Engineer Data

Designed and ran 6 ETL/ELT pipelines

Designed, implemented, and administered 6 ETL/ELT pipelines, utilizing Google BigQuery, MSSQL, PostgreSQL, Shell scripting, PL/pgSQL, and Transact-SQL, integrating data for efficient Python API processing.

2022 – 2024

6

ETL pipelines across three database engines

Situation. The organization needed to consolidate and process large volumes of tabular data from a remote data warehouse located behind an IPsec VPN. This data was crucial for powering internal analytics dashboards and external Python APIs used by clients and partners.

Task. The objective was to design, implement, and maintain a set of robust and automated ETL/ELT pipelines to securely retrieve, transform, and load data into Google BigQuery, ensuring data accuracy, performance, and scalability across systems.

Action. 6 end-to-end ETL/ELT pipelines were designed, implemented, and administered, spanning multiple technologies including Google BigQuery, MSSQL, PostgreSQL, Shell scripting, PL/pgSQL, and Transact-SQL.

- Established secure connections to a remote server protected by an IPsec VPN to automate data retrieval.
- Scheduled and executed downloads of compressed data archives containing Parquet, CSV, and .bak files.
- Developed Bash and Python scripts to extract and classify files based on type and schema.
- For .bak files, backups were restored into a local MSSQL Server instance using RESTORE DATABASE workflows and schema integrity validated.
- Loaded tabular data into staging schemas in MSSQL and PostgreSQL, using bcp, psql, and SSIS tools, depending on the source format.

- Wrote modular and reusable PL/pgSQL and T-SQL procedures to clean, normalize, and enrich the data based on business logic.
- Exported curated datasets and tables into intermediate formats, compressed them using gzip, and securely transferred them to a Google Cloud Storage bucket.
- Automated the upload and schema mapping processes to Google BigQuery, using bq CLI and Python-based data ingestion scripts.

Result. These automated pipelines significantly reduced the manual overhead and processing time — from 3+ hours of manual work to under 20 minutes end-to-end, while improving data freshness from weekly to daily syncs. The Python APIs consuming this data saw a 30% performance improvement, and the enhanced visibility helped business analysts deliver faster insights to stakeholders. The solution remains scalable and extensible for onboarding new data sources as the business grows.

Bring this experience to your project

Use “Designed and ran 6 ETL/ELT pipelines” as a starting point for a similar challenge in your organisation. The services behind this work are available for your project. A structure that matches how the business actually works, so the questions you need answered later are still possible to ask.

Explore our services Get in touch →

Look behind the result

Read the situation, decisions and result in this example. We can discuss which parts transfer to your constraints and which need another approach. Designed and ran 6 ETL/ELT pipelines — 6 ETL pipelines across three database engines

Explore our services Get in touch →

Need to bring data from many sources together?

Talk through your project with the engineer responsible for delivery. We agree the scope, price basis and handover in writing before work begins.

Role context: Data Engineer. It draws on Databases, PostgreSQL, SQL, Data Warehousing, Data Engineering, Data Pipelines (ETL/ELT), Cloud, Automation & CI/CD, Networking & VPN and Python. And while you are here, we are rather good at Backend & API Development, Database Design & Modeling, Data Pipeline Development (ETL/ELT), Data Warehouse Design and AI Data Readiness.

Get in touch Explore our services

One piece of **a longer record** — and it is all on this site.

Every entry is written the same way: the situation, the task, what we did and what changed.

See the whole record

Define the improvement you need

Bring a representative dataset, the questions it must answer and the current bottleneck. We can agree how to measure quality, throughput and recovery.

See what it costs Get in touch →

Need to bring data from many sources together?

Ask about the decisions behind “Designed and ran 6 ETL/ELT pipelines”. You will speak with the people who can explain the work and assess your own project.

Explore our services

welcome@engineer.company +45 50 10 61 84 Copenhagen, Denmark

Take this page with you

Contact card

Carefully Crafted With Love by D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Copyright © 2025 – present · Engineer ApS · All rights reserved.

Printed from https://data.engineer.company/portfolio/designed-and-ran-6-etl-elt-pipelines/