

Consolidating Insights Securely with a Data Clean Room for Semiconductor Major

Business Problem



- Establish a secure, governed Data Clean Room (DCR) environment with privacy, within Google BigQuery
- Enable safe and aggregated data analysis across organizational datasets
- Ensure collaborative, data-driven insights both within the organization and with trusted external partners
- Optimize manufacturing supply chain, improving product quality, and enhancing operational efficiency

Solution



- Application built using Google BigQuery's data-sharing capabilities
- Web interface using React JS using a Python API for orchestration and logic
- Implemented attribute-level masking, obfuscation, and granular access controls to enable joint analysis without exposing sensitive raw data
- Infrastructure provisioned via Terraform and hosted on a Google Cloud VM
- Utilised Vertex AI Workbench for analytics and data ingestion through AWS S3, Azure ADLS, GCS
- Utilized Google Cloud IAM for access control

Value Delivered



- ✓ **Faster time-to-insight:** Reduced dependency on engineering teams
- ✓ **Improved data ROI:** Safe reuse and monetization of sensitive datasets
- ✓ **Operational efficiency:** Streamlined partner onboarding and governance workflows
- ✓ **Regulatory confidence:** Built-in controls for privacy, security, and auditability
- ✓ **Collaborative analysis:** Enabled secure data analysis with external partner without exposing confidential data



50K TPS

Data Successfully Migrated to Snowflake on AWS



3500+

ETL Scripts Migrated, Including 700+ Converted in Under 7 Weeks



**~10 Hrs
→ 5 Min**

Query Performance Improvement After Platform Modernization