Engineer Daten

6 ETL/ELT-Pipelines entworfen und betrieben

6 ETL/ELT-Pipelines mit Google BigQuery, MSSQL, PostgreSQL, Shell-Skripten, PL/pgSQL und Transact-SQL entworfen, umgesetzt und administriert sowie Daten für eine effiziente Verarbeitung durch die Python-API integriert.

2022 – 2024

6

ETL-Pipelines über drei Datenbankengines hinweg

Situation. Das Unternehmen musste große Mengen tabellarischer Daten aus einem entfernten Data Warehouse hinter einem IPsec-VPN zusammenführen und verarbeiten. Diese Daten waren wesentlich für interne Analysedashboards und externe Python-APIs für Kunden und Partner.

Aufgabe. Belastbare, automatisierte ETL/ELT-Pipelines sollten entworfen, umgesetzt und gepflegt werden, um Daten sicher abzurufen, zu transformieren und nach Google BigQuery zu laden. Datenrichtigkeit, Performance und systemübergreifende Skalierbarkeit mussten gewährleistet sein.

Massnahmen. 6 durchgängige ETL/ELT-Pipelines wurden entworfen, umgesetzt und administriert. Sie verbanden Google BigQuery, MSSQL, PostgreSQL, Shell-Skripte, PL/pgSQL und Transact-SQL.

- Sichere Verbindungen zu einem durch IPsec-VPN geschützten entfernten Server automatisierten den Datenabruf.
- Downloads komprimierter Archive mit Parquet-, CSV- und .bak-Dateien wurden geplant und ausgeführt.
- Bash- und Python-Skripte entpackten und klassifizierten Dateien nach Typ und Schema.

- .bak-Sicherungen wurden mit RESTORE-DATABASE-Abläufen in einer lokalen MSSQL-Server-Instanz wiederhergestellt; die Schemaintegrität wurde geprüft.
- Tabellarische Daten wurden abhängig vom Quellformat mit bcp, psql und SSIS in Staging-Schemata von MSSQL und PostgreSQL geladen.
- Modulare, wiederverwendbare PL/pgSQL- und T-SQL-Prozeduren bereinigten, normalisierten und ergänzten die Daten gemäß der Geschäftslogik.
- Aufbereitete Datensätze und Tabellen wurden in Zwischenformate exportiert, mit gzip komprimiert und sicher in einen Google-Cloud-Storage-Bucket übertragen.
- Upload und Schemazuordnung nach Google BigQuery wurden mit der bq-CLI und Python-Skripten zur Datenübernahme automatisiert.

Ergebnis. Die automatisierten Pipelines verringerten manuellen Aufwand und Verarbeitungszeit erheblich — von über 3 Stunden Handarbeit auf unter 20 Minuten für den gesamten Ablauf. Zugleich wurden die Daten statt wöchentlich nun täglich abgeglichen. Die Performance der nutzenden Python-APIs verbesserte sich um 30 %, und die bessere Transparenz half Geschäftsanalysten, Stakeholder schneller mit Erkenntnissen zu versorgen. Die Lösung bleibt skalierbar und erweiterbar, wenn mit dem Unternehmenswachstum neue Datenquellen hinzukommen.

Diese Erfahrung für Ihr Projekt nutzen

Nutzen Sie „6 ETL/ELT-Pipelines entworfen und betrieben“ als Ausgangspunkt für eine ähnliche Herausforderung in Ihrer Organisation. Die Leistungen hinter dieser Arbeit stehen auch für Ihr Projekt zur Verfügung. Eine Struktur, die zur tatsächlichen Arbeitsweise Ihres Unternehmens passt, damit sich auch spätere Fragen noch beantworten lassen.

Unsere Leistungen entdecken Kontakt aufnehmen →

Hinter das Ergebnis schauen

Lesen Sie die Situation, die Entscheidungen und das Ergebnis dieses Beispiels. Wir können besprechen, welche Teile sich auf Ihre Rahmenbedingungen übertragen lassen und welche einen anderen Ansatz erfordern. 6 ETL/ELT-Pipelines entworfen und betrieben — 6 ETL-Pipelines über drei Datenbankengines hinweg

Unsere Leistungen entdecken Kontakt aufnehmen →

Müssen Daten aus vielen Quellen zusammengeführt werden?

Besprechen Sie Ihr Projekt mit der Ingenieurin, die für die Umsetzung verantwortlich ist. Wir vereinbaren Umfang, Preisgrundlage und Übergabe schriftlich, bevor die Arbeit beginnt.

Rollenkontext: Dateningenieurin. Die Arbeit baut auf Datenbanken, PostgreSQL, SQL, Data Warehousing, Data Engineering, Datenpipelines (ETL/ELT), Cloud, Automatisierung und CI/CD, Netzwerke und VPN und Python. Und wenn Sie schon hier sind: Wir sind auch ziemlich gut in Backend- und API-Entwicklung, Datenbankdesign und Datenmodellierung, Entwicklung von Datenpipelines (ETL/ELT), Data-Warehouse-Design und Daten für KI nutzbar machen.

Kontakt aufnehmen Unsere Leistungen entdecken

Ein Teil einer **längeren Arbeitsbilanz** — vollständig auf dieser Website zu finden.

Jeder Eintrag folgt demselben Aufbau: die Situation, die Aufgabe, unser Vorgehen und die Veränderungen.

Alle Arbeiten ansehen

Die benötigte Verbesserung definieren

Bringen Sie einen repräsentativen Datensatz, die Fragen, die er beantworten muss, und den aktuellen Engpass mit. Wir können vereinbaren, wie Qualität, Durchsatz und Wiederherstellung gemessen werden.

Preise ansehen Kontakt aufnehmen →

Müssen Daten aus vielen Quellen zusammengeführt werden?

Fragen Sie nach den Entscheidungen hinter „6 ETL/ELT-Pipelines entworfen und betrieben“. Sie sprechen mit den Menschen, die die Arbeit erklären und Ihr eigenes Projekt einschätzen können.

Unsere Leistungen entdecken

welcome@engineer.company +45 50 10 61 84 Kopenhagen, Dänemark

Nehmen Sie diese Seite mit

Kontaktkarte

Mit Sorgfalt und Liebe gestaltet von D.

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

Copyright © 2025 – heute · Engineer ApS · Alle Rechte vorbehalten.

Gedruckt von https://data.engineer.company/de/portfolio/designed-and-ran-6-etl-elt-pipelines/