, Cloudian HyperStore provides scalable data infrastructure. Moreover, it supports the storage of machine learning models, ensuring model reproducibility, governance, and secure serving. With Cloudian HyperStore, organizations can streamline their data pipelines by leveraging a single storage system for different AI components.



An AI pipeline uses massive amounts of data for model training and running predictive analytics. Cloudian HyperStore provides an exabyte scalable, highly concurrent, affordable and fault tolerant object storage system which can be used as optimal foundation for building stream and batch AI pipelines.

Protecting Valuable Data

Raw data is a valuable asset that cannot be easily replicated. Cloudian HyperStore ensures data protection through features such as data replication and erasure coding. It provides military-grade data security, object versioning, fault-tolerance, and data immutability, ensuring the integrity and durability of critical AI data assets.

Streamlining Data Management

Effective data management is crucial for successful AI applications. Cloudian HyperStore offers rich object metadata, object versioning, and object tags, allowing organizations to organize and catalog their data efficiently. HyperStore's multi-tenancy feature enables multiple data scientists or teams to work simultaneously with the same data, fostering collaboration and accelerating AI workflows. Additionally, Cloudian HyperSearch simplifies data discovery, providing an easy-to-use interface for finding the needed data quickly.

Enhancing Performance and Compatibility

Cloudian HyperStore optimizes AI workloads by supporting popular machine learning frameworks like TensorFlow, PyTorch, and Spark ML. These frameworks are specifically designed for parallel training from object storage, providing improved performance and compatibility. Organizations can harness the power of GPUs without storage limitations, maximizing the utilization of expensive and high-demand resources.

NVIDIA GPU and Triton Support

NVIDIA's powerful inference server, NVIDIA Triton™, supports S3-compatible storage, Cloudian HyperStore's primary storage type. This means that organizations can harness the capabilities of NVIDIA GPUs for accelerated and scalable AI model deployment and execution, while also benefiting from the robust and scalable storage infrastructure provided by Cloudian HyperStore. By combining the power of NVIDIA AI Enterprise, including Triton, with the storage capabilities of Cloudian HyperStore, organizations can streamline the data science pipeline, expedite the development and deployment of AI models, and achieve efficient and reliable production AI.

Support for Multiple Types of Stores

Cloudian HyperStore may be deployed as a data platform for multiple types of stores, including feature stores, vector databases, and model stores, providing organizations with the necessary infrastructure to effectively manage curated features, conduct large-scale similarity searches, and securely store machine learning models within their AI systems.

Different Types of Stores



Feature Store

- A repository for curated features
- Features to be registered and discovered
- Used both as part of model training, as well as inference pipelines

HyperStore supports rich object metadata. Store features in the metadata layer for fast online access. Use HyperStore with external tables for offline feature storage and to manage and discover features.



Vector Database

- Enables large-scale similarity search
- Semantic search
- Cross Modality Data
- Data embedded into a common vector space

HyperStore can provide data infrastructure for Milvus, an open-source vector database.



Model Store

- A repository for ML models
- Stores model files, artifacts, and metadata
- Guarantees the reproducibility of models
- Provides governance and security of models

HyperStore can be used to store, search, and discover ML models. Object versioning, IAM, Object Lock, and tagging features to store, share, and protect models.

Conclusion

Choosing the right storage platform is paramount for the success of AI systems. Cloudian HyperStore emerges as the ideal solution, providing a scalable, affordable, and resilient data storage foundation. With its comprehensive data management capabilities, robust data protection features, and seamless integration with AI workflows, Cloudian HyperStore empowers organizations to store massive amounts of data, automate machine learning workflows, and unlock the full potential of their AI initiatives.



Cloudian turns information into insight with an infinitely scalable platform that consolidates, manages and protects enterprise data. Cloudian enterprise file and object storage brings cloud technology and economics to the data center with uncompromising data durability, intuitive management tools, and the industry's most compatible S3 API. Cloudian and its ecosystem partners help Global 1000 customers simplify operations in enterprise use cases including data management, analytics, storage services, and AI/ML.

More at cloudian.com/ai-workflows