Engineer Datos

Diseño y operación de 6 flujos ETL/ELT

Diseñé, implementé y administré 6 flujos ETL/ELT con Google BigQuery, MSSQL, PostgreSQL, scripts de Shell, PL/pgSQL y Transact-SQL, integrando los datos para su procesamiento eficiente mediante una API en Python.

2022 – 2024

6

flujos ETL en tres motores de bases de datos

Situación. La organización necesitaba consolidar y procesar grandes volúmenes de datos tabulares procedentes de un almacén remoto situado detrás de una VPN IPsec. Estos datos eran esenciales para los paneles analíticos internos y las API Python externas utilizadas por clientes y socios.

Tarea. El objetivo consistía en diseñar, implementar y mantener flujos ETL/ELT sólidos y automatizados para recuperar, transformar y cargar datos en Google BigQuery de forma segura, manteniendo precisión, rendimiento y escalabilidad entre sistemas.

Acción. Se diseñaron, implementaron y administraron 6 flujos ETL/ELT completos con Google BigQuery, MSSQL, PostgreSQL, scripts de Shell, PL/pgSQL y Transact-SQL.

- Se establecieron conexiones seguras con un servidor remoto protegido por una VPN IPsec para automatizar la recuperación de datos.
- Se programaron y ejecutaron descargas de archivos comprimidos con datos Parquet, CSV y .bak.
- Se desarrollaron scripts Bash y Python para extraer y clasificar los archivos según su tipo y esquema.
- Las copias .bak se restauraron en una instancia local MSSQL Server mediante procesos RESTORE DATABASE y se validó la integridad del esquema.
- Se cargaron datos tabulares en esquemas intermedios de MSSQL y PostgreSQL con bcp, psql y SSIS, según el formato de origen.

- Se escribieron procedimientos modulares y reutilizables PL/pgSQL y T-SQL para limpiar, normalizar y enriquecer los datos conforme a la lógica de negocio.
- Se exportaron las tablas y conjuntos depurados a formatos intermedios, se comprimieron con gzip y se transfirieron de forma segura a un bucket Google Cloud Storage.
- Se automatizaron la carga y la correspondencia de esquemas en Google BigQuery mediante bq CLI y scripts de ingesta Python.

Resultado. Los flujos automatizados redujeron considerablemente el esfuerzo manual y el tiempo de procesamiento: de más de 3 horas de trabajo manual a menos de 20 minutos de principio a fin, mejorando además la actualidad de los datos al pasar de sincronizaciones semanales a diarias. Las API Python que consumían los datos mejoraron su rendimiento un 30 % y la mayor visibilidad permitió a los analistas de negocio proporcionar información con más rapidez a las partes interesadas. La solución sigue siendo escalable y ampliable para incorporar nuevas fuentes a medida que crece el negocio.

Aplique esta experiencia a su proyecto

Utilice «Diseño y operación de 6 flujos ETL/ELT» como punto de partida para un reto similar en su organización. Los servicios que respaldan este trabajo están disponibles para su proyecto. Una estructura que refleja cómo funciona el negocio para poder responder también a las preguntas que surjan más adelante.

Explore nuestros servicios Póngase en contacto →

Mire más allá del resultado

Lea la situación, las decisiones y el resultado de este ejemplo. Podemos hablar de qué partes se adaptan a sus restricciones y cuáles necesitan otro enfoque. Diseño y operación de 6 flujos ETL/ELT — 6 flujos ETL en tres motores de bases de datos

Explore nuestros servicios Póngase en contacto →

¿Necesita reunir datos de muchas fuentes?

Converse sobre su proyecto con la ingeniera responsable de la entrega. Acordamos por escrito el alcance, la base del precio y el traspaso antes de empezar a trabajar.

Contexto de la función: Ingeniera de datos. Se apoya en Bases de datos, PostgreSQL, SQL, Almacenes de datos, Ingeniería de datos, Flujos de datos (ETL/ELT), Nube, Automatización y CI/CD, Redes y VPN y Python. Y ya que está aquí, también se nos da bien Desarrollo de backend y API, Diseño y modelado de bases de datos, Desarrollo de flujos de datos (ETL/ELT), Diseño de almacenes de datos y Preparación de datos para IA.

Póngase en contacto Explore nuestros servicios

Una parte de **una trayectoria más amplia** — y toda ella está en este sitio.

Cada entrada sigue la misma estructura: la situación, la tarea, lo que hicimos y lo que cambió.

Vea la trayectoria completa

Defina la mejora que necesita

Traiga un conjunto de datos representativo, las preguntas que debe responder y el cuello de botella actual. Podemos acordar cómo medir la calidad, la capacidad de procesamiento y la recuperación.

Vea cuánto cuesta Póngase en contacto →

¿Necesita reunir datos de muchas fuentes?

Pregunte por las decisiones detrás de «Diseño y operación de 6 flujos ETL/ELT». Hablará con las personas que pueden explicar el trabajo y valorar su propio proyecto.

Explore nuestros servicios

welcome@engineer.company +45 50 10 61 84 Copenhague, Dinamarca

Llévese esta página

Tarjeta de contacto

Creado con esmero y cariño por D.

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

Derechos de autor © 2025 – actualidad · Engineer ApS · Todos los derechos reservados.

Impreso desde https://data.engineer.company/es/portfolio/designed-and-ran-6-etl-elt-pipelines/