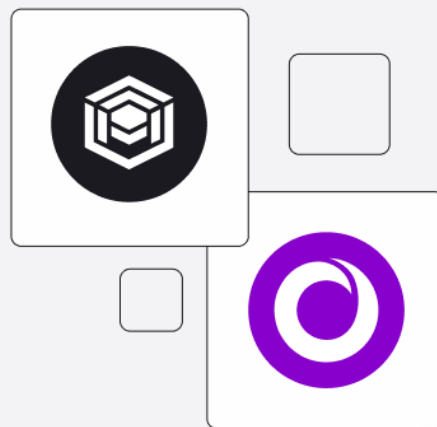


Technical Comparison Report

AlloyDB vs. SingleStoreDB

Abhinav Srivastava, Director, Product Marketing



Introduction

AlloyDB™ is a managed database service that is a fork from PostgreSQL and aims to provide high performance, scalability and availability for transactional (OLTP) and analytical (OLAP) workloads with a multi-node architecture as its foundation. While AlloyDB and SingleStoreDB™ share some similarities with respect to HTAP and multimodel capabilities, they differ in platform maturity, scalability, deployment choices and compatibility with open-source databases.

Similarities

1. **HTAP.** Both AlloyDB and SingleStoreDB are SQL-based, HTAP databases with columnstores for analytical queries.
2. **Multi-model.** Both platforms support a wide range of data types, including JSON and geospatial data.

Key differences

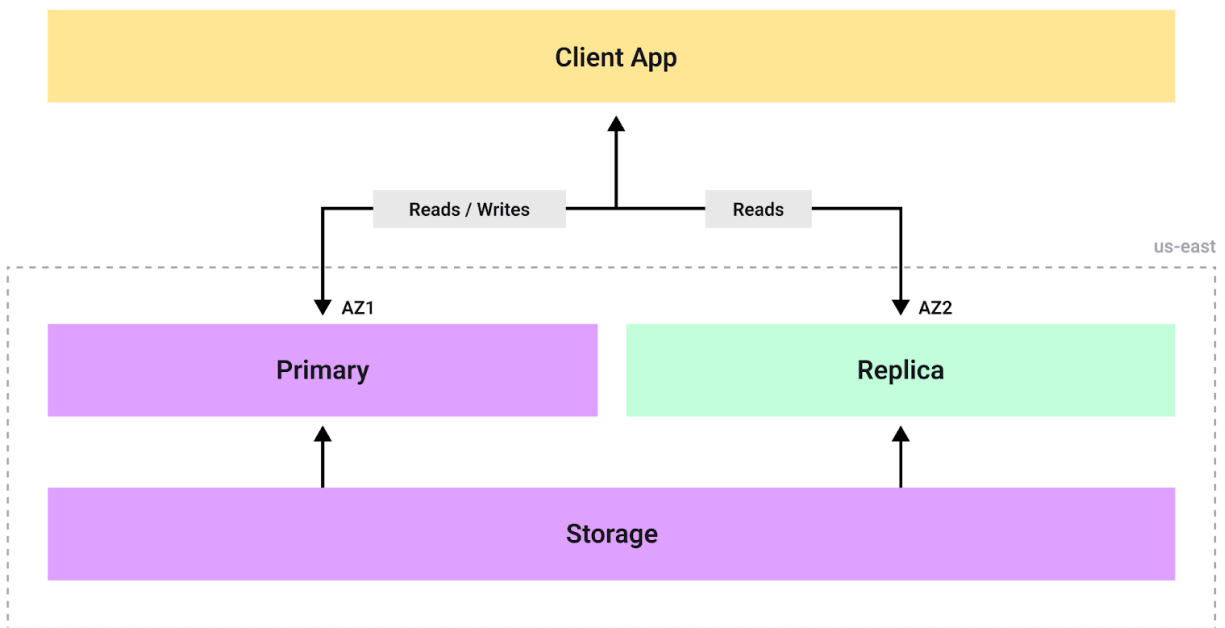
1. **Product maturity.** AlloyDB launched in 2022, so it is a relatively new and untested platform. In particular, its unique functionality built beyond open-source PostgreSQL is questionable — the separation of compute and storage, horizontal scalability, columnstore, etc. In contrast, SingleStoreDB has been built over a decade and deployed in production in F100, as well as fast-growing startups.
2. **Costs.** AlloyDB's reliance on an in-memory columnstore for analytical performance means scaling OLAP workloads can be very expensive. SingleStoreDB's columnstore is on-disk.
3. **Scalability.** While it can scale storage independently, AlloyDB is essentially a single-node database. Adding replica nodes to AlloyDB helps with scaling RO throughput, but latency cannot be improved as there can be only a single primary node servicing a query. SingleStoreDB is a truly distributed database with a shared-nothing architecture that decouples storage and compute, delivering query latency in the milliseconds.
4. **Deployment choices.** AlloyDB is Generally Available (GA) only on Google Cloud Platform (GCP; AlloyDB Omni is currently in preview). SingleStoreDB is available as a managed service on all three cloud hyperscalers, or on-premises/ self-managed on public cloud infrastructure to suit customer needs.

