

Data Lakehouse as a Service for Sovereign VMware Cloud

VMware Cloud Provider-Ready Greenplum and Cloudian

Challenge

Modern analytics applications in the cloud have evolved by combining the flexibility, cost-efficiency, and scale of S3 data lakes with the data management and ACID transactions of data warehouses to create “Data Lakehouses.” However, organizations that need to comply with sovereignty laws where their data resides have not been able to take advantage of these paradigm-shifting analytics applications in the public cloud.

It is no surprise, therefore, that such organizations are increasingly looking at cloud service providers and managed service providers (xSPs) to help them modernize high value use cases like analytics while maintaining sovereignty, security, and data protection measures. xSPs running a sovereign cloud are also well positioned to help enterprises navigate regulations and ensure compliance with data privacy legislation.

The opportunity to offer a sovereign Data Lakehouse as a Service that is secure, affordable, and scalable has never been greater. Essential to this is an S3 Data Lake along with an analytics tool built on open standards, for xSPs, with in-built multi-tenancy, billing, metering, and the ability to expand to petabyte scale on-demand.

Solution

VMware Tanzu Greenplum, a massively parallel processing (MPP) data warehouse platform, seamlessly integrates with Cloudian HyperStore S3-compatible object storage to provide enterprise customers the same Data Lakehouse architectures on-premises. This VMware-certified solution enables new efficiencies and savings and is ideal for the creation and deployment of advanced analytics models for complex enterprise applications.

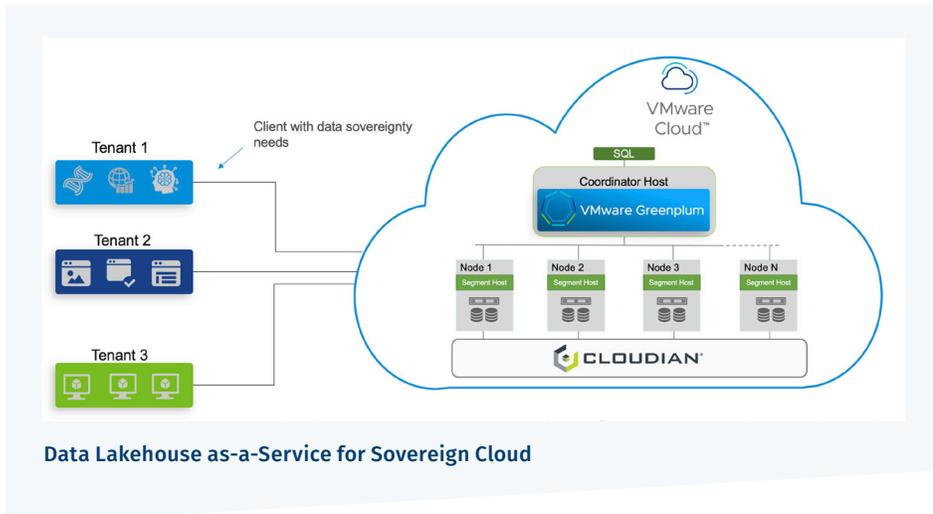
Greenplum and Cloudian HyperStore are both available as part of the VMware Cloud Provider Program. Cloudian for VMware Cloud Director is an integrated storage platform that lets xSPs and their tenant users deploy, manage, and consume S3-compatible storage within their services environment. For cloud providers, this opens the door to new service revenue streams in use cases such as data protection, big data, healthcare, media archives, video surveillance, and more. For their tenant customers, it delivers a best-in-class object storage service fully integrated and delivered through their preferred cloud management platform.

