snowflake®

THE PRODUCT MANAGER'S GUIDE TO BUILDING DATA APPS ON THE DATA CLOUD

5 COMMON USE CASES FOR BUILDING HIGHLY PERFORMANT SAAS APPLICATIONS

CHAMPION GUIDES

TABLE OF CONTENTS

BUILD YOUR APPLICATION ON THE DATA CLOUD

Data-intensive applications require a modern data architecture to meet customer expectations. That's why fast-growing software companies are turning to cloud data platforms. Although some application builders originally **launched applications on a legacy data stack, many have now migrated to the cloud to overcome growing pains**. Others are building apps on Snowflake's platform from the start because they are aware that Snowflake helps address the biggest challenges in building highly performant data applications.

This guide describes five of the most common use cases for building data applications and how Snowflake addresses the key challenges that builders of such applications face:

- **Customer 360:** Marketing or sales automation that requires a complete view of the customer relationship
- **IoT:** Near real-time analysis of large volumes of time-series data from IoT devices and sensors
- **Application health and security analytics:** Identification of potential security threats and monitoring of application health through analysis of large volumes of log data
- **Machine learning (ML) and data science:** Training and deployment of machine learning models in order to build predictive applications, such as recommendation engines
- **Embedded analytics:** Branded analysis and visualizations delivered within an app

CUSTOMER 360 DATA APPS

Marketing or sales automation data apps require a complete view of the customer relationship to deliver fresh insights at the time of an interaction. All relevant customer data must be analyzed to find profitable customer segments, build effective marketing campaigns, make personalized offers that convert, and contact the right customers at the right time.

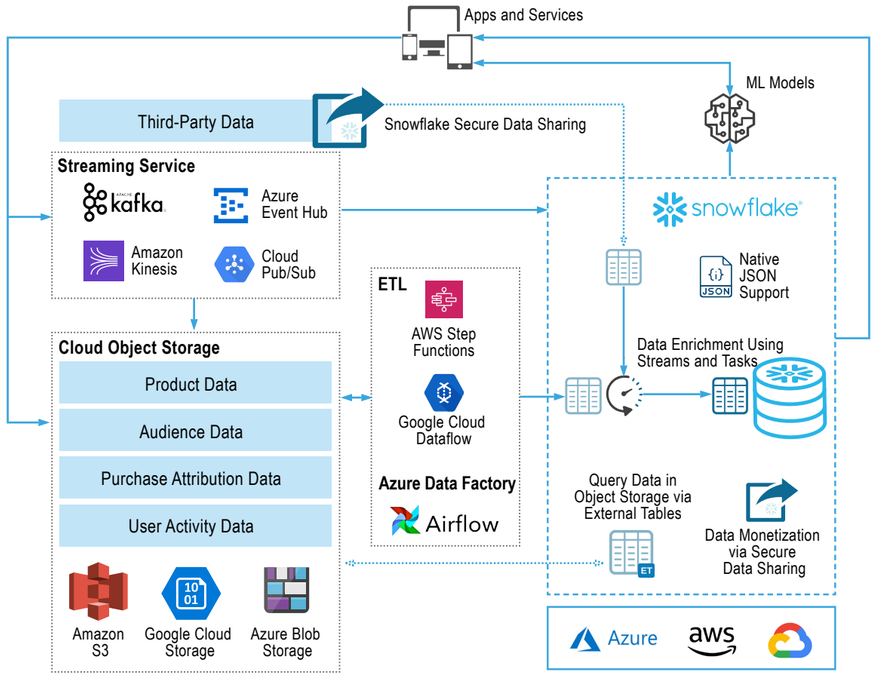
Developing a 360-degree view of the customer is often challenging because customer data is typically fragmented across many systems and consists of different data types. For example, recent customer purchase history is stored in operational databases; in contrast, real-time event data, such as clickstreams, isn't structured and requires complex data pipelines that are difficult to maintain. Intent and demographics data is typically acquired from third parties and requires complex integrations and expensive data duplication, which results in stale data. To make matters worse, each of your customers may source customer data from a different set of systems.

The following Snowflake capabilities meet the needs for building apps that provide a 360-degree view of customers:

- Integration with a streaming infrastructure, such as Kafka, enables reliable ingestion of customer activity, such as clickstreams and browser cookies.
- Native support for semi-structured data simplifies the maintenance of the data pipeline and the analysis of event data with historical data.
- External tables and Snowflake Secure Data Sharing allow access to raw customer external and third-party data for a more complete view of the customer.
- Commodity storage prices enable cost-effective storage for very large amounts of data over longer periods.

BUILD APPS THAT:

- Uncover new, profitable customer microsegments using historical and near real-time data.
- Send targeted emails that engage customers and then analyze user behavior such as opens, clicks, and device usage to optimize email campaigns.
- Deliver cashback offers to consumers in near real time.
- Help merchant partners acquire more customers and increase order size.
- Deliver coupons to consumers shortly after they abandon their shopping carts and personalize coupons based on cart contents, customer history, and behavior.
- Make product recommendations to cross-sell and upsell customers.
- Deliver offers across web, mobile, and contact center channels in near real time.



Customer 360 Reference Architecture. **Download high-resolution image and annotations here.**

IOT DATA APPS

The value of IoT apps is tied directly to the timely delivery of continuous time-series data from large numbers of smart devices and sensors. Although some apps depend on immediate data, such as those that route traffic or detect unsafe situations, other apps rely on the availability of massive amounts of data, such as those that use ML to improve health outcomes or predict maintenance schedules.