5. **Streaming analytics/ ingestion performance.** SingleStoreDB’s ingestion feature called Pipelines enables fast, high-throughput, parallelized data ingestion – with performance of up to millions of rows ingested per second. Pipelines can extract, transform (modify) and load data without the need for third-party tools or middleware. SingleStoreDB will also soon support CDC out, meaning data can be flowed out into a data lake or lakehouse.
6. **Compatibility/ecosystem.** AlloyDB is based on **PostgreSQL**, while SingleStoreDB is wire-compatible with **MySQL** and **MongoDB**. While both PostgreSQL and MySQL offer large ecosystems of tools and extensions, SingleStoreDB’s compatibility with **MongoDB** could be valuable for customers looking for fast analytics on applications built on MongoDB.

Architecture comparison

AlloyDB architecture

- **Scaling.** To overcome the scaling and analytical limitations of PostgreSQL, AlloyDB separates the database layer from the storage layer.
- **Analytics/HTAP:** AlloyDB uses caching to achieve performance improvements. While the internals of AlloyDB architecture are not well publicized, it can be inferred that AlloyDB’s rowstore is disk-based and columnstore is in-memory.
- **Multi-AZ, not multi-region:** AlloyDB primary and replica instances can only be deployed within the same region. [Source: Google [documentation](#)]

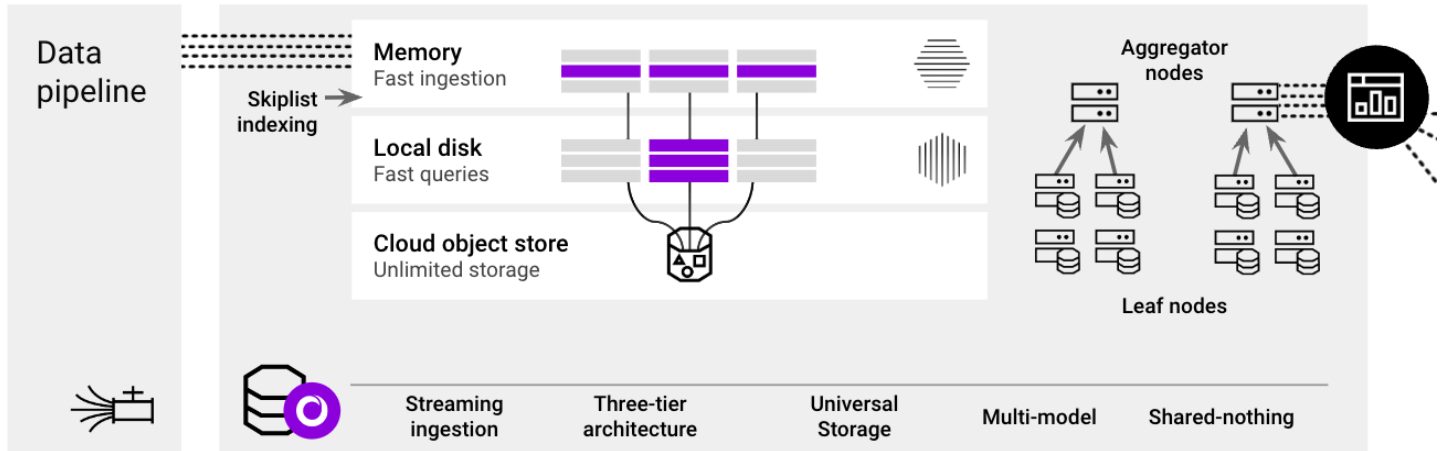


AlloyDB High-level architecture

([Source](#): “Deep Dive into Google’s AlloyDB” by Srini Vadlamani, 2022)

SingleStoreDB architecture

- **HTAP.** SingleStoreDB offers both in-memory rowstore and on-disk columnstores, enabling it to handle both OLTP and OLAP use cases while minimizing ETL and data duplication.
- **Distributed architecture:** A SingleStoreDB cluster is made up of nodes, which hold partitions of data and are responsible for query processing. Each node holds several partitions of data. Each partition is either a primary which can serve both reads and writes, or a replica which can only serve reads and is used for high availability.



SingleStoreDB's three-tier, distributed SQL architecture

Performance and Scalability

Both SingleStoreDB and AlloyDB are horizontally scalable. However, scaling in AlloyDB has some limitations:

- **Single node compute database.** Compute and storage cannot be scaled independently – and autoscaling with AlloyDB works partially, only for storage.
- Disabling an AlloyDB instance is not possible.
- AlloyDB's performance for write-heavy workloads is reportedly poor.
- AlloyDB uses tiered caching to improve performance.
- While it can scale storage independently, AlloyDB is essentially a single-node database. Adding replica nodes to AlloyDB helps with scaling RO throughput, but latency cannot be improved as there can be only a single primary node servicing a query.

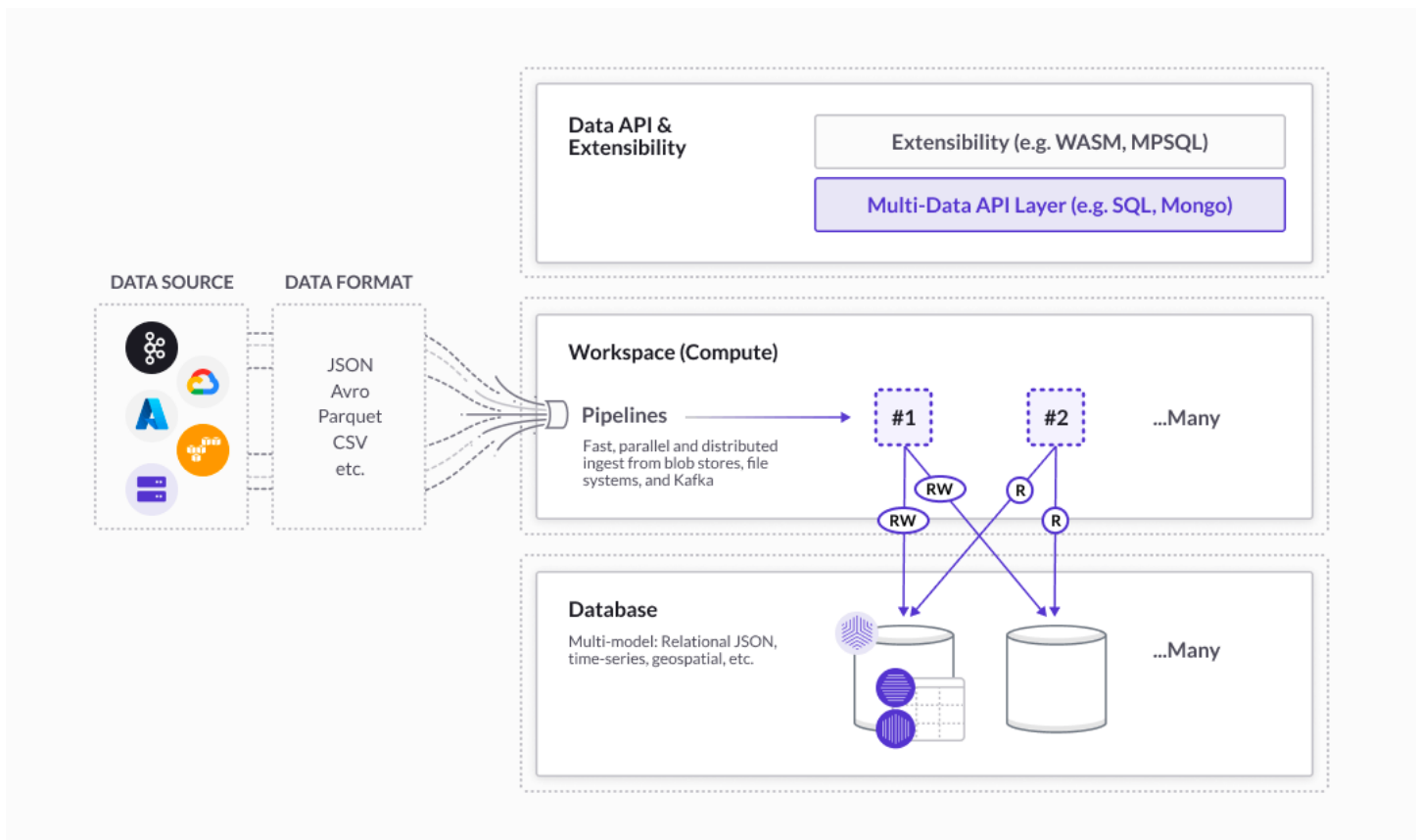
SingleStoreDB decouples storage and compute with its [unlimited storage](#) architecture that leverages three data storage tiers: in-memory, disk and cloud object storage. This allows users to effortlessly scale compute resources to meet the needs of any workload while managing the storage needs completely independently. SingleStoreDB's [Workspaces](#) feature allows isolation of workloads to power multiple workloads (read-only or read-write) on the same shared database.

Multi-model capabilities

- SingleStoreDB and AlloyDB (as an incarnation of PostgreSQL) are both multi-model databases and support relational, geospatial, time series, JSON and text data.
- SingleStoreDB natively supports vector data. With AlloyDB, a PostgreSQL extension might be needed to achieve similar functionality.

Extensibility

- AlloyDB's extensibility is similar to PostgreSQL and includes UDFs, custom data types, Foreign Data Wrappers (FDWs) and extensions like PostGIS or pgcrypto.
- SingleStoreDB provides multiple choices to increase the extensibility of your database operations. These include user-defined functions (**UDFs**), user-defined aggregate functions (**UDAFs**), table-valued functions (**TVFs**) and stored procedures (**SPs**) – all of which can be created, executed, replaced or deleted with the appropriate privileges. And with [Code Engine – Powered by Wasm](#), SingleStoreDB also supports extensibility for high-level languages including C, C++ and Rust.



Native data formats and extensibility of SingleStoreDB

Capability comparison

Capability	Why this matters	SingleStoreDB OLTP + OLAP	AlloyDB OLTP + OLAP
Speed	Immediate insights; responsive application experience for end users		
In-memory performance	Responsive applications and instantaneous analytics	●	◐
Streaming data ingestion	Immediate queryability of streaming data from multiple sources (files, Kafka, Spark, HDFS or object stores such as S3)	●	◐
Columnstore	Low-latency (~10s of milliseconds) on complex queries	●	◐
Analytics on semi-structured data	High-performance analytics on relational and semi-structured data in the same database engine	◐	◐
Scale	Adapt to growing needs		
Horizontal scalability	Distributed (shared-nothing) architecture that decouples storage and compute to allow scaling using low-cost infrastructure	●	◐
Read replicas/ multi-region deployments	Run multiple workloads and scale compute instances across shared databases	●	●
Resiliency	Run critical applications and workloads; mitigate risks on business operations and reputation.	●	●
Deploy anywhere	Ability to deploy both as a fully managed cloud service or self-managed on-premises	●	◐
Simplicity	Minimize complexity and costs		
SQL-powered OLTP + OLAP with zero ETL	Minimize data movement and duplication; minimize complexity and costs emanating from sprawl; power and simplicity of SQL for CRUD and rich query operations.	●	●
Multi-model	Ability to store and query multiple secondary data formats (JSON, time series, geospatial, full-text search, etc.)	●	●
Vector search engine	Efficiently handle large amounts of vector data for vector similarity search	◐	◐
Open-source software	Community-developed software that's typically free to use and distribute	○	◐

Recommendations

- As with many other GCP services, AlloyDB lacks product maturity. Therefore, it is recommended that customers considering AlloyDB implement a **Proof of Concept (POC)** to ensure that SLA, scalability, performance and cost expectations are met.
- Customers seeking to avoid cloud vendor lock-in may want to consider a data platform that provides more deployment flexibility.
- SingleStoreDB's capabilities make it suitable for diverse use cases including Customer 360, real-time analytics, enterprise generative AI applications, user-facing analytics, recommendation systems, fraud detection, IoT and more.

SingleStoreDB: Transact, analyze + contextualize data in real time.

SingleStoreDB is a real-time distributed SQL database that empowers users to transact, analyze and contextualize data in real time. It delivers stellar performance for both transactional applications (OLTP) and analytical workloads (OLAP).

With support for **streaming data ingestion** (Pipelines), a unique table type that supports **both transactions and analytics** (Universal Storage), **separation of compute and storage** (unlimited storage), limitless **point-in-time recovery** (PITR) and a **distributed** (shared-nothing), **MySQL-compatible** architecture, SingleStoreDB provides fast-growing companies the ability to build and scale real-time applications and analytics, generative AI applications and more.

All this comes with **compelling price performance**. SingleStore customers often report **2-3x better TPC-H performance at a fraction of the cost** of multiple technologies for OLTP, OLAP, ETL, NoSQL, etc.

Capabilities

SingleStoreDB empowers the world's makers to build, deploy and scale modern, intelligent applications — leading to faster, real-time decisions, lasting customer experiences and more cost-efficient operations.

Transact

Drive real-time analytics with super-fast ingest pipelines and a horizontally scalable architecture.

Analyze

Columnstore capabilities and aggregate functions for the most demanding analytics.

Contextualize

Vector functions, semantic and keyword search, and Notebooks provide a foundation for building AI applications.

Universal Storage

Unique single table type for transactions and analytics. Compatibility with ANSI SQL, MySQL + MariaDB ecosystem; NoSQL



Pure speed

Fast transactions, analytics, vectorization and query compilation.



Fast ingestion

SingleStore Pipelines — load data with updates.



Unlimited storage

Separation of storage + compute in a unified database



SingleStore Kai™

Power up to 1,000x faster analytics on JSON for applications built on MongoDB



Multi-model

Support for various data types, including vectors.



Run anywhere

Hybrid, multi-cloud, SaaS, on-premises, Kubernetes operator



Intelligence

Build generative AI and other apps with Notebooks on top of SingleStoreDB

