Jinmu Leverages Couchbase for Real-time Data Generated by AI Assistant

Jinmu Information is a professional IT data consulting and service provider, committed to providing users with high-quality information products, consulting, and technical services. The company supports NoSQL database solutions across China, Hong Kong, and Singapore for over 100 leading clients in finance, insurance, gaming, e-commerce, and other industries. Jinmu needed a document based NoSQL database that could provide a consistent and highly-available key value store for its client.

Challenges

- Needed a document based NoSQL database that could provide a consistent and highly-available key value store
- Solution with vector search capabilities, SQL syntax support, and the ability to store and process time series data

Outcomes

- The database supports query acquisition under high concurrency
- Improved application performance and stability
- 70 million documents, 8,000 operations per second with 10ms response times

Industry
- High Tech

Customer application
- AI Assistant for virtual meetings

Use case
- Artificial intelligence

Product
- Couchbase Server

Key Features
- In-memory/caching
- SQL++
- Full Text Search

"Couchbase stores all the real-time communications data and supports query acquisition under high concurrency. This has helped greatly improve the performance and stability for the AI Assistant application."

— Andy Qiu, CEO, Jinmu

One of Jinmu's clients is a company dedicated to audio and video solutions. The client implemented a new innovative project – an AI Assistant – that combines artificial intelligence capabilities with real-time audio and video features. To support this new project, Jinmu needed a document based NoSQL database that could provide a consistent and highly-available key value store.

Jinmu selected Couchbase over MongoDB primarily due to Couchbase's vector search capabilities, SQL syntax support, and the ability to store and process time series data. The AI Assistant is designed to capture and organize conversations between users in chronological order. It summarizes the main discussion points, draws conclusions, and provides notifications for when certain topics were discussed. By leveraging Couchbase's Vector Search, the assistant can quickly and accurately retrieve relevant context from past exchanges. It also supports Retrieval-Augmented Generation (RAG) to ensure that the assistant can generate responses that are both contextually accurate and timely, helping users navigate through discussions with seamless, intelligent support.

All this real-time communications data is stored within Couchbase. The database supports query acquisition under high concurrency and helps improve the application performance and stability. Now with Couchbase, the client's AI Assistant can offer external services to users safely and efficiently – totaling 70 million documents, 8,000 operations per second with 10ms response times. In the future, Jinmu will use Couchbase even further to integrate time series data, vector search, and LLM capabilities within the application.

Modern customer experiences need a flexible database platform that can power applications spanning from cloud to edge and everything in between. Couchbase's mission is to simplify how developers and architects develop, deploy and run modern applications wherever they are. We have reimagined the database with our fast, flexible and affordable cloud database platform Capella, allowing organizations to quickly build applications that deliver premium experiences to their customers – all with best-in-class price performance. More than 30% of the Fortune 100 trust Couchbase to power their modern applications. For more information, visit www.couchbase.com and follow us on X (formerly Twitter) @couchbase.