

Bobble AI reduced data processing cost by 33% by migrating to Google Cloud



Client profile

Bobble AI is a cutting-edge conversational media platform that enhances communication through personalized and expressive content, such as stickers, emojis and GIFs. Leveraging AI and deep learning, Bobble AI integrates seamlessly into keyboards, messaging apps and social platforms, enabling users to create engaging, context-aware conversations. Their focus on personalization and innovation makes Bobble AI a leader in the digital content space, transforming how people interact through mobile devices.

Dealing with a high volume of keyboard events from its extensive user base, Bobble AI sought to optimize the performance and cost-efficiency of its data processing. Their goal was to reduce expenses while enhancing the speed and efficiency of the processing layer.

We helped them migrate their data processing to Google Cloud Dataproc and Cloud Composer. This enabled them to leverage the high availability of spot instances on Google Cloud and reduce the cost of Spark clusters with up to 90% discounts. We also helped them build a robust continuous integration/continuous delivery (CI/CD) pipeline which delivered a 5% improvement in pipeline runtime.



NTT DATA transformed our cloud infrastructure, reducing costs and improving efficiency. Their expertise and tailored approach made the migration seamless and highly effective.”

Kunal Dawn, VP of Engineering - Backend, Bobble.ai

33% reduction in data-processing costs

5% improvement in pipeline runtime

Business need

Reduce expenses while enhancing the speed and efficiency of processing

Data processing workloads were using spot instances, which were frequently unavailable in required regions, leading to processing delays.

Bobble AI preferred open source tools over cloud provider solutions but hesitated to use managed services due to concerns about losing control over the process.

Without a CI/CD pipeline, deployments were done manually, making the process slow and cumbersome. They were looking to implement automated deployment pipelines to streamline and accelerate the deployment process.



Google Cloud and NTT DATA helped migrate data processing and analytics infrastructure by adopting highly modern, scalable infrastructure solutions. We achieved this with lower cost. Reduced job run times helped in our timely business delivery, meeting future demands for analytics and machine learning outcomes.”

Vipul Singhal, VP - Data Intelligence, Bobble.ai

Solution

Migrating data processing to Google Cloud Dataproc and Cloud Composer

NTT DATA recommended migrating Bobble AI's workloads to Google Cloud's **Dataproc** and **Cloud Composer**. Cloud Composer, which is based on Apache Airflow, replaced the existing Step Functions for extract, transform, load (ETL) orchestration, addressing Bobble AI's initial concerns about adopting managed services — and highlighting the advantages of this transition.

The aim was to leverage the high availability of spot instances on Google Cloud Platform (GCP), which offered up to 90% cost savings for Spark clusters.

The migration was carefully executed by transitioning each Spark job individually, selecting optimal machine types and configurations based on the job type. At each step, the cost and runtime of the jobs on GCP were compared with those on their existing cloud provider, ensuring efficiency improvements. A custom utility was built in Cloud Composer to orchestrate Dataproc cluster creation and updates.

The utility prioritized spot instances, automatically switching to on-demand instances when necessary, ensuring high availability and cost-efficiency.

Data from the Bobble Keyboard application flows into **Google Cloud Storage** buckets, where it is processed at scheduled intervals. Cloud Composer, serving as the ETL orchestration tool, dynamically creates Dataproc clusters, optimizes spot instance usage and submits Spark jobs. All sensitive information is securely stored in **Secret Manager**, and a CI/CD pipeline was built using Cloud Build for seamless deployment. Postmigration, the availability issues with spot instances were resolved, leading to faster completion of data processing.

Outcomes

NTT DATA provided Bobble AI with the expertise needed for a seamless data processing migration to Google Cloud. By leveraging cutting-edge technologies and best practices, we helped Bobble AI reduce data processing costs through the optimized use of GCP spot instances. Additionally, NTT DATA implemented automated CI/CD pipelines, streamlining deployments and improving efficiency. This partnership addressed Bobble AI's concerns about managed services while enhancing their cloud capabilities for future growth.

Highlights:

- **33% reduction in data processing cost**

The high availability of spot instances on GCP led to a 33% reduction in the cost of running Spark clusters with discounts of up to 90%.

- **Faster and more efficient data analytics**

With BigQuery Serverless application programming interface (API), the data can now be analyzed faster and more efficiently.

- **Modernized data processing**

By using Dataproc and Cloud Composer we enabled a more modernized approach to data processing, putting future-proof strategies in place.

- **Improved availability and efficiency**

GCP's consistent availability of spot instances eliminated processing delays, allowing workloads to be completed faster and more efficiently.

- **CI/CD pipeline implementation**

A new CI/CD pipeline helped speed up the development process and reduce manual errors.

- **Streamlined ETL orchestration**

Cloud Composer, based on Apache Airflow, was implemented for efficient ETL orchestration, automating the data flow and Spark job submission on Dataproc clusters.

- **Enhanced security and compliance**

Sensitive credentials were securely managed through Google Cloud Secret Manager, ensuring data security and regulatory compliance throughout the process.

- **Flexibility with open source tools**

The adoption of open source technologies like Apache Airflow and Spark allowed Bobble AI to maintain control over their stack while benefiting from the scalability and cost-efficiency of managed GCP services.

Visit nttdata.com to learn more.

NTT DATA is a global innovator of digital business and technology services, helping clients innovate, optimize and transform for success. As a Global Top Employer, we have experts in more than 50 countries and a robust partner ecosystem. NTT DATA is part of NTT Group.