Solution Benefits

- Provide your customers the capability to perform data analytics in place
- Build a highly profitable Data Lakehouse as a Service offering to compete with public cloud-based offers
- Built on enterprise-grade object storage software with proven VMware Tanzu Greenplum platform
- Compatible with VMware Sovereign framework
- Part of the VMware Cloud Provider Program (purchase with points)
- Military-grade security and compliance certifications
- Integrated data immutability with S3 Object Lock
- xSP ready with multi-tenancy, metering, billing, QOS in-built



VMware Tanzu Greenplum



Solution Advantages

Sovereign Data Lakehouse as-a-Service

VMware Tanzu Greenplum integrates seamlessly with Cloudian HyperStore, the industry’s leading on-prem S3 storage system, with native support for the S3 API. Only Cloudian offers a 100% native S3 API for objects. Together, Greenplum and Cloudian enable xSPs to build and offer robust Data Lakehouse solutions, allowing customers to modernize their data analytics architecture, offer high data durability, and more easily comply with local data sovereignty laws by storing data in scalable storage in-region at up to 70% lower total cost than public cloud.

VMware Cloud Provider Program-Enabled

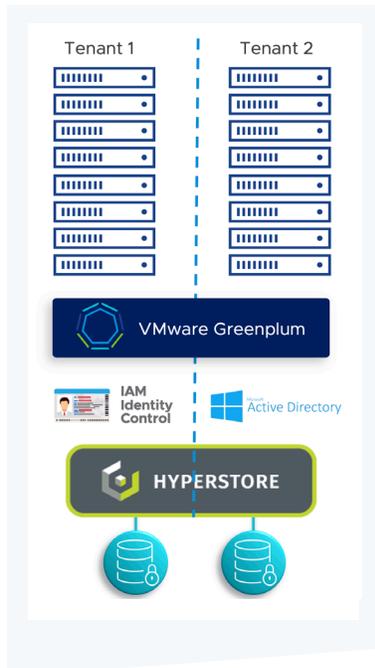
Greenplum and Cloudian are both part of VMware’s Cloud Provider Program, enabling xSPs to build and offer Data Lakehouse-as-a-Service to enterprise customers in a true consumption model.

Exabyte-Scalable to Handle the Largest Greenplum Environments

The Greenplum-Cloudian Data Lakehouse separates compute and storage scalability, enabling xSPs to scale either as required and create a profitable business offering. Cloudian HyperStore, a scale-out object storage system, provides a fast, cost-effective storage platform ideal for data lakes of semi-structured or unstructured data. Available as an appliance or as software, HyperStore lets customers start small and expand seamlessly from TBs to PBs without interruption.

Multi-Tenancy

The solution allows multiple users to analyze data sitting on a single shared data lake infrastructure without compromising security. HyperStore gives role-based access to system and group administrators and to end users. Users can select and provision storage services on-demand from a service catalog. Billing, metering, and QOS are built into the HyperStore platform to enable a full multi-tenant deployment.



Secure Sovereign Platform

The solution provides extensive data security features starting with secure shell, integrated firewall and RBAC/IAM for access controls that allow xSPs to implement a truly secure sovereign cloud service. Cloudian’s S3 Object Lock feature employs write once, read many (WORM) technology to provide immutable storage. Using the industry standard API, S3 Object Lock protects data for the retention period specified by the customer. It is also certified to meet the non-rewriteable, non-erasable storage requirements of SEC Rule 17a-4(f) and is certified to meet the principles-based requirements of CFTC Rule 1.31(c)-(d) and the requirements of FINRA Rule 4511. In addition, the solution supports AES-256 server-side encryption for data at rest and SSL for data in transit (HTTPS). Fine-grained storage policies, including encryption at object and bucket levels, permit security settings to be individually configured for different users or data types in a shared-storage environment.

Profitable Services Offering

Running on standard x86 hardware, xSPs can build the solution on-prem at up to 70% lower cost than in public cloud, leading to highly profitable service offerings.



Award-Winning

Proven at over 700 enterprise and MSP customers worldwide with nearly 2 exabytes of capacity under management, Cloudian Scored Highest for All Use Cases in the Gartner 2020 Critical Capabilities for Object Storage Report and was named a Gartner Peer Insights Customers' Choice in 2020, 2021, and 2022.