

# OKTA ELEVATES DATA ANALYTICS BY INGESTING 500 MILLION EVENTS PER DAY INTO SNOWFLAKE

## SOFTWARE

okta

**COMPANY** Okta

**LOCATION** San Francisco, California

## SNOWFLAKE WORKLOADS USED



Identity company Okta helps organizations securely connect people and technology. More than 9,400 organizations, including JetBlue and Slack, use the Okta Identity Cloud to manage access and authentication for employees, contractors, partners, and customers. To enable data-driven decision-making across the company, Okta ingests and analyzes large amounts of product configuration and usage data.

## STORY HIGHLIGHTS:

### Multi-cluster shared data architecture with instant scalability

Okta ingests and processes up to 500 million events per day with Snowflake without experiencing latency or resource contention.

### Increased revenue from data-driven insights

Okta's data model converts product configuration and usage into actionable insights for sales.

### Secure, governed access to data

Business users across the organization have secure and governed access to data for data-driven decision-making.

## CHALLENGE:

### Reduced insights because of latency and siloed data

Okta's previous legacy cloud data architecture could not affordably scale to handle up to 500 million events per day from the Okta Identity Cloud. Resource contention led to multi-day data processing delays. Basic event stream queries took minutes to finish running, which negatively impacted data analyst productivity. Large, month-end processes took up to nine hours to finish running.

Customer growth compounded these data engineering challenges and prevented the integration of new data sources, which could have overwhelmed Okta the data architecture. "We wanted to maintain the ability we had, even if it was limited, and not break the system. This meant we could not surface the insights that people were asking for," Okta's Director, Data and Analytics, Lauren Anderson, said.

Lacking a single source of truth, Okta's marketing, product, and finance teams relied on a variety of siloed data sets and manual processes for data analytics. Okta's sales and customer service teams had minimal access to product usage data, which lessened the impact of renewal, upselling, and cross-selling activities.

Time-consuming data warehouse administration diverted attention from increasing data analytics.

**“Snowflake's Data Cloud is our central nervous system that enables data sharing and self-service analytics at Okta.”**

—LAUREN ANDERSON, Director, Data and Analytics, Okta

**500 Million**

Number of events ingested per day

**5x**

Increase in data sources

**15x**

Increase in amount of data stored

## SOLUTION:

### Lower TCO and accelerated analytics with Snowflake

Realizing the need for a modern data environment, Okta turned to Snowflake on AWS. According to Anderson, "Switching to Snowflake was an extremely smooth and quick process, and we experienced no issues whatsoever." Snowflake's multi-cluster shared data architecture scaled instantly to eliminate Okta's resource contention and stored 15 times more data for a lower TCO. Snowflake's near-zero maintenance enabled Okta to redirect DBA resources to focus on more important tasks.

Connecting Snowflake to Apache Airflow and Fivetran enabled Okta's data engineering team to extract data from 17 sources, including Salesforce, Marketo, NetSuite, Workday, Greenhouse, Jira, and ServiceNow. Bulk loading from Amazon S3 and continuous data loading via Snowpipe streamlined data ingestion.

Snowflake views and role-based access control (RBAC) provided a convenient solution for sharing data with granular control. Tableau integration accelerated the development of meaningful reports to support every analytics use case.

## RESULTS:

### Faster data-driven insights supporting more business units

Eliminating multi-hour data processing delays resulted in increased analyst productivity. Near real-time querying of event stream data empowers analysts to provide richer insights in less time. "Architecting our ecosystem on Snowflake allows us to support more business units, create more reports, and report on metrics more frequently," Anderson said.

**“With Snowflake, we experienced a substantial improvement in query time—at least 20%—compared to other solutions that we POC tested.”**

—VICTOR DUAN, Sr. Data Engineer, Okta

### Snowflake as a single source of truth for BI and analysis

Ingesting data from 17 sources into Snowflake provides Okta with a single source of truth for BI reporting and ad hoc analysis. Marketers rely on Tableau dashboards, powered by Snowflake, to achieve a unified view of advertising performance and attribution across all platforms, including Google, LinkedIn, and Facebook. Product teams monitor configuration and usage data to measure feature adoption and guide development decisions.

Automated reports keep Okta's finance team informed about key metrics such as total customer count and net retention rate. Securely sharing financial insights with other groups is also easy. "Snowflake allows us to keep financial data separate and then share data between warehouses, which provides an additional layer of control," Anderson said.

### Increased revenue from data-driven insights

Okta's data model converts product configuration and usage data into actionable insights for understanding customer license compliance. Customer journey data in Snowflake protects renewals by proactively identifying customers who are underutilizing the Okta Identity Cloud. Combining product configuration and usage data with Salesforce data surfaces pipeline opportunities for Okta's sales team.

**“With Snowflake, we're able to identify customers who need additional seats or may be interested in other products, which has resulted in millions of dollars in revenue.”**

—LAUREN ANDERSON, Director, Data and Analytics, Okta

## FUTURE:

### Data science and data lake strategies driven by Snowflake

Connecting Snowflake to DataRobot will accelerate data exploration, simplify hypothesis testing, and enable automated models that surface sales opportunities at scale. "Our goal will be to shorten sales cycles, increase upsells, and ensure the right customers are getting the right products," Anderson said.

Centralizing Okta's data lake strategy on Snowflake is also on the roadmap. "We're already processing a lot of configuration and feature adoption data, which we plan to share to the data lake to power self-service customer reporting," Anderson said.

## ABOUT SNOWFLAKE

Snowflake delivers the Data Cloud—a global network where thousands of organizations mobilize data with near-unlimited scale, concurrency, and performance. Inside the Data Cloud, organizations unite their siloed data, easily discover and securely share governed data, and execute diverse analytic workloads. Wherever data or users live, Snowflake delivers a single and seamless experience across multiple public clouds. Join Snowflake customers, partners, and data providers already taking their businesses to new frontiers in the Data Cloud. [snowflake.com](https://snowflake.com)