db insight

RESEARCH NOTE

AlloyDB: Google forges its own new PostgreSQL blend

Extending PostgreSQL with ML and operational analytics

Published July 2022

Executive Summary

Trigger

Google Cloud is taking another step to further round out its database portfolio. Taking a route that is similar to other cloud providers, it has chosen a flexible target, PostgreSQL, to develop a new API-compatible database with enhanced functionality. The new release, AlloyDB for PostgreSQL, takes the PostgreSQL API and then extends it under the hood with Google technologies to deliver a database that is not only faster and more scalable than the vanilla open source version available with Cloud SQL, Google Cloud's existing PostgreSQL service, but extends PostgreSQL into new territory with analytics and in-database machine learning (ML). So, what is the gap that AlloyDB fills within the Google Cloud database portfolio, and how does it stack up against other cloud PostgreSQL derivatives? And with other hybrid transaction/analytic cloud databases emerging, how does AlloyDB differentiate?

Our Take

AlloyDB is aptly named, as the definition of alloy is combining two metals to forge a new material with unique properties. In this case, AlloyDB takes the PostgreSQL API, which makes it a compatible platform appealing to the growing base of PostgreSQL developers, and under the hood, substitutes core Google technologies at the storage and AI/ML analytics layers. The result is a database designed for high throughput transaction processing and operational analytics. According to Google benchmarks, AlloyDB outperforms the standard PostgreSQL database by more than 4x, while leveraging Google distributed storage, ML-enabled autopilot systems, and Vertex AI for ML. Google's targeting of AlloyDB should sound familiar: as a destination from traditional OLTP databases to cloud-native platforms, which may sound similar to AWS's positioning of Amazon Aurora for PostgreSQL. But we believe that minimizes the added analytics functionality that differentiates it. Instead, we would compare it to Oracle's recently introduced MySQL HeatWave, along with other cloud database services that are also seeking to add operational workloads to analytics or vice versa.

We're not surprised that Google chose to introduce an extended PostgreSQL product, not only because AWS and Azure have done the same, but also because of the open-ended nature of the PostgreSQL open source project that encourages technology providers to reshape PostgreSQL. AlloyDB will take its place as a transaction processing platform with inline operational analytics; those analytics will complement, not replace BigQuery or Cloud Dataproc, which are designed for greater scale and more complex analytical workloads. Nonetheless, leveraging a common storage engine, we would like to see greater blending of these services because, at the end of the day, we believe that use case and programming environments are simply front ends that can leverage common building blocks to appeal to different target audiences and use cases.

AlloyDB is a fitting name

An alloy is a classic example of a 1 + 1 = 3 process, where the goal is generating a new metal with superior properties from the physical mixing of two different source metals, such as zinc and copper to make brass. And so, AlloyDB is a fitting name for Google's reimagining of PostgreSQL.

AlloyDB is a PostgreSQL-compatible offering that is API-compatible to PostgreSQL itself. It combines several Google core technologies, such as the disaggregated compute and storage architecture that underlies BigQuery, Spanner, and many other Google services, along with harnessing of Google ML technologies, both to optimize running of the database, and for developers and business analysts to use for running operational analytics models. Google benchmarks AlloyDB at more than 4x faster with transactions and up to 100x for analytic workloads compared to standard PostgreSQL. As we'll note later, as an API-compatible implementation of PostgreSQL, AlloyDB is hardly unique, especially among cloud database services. But the key is the use cases for which it targets: high-throughput transaction processing with inline operational analytics. Hold that thought.

At first glance, AlloyDB has parallels with BigQuery, in that both separate compute and storage (making them elastic), both use Google distributed storage, both are fully managed, both can integrate, and both perform analytics and in-database ML, but that's about as far as the similarities lie. AlloyDB is designed as a hybrid, combining high-throughput transaction processing (OLTP) with analytics. Spoiler alert: we think the two could be paired together as a blended service, but we're getting ahead of ourselves.

AlloyDB tiers storage and uses ML to optimize it. Base data is persisted to storage, and there are several levels of memory and flash for hotter data. It also adopts a common practice among cloud-native database services for multiple read replicas.

Let's dive deeper. It starts with low latency regional log storage for append-only write-ahead logs (this is the key to fast writes, and is much faster than writing pages), then a log processing service that asynchronously streams writes into primary and block storage replicas (where updates happen), and then for long term storage, sharded regional block storage across three zones. And because storage is separated from compute, write log processing can scale up and down with load, and there are multiple instances of log processors in each region.

Tiered storage is used for reads, where the hottest data is cached in memory, with the next tier in what Google terms "ultra-fast cache," which is actually flash, followed by long term distributed storage. Storage is sharded to balance IOPS. Since AlloyDB was designed also to handle analytic workloads, there is also an accompanying columnar storage.

The underlying smarts to all this is that AlloyDB uses ML to tier storage and direct queries to row or columnar tables based on analysis of the workload. AlloyDB also allows users to run ML models in-database. But unlike BigQuery ML (which is also an in-database ML capability), the models are not curated by Google; instead, it's the customer's responsibility to develop models in Vertex AI, from which they can be deployed to AlloyDB.

Where AlloyDB fits in the Google Cloud database portfolio

Ironically, the original design point was high throughput OLTP, with Google initially targeting customers migrating enterprise workloads to the cloud, who might otherwise consider Amazon Aurora for PostgreSQL. We believe that it is more than that; hold that thought.

In a nutshell, AlloyDB:

- Provides a higher performance, higher throughput replacement for vanilla PostgreSQL. Google's benchmarks estimate a 4x OLTP performance improvement over vanilla PostgreSQL. Conversely, Cloud SQL for PostgreSQL will continue to be suited for small/midsized workloads.
- Won't replace Cloud Spanner, as Spanner is aimed at globally distributed, active-active ACID transaction workloads traversing multiple regions.
- Won't replace BigQuery or other Google analytic services

There are close parallels with BigQuery, in that both run off the same underlying storage, have caching capabilities, and can run ML models in-database. But the elemental difference is that BigQuery currently does not support a PostgreSQL API (of course, it wouldn't be impossible to add one as an option, like Spanner just did). BigQuery is serverless, whereas AlloyDB is not. BigQuery has always been designed for analytics at scale (and has just added serverless Spark support to bolster that), and it also has some multimodel support (e.g., JSON data), whereas AlloyDB was primarily designed for high throughput transaction processing supplemented with lighter weight operational analytics and ML.

One thing that surprises us is that unlike BigQuery ML, Google is not limiting the ML algorithms that could run in AlloyDB. That places the burden on the developer's shoulders not to run a model that is overly complex or compute-intensive.

Our take is that the analytics face of AlloyDB will be best suited for in-line operational analytics and predictive models, which could likely cover use cases like customer churn, next-best answer, predictive maintenance, and so on. On that score, AlloyDB could conceivably run analytics with transactions in a closed loop manner because it has the high performance backplane to support it. But note that supported ML workloads won't run in cluster, but use the Vertex AI service that is integrated while running on separate nodes. As for Spark support, there's no reason that Google couldn't add it down the road, but as currently designed, it would not run in-cluster because AlloyDB is a provisioned service while Google's Spark implementation runs serverless. In our eyes, adding Spark wouldn't greatly change AlloyDB's addressable use cases.

In the Google Cloud portfolio, AlloyDB could perform closed-loop operational analytics. Conversely, BigQuery, Cloud Dataproc, Serverless Spark would be utilized for more complex analytics, while Dataflow would be used for streaming analytics. As we'll state below, we believe that AlloyDB could be blended with some of these services to provide "the PostgreSQL-friendly face" for developers versed on that platform.

How it compares to the usual suspects… and others

As we noted above, AlloyDB takes a path that is becoming more common as cloud providers rethink how they deliver databases. Put an API-compatible face, so that it can be utilized by the growing skills pools of open source (read: MySQL, PostgreSQL, MongoDB) developer bases. Hyperscaler and third parties alike offer PostgreSQL API-compatible services. That reflects the huge richness of the PostgreSQL ecosystem, all due to the fact that the open source project is so loosely licensed. As long as you hold UC Berkeley (which administers the PostgreSQL project) harmless, forking is widely accepted. And so, you have variants of PostgreSQL implemented as analytic column stores, time series databases, embedded databases, and other forms. There are just too many to list here.

Google initially designed AlloyDB as an answer for customers migrating legacy databases to the cloud. Since similar services are offered by AWS, among others, we believe that reduces the argument to a matter of feeds and speeds, where each provider will be constantly playing a game of leapfrog. Instead, we believe that Google should emphasize AlloyDB's hybrid nature, combining high-performance OLTP with analytics, even if that's not yet on the minds of the target audience.

Actually, the notion of blending operational databases and analytics together is hardly unique, although the cloud is reinventing it. The idea dates back to the hybrid row/column stores that debuted with on-premise incumbents such as IBM and Oracle roughly a decade ago. Because of the limits of on-premises implementation, capacity was a consideration, and so the analytic column stores typically only replicated a subset of transaction data.

Cloud-native architecture has reinvented the hybrid row/column store database, removing upper limits on compute and column storage; reducing or eliminating resource contention issues; and providing more operational and cost flexibility. In our eyes, that's where the real comparisons for AlloyDB come.

In that light, Oracle MySQL HeatWave is the closest counterpart. Admittedly, it works off a different database – MySQL – but has parallels in applying an API-compatible implementation atop a fit-for-purpose storage engine pairing row and column stores. Both also use ML for helping run the database (in Oracle's case, it's for recommendations rather than closed-loop automation that Google provides) and for in-database running of lightweight predictive models. Aside from the fact that AlloyDB is PostgreSQL and HeatWave is MySQL, the biggest difference is the design point for OLTP: AlloyDB targets high-throughput transaction processing whereas Oracle points such use cases to its flagship database (although it has been raising upper limits for HeatWave cluster and node size).

Google and Oracle are hardly alone in converging operational transactions with analytics in a cloud service. There are some parallels with MariaDB SkySQL that pairs transaction and analytic processing, but without the ML. There are other cloud database providers just starting this journey. Snowflake adding a lightweight hybrid transaction database to its core analytics cloud service that at this point lacks the scale or performance of AlloyDB, while MongoDB is taking its first steps toward building analytic capability with its first SQL query engine, a column store, federated query, and node partitioning. SingleStore, the former MemSQL, has recast its in-memory platform to converge analytics and streaming with transactions.

At this point, Oracle MySQL HeatWave is AlloyDB's only serious competition for analytics and in-database ML at scale in a converged transaction/analytics cloud service.

Takeaways

AlloyDB is very much a logical step forward for Google. As PostgreSQL is a very flexible platform, we were wondering when Google would finally get around to flexing its technology muscles to extend the popular database. It's an auspicious start. Not surprisingly, there are still functional gaps to be filled in. For instance, with columnar data, it offers min/max/average to reduce scanning overhead, but we expect Bloom filters to come later. We would also like to see a serverless option come down the track.

AlloyDB will provide a higher-performance alternative to Cloud SQL for PostgreSQL. But as noted earlier, AlloyDB won't replace Google Cloud's analytic services or distributed Spanner transaction platform.

As noted earlier, while Google originally designed AlloyDB for migrating high-throughput transactional workloads, we think that sells the database short.

We view AlloyDB as a platform for bringing high-throughput transaction processing and operational analytics together, whereas products like Aurora are designed for transactions. Typical use cases could involve next-best actions in e-commerce transaction systems or identifying financial fraud in retail banking or capital market transactions. Trained offline in the data lake, the ML models used in AlloyDB would be lightweight, delivering smart yes/no or classification decisions on fresh data in an instant.

Here's where we think things could get interesting. We've been on record that cloud providers must make their clouds easier to use. Typically, that has been done within the confines of a managed service, but when that portfolio of managed services grows wide enough, customers with divergent needs must shoulder the burden of not only choosing from large menus, but then integrating those services together. As AlloyDB shares common long term storage with BigQuery and Dataproc, maybe AlloyDB could just become another way of surfacing the same data. That of course requires common data governance and security, but that's where Dataplex could come in. And by the way, that's where integration with Serverless Spark could come in for data engineering, or Dataflow for closed-loop filtering and analytics of streaming data. At preview launch, AlloyDB will be able to stream from its write ahead log processing layer to BigQuery, which is a start.

Author

Tony Baer, Principal, dbInsight

tony@dbinsight.io

Twitter @TonyBaer

About dbInsight

dbInsight LLC® provides an independent view on the database and analytics technology ecosystem. dbInsight publishes independent research, and from our research, distills insights to help data and analytics technology providers understand their competitive positioning and sharpen their message.

Tony Baer, the founder and principal of dbInsight, is a recognized industry expert on data-driven transformation. *Onalytica* named him as a Top Cloud Influencer for 2022 for the fourth straight year. *Analytics Insight* named him one of the 2019 Top 100 Artificial Intelligence and Big Data Influencers. His combined expertise in both legacy database technologies and emerging cloud and analytics technologies shapes how technology providers go to market in an industry undergoing significant transformation.

dbInsight® is a registered trademark of dbInsight LLC.