ANNUAL COST SAVINGS FOR PORTLAND GENERAL ELECTRIC WITH THE SNOWFLAKE DATA CLOUD

UTILITIES

PGE

COMPANY Portland General Electric

LOCATION Portland, Oregon

SNOWFLAKE WORKLOADS USED

DATA ENGINEERING | DATA LAKE | DATA WAREHOUSE | DATA SCIENCE | DATA APPLICATIONS | DATA SHARING

Portland General Electric (PGE) is a fully integrated energy company based in Portland, Oregon, with operations across the state. The company serves 901,000 customers and a service area population of 1.9 million Oregonians in 51 cities.

PGE has 16 generation plants in five Oregon counties, and it maintains and operates 14 public parks and recreation areas. For over 130 years, PGE has delivered safe, affordable, and reliable energy to Oregonians.

STORY HIGHLIGHTS:

Near-zero maintenance
Snowflake helped Portland General Electric eliminate the need for DBAs.

Secure data sharing
Secure data sharing lets Portland General Electric seamlessly share data with internal teams and external partners.

Secure views
Secure views help Portland General Electric implement user-level data security.

> "Our data science team works with a lot of data. In the past, providing them a copy of our production data required the work of a full-time DBA. With Snowflake Secure Data Sharing, it's as easy as the click of a button. We securely share production data without doing any copying. It's helped the data science team tremendously."
>
> **—ARAVIND MURUGESH,**
> Principal Data Architect, Portland General Electric

CHALLENGE:

Managing an expensive, inflexible legacy data warehouse
PGE managed a legacy, on-premises data warehouse that was expensive to maintain and had performance issues. The system's tightly-coupled architecture was inflexible. In addition, the architecture caused copies of data to proliferate across the organization, making it difficult to identify the authoritative source of data and increasing storage costs.

According to PGE's Principal Data Architect, Aravind Murugesh, "Because our data was in multiple locations and we had integration tools and complex business logic to support them, we were not able to deliver self-service tools to our users."

Updating the data warehouse required bringing the entire system down for a day, which was labor-intensive. In addition, the system had a one-day backlog, preventing users from accessing real-time data.

SOLUTION:

Snowflake and AWS for high performance and low maintenance
Realizing the need for a modern data environment, PGE turned to Snowflake.

PGE evaluated numerous cloud data platforms, including those from Google, Amazon, and Microsoft. It chose the Snowflake Data Cloud on AWS for its high performance, separation of storage from compute, near-zero maintenance, and micro-partitioning.

Data migration into a Snowflake Data Lake
As Figure 1 shows, PGE used the AWS Database Migration Service (DMS) to migrate data from its on-premises Oracle database to Amazon S3. API calls retrieved external data, which was fed into the ETL tool, Matillion, landing in S3. Data from a SQL Server database
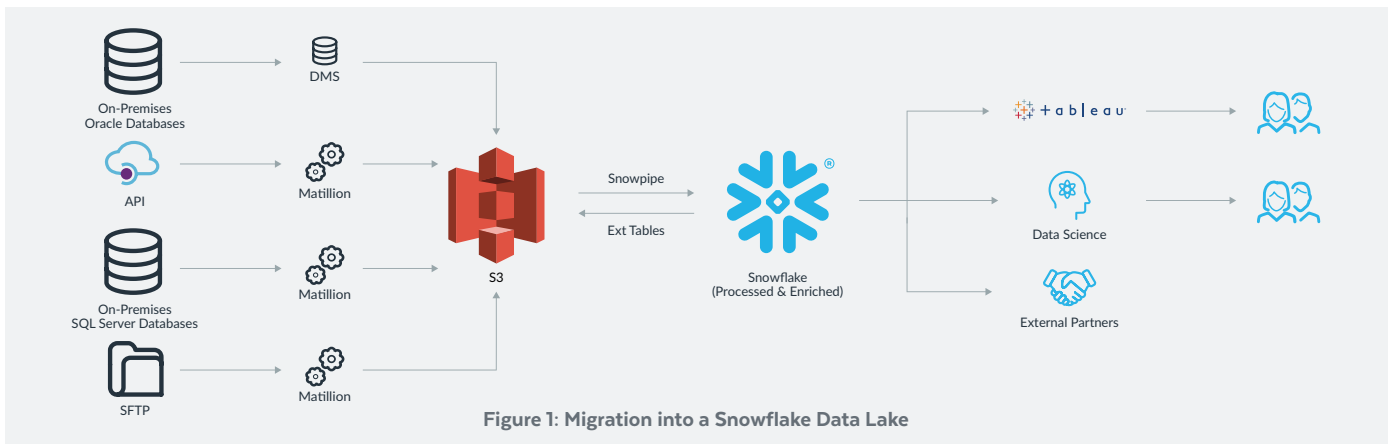
168
Number of manual hours saved by automating customer reporting processes

$500,000
Cost savings from decommissioning a data transfer tool

$1 Million
Expected annual cost savings from moving to the Snowflake Data Cloud

Figure 1: Migration into a Snowflake Data Lake

and from SFTP feeds also used Matillion for ETL into S3. As soon as new files were available in S3, Snowpipe pushed that data into Snowflake. The landing tables in Snowflake are populated in near real time.

Snowflake Secure Data Sharing
Providing data to external partners was a challenge for PGE in the past. Now PGE uses Snowflake Secure Data Sharing to provide data to external groups as well as internal groups such as the data science team.

Data Vault methodology and data mart
PGE implemented the Data Vault methodology to model data for each source system and reduce maintenance. PGE also implemented a data mart. Tableau and related reporting tools plug in to the data mart to enable report creation. Snowflake users have access to both the data mart and the raw landing tables.

> **Users can access both the modeled and raw data independently. We give them access to the data mart to create their own views and reports. It's a great self-service model."**
>
> **—ARAVIND MURUGESH,**
> Principal Data Architect, Portland General Electric

Data security using secure views
Secure views in Snowflake provide row-level and column-level access based on user ID. Users are authenticated via a single sign-on into Tableau. Data requested from Tableau dashboards have access controlled by user-level privileges managed by secure views.

RESULTS:

168 hours of time savings from automating manual tasks
In the past, sharing energy consumption and billing data with customers required a lot of manual effort. Employees exported data into Excel, then emailed results to customers. When PGE migrated to Snowflake, it created an automated process to post data using a web service. Consumption and billing data was posted to customers' accounts automatically. This saved 168 hours of manual labor annually, and it gave customers more immediate access to their data.

$500,000 cost savings in software licenses
Before implementing Snowflake, PGE used commercial software for transferring data to its data science team. After migrating to a Snowflake data lake, the data science team implemented the data transfer in-house. The decommissioning of the data transfer software saved the company a total of $500,000 in software license cost.

$1 million in expected cost savings by the end of 2021
The TCO with Snowflake is far more favorable than with PGE's on-premises warehouse. The company is continuing to migrate more on-premises databases into the Snowflake data lake. By decommissioning software licenses and recovering storage and server costs, PGE expects to save over $1 million by the end of 2021.

FUTURE:

Continued efficiency gains with Snowflake
PGE plans to continue using Snowflake's Secure Data Sharing and scalability and security features to increase its data-driven decision making.