

TAVANT

Transforming enterprise data platforms

ADVERTISING DATA WAREHOUSE RE-ENGINEERED

*A 1.5 PB advertising warehouse processing
roughly 30 TB of input data daily and serving
75M daily and 190M monthly active users.*



THE CLIENT AND THE CHALLENGE

A major global video game publisher monetizing at massive scale, with hundreds of millions of players, extremely high event velocity and revenue that is sensitive to advertising performance across its titles.

The publisher needed to migrate its advertising data warehouse and ETL from a brittle, Jenkins-centric setup to Google Cloud, take ownership of its pipelines and data model, and get better insight into advertising performance, including eCPM, fill rate, and revenue attribution, to monetize more titles.

THE SOLUTION

Tavant re-engineered the advertising data warehouse on Google Cloud. Groovy- and Jenkins-based ETL was migrated to Airflow and Cloud Composer, ingestion was standardized on Dataflow and Pub/Sub, and BigQuery became the analytical backbone. Tavant set up CI/CD with Jenkins and Kubernetes on GKE, redesigned a generic advertising data model, integrated new revenue and telemetry sources, and strengthened monitoring and alerting across the platform using BigQuery, Dataflow, Looker, Cloud Storage, GKE, and Pub/Sub.

THE IMPACT

A 1.5 PB advertising warehouse processing roughly 30 TB of input data daily and serving 75M daily and 190M monthly active users. Enabled monetization of more game titles. Business teams monitor key advertising KPIs and optimize floor prices, while data science runs A/B experiments on a scalable, automated Google Cloud platform with near-real-time analytics.



ARCHITECTURE & PLATFORM

CORE ARCHITECTURE

A BigQuery-centered advertising warehouse on Google Cloud, with Airflow on Cloud Composer orchestrating pipelines migrated off Groovy and Jenkins ETL.

CLOUD & INTEGRATION

Standardized ingestion on Dataflow and Pub/Sub, a generic advertising data model spanning revenue and telemetry sources, and CI/CD, monitoring, and alerting running on GKE.

TECHNOLOGY STACK

CORE STACK

BigQuery, Cloud Composer, Airflow, Dataflow, Pub/Sub, Cloud Storage, GKE, Kubernetes, Jenkins, Looker

SUPPORTING STACK

Cross-cutting tooling and platforms supporting the core stack.

BUSINESS VALUE DELIVERED

OUTCOMES

A 1.5 PB advertising warehouse processing roughly 30 TB of input data daily and serving 75M daily and 190M monthly active users.

OPERATING IMPACT

Enabled monetization of more game titles.

ACCELERATORS & CAPABILITY

REUSABLE ASSETS

Generic advertising data model reusable across titles and revenue sources.

CAPABILITY BUILT

Standardized Dataflow and Pub/Sub ingestion pattern and hardened GKE-based CI/CD, monitoring and alerting for event-heavy, high-cardinality datasets.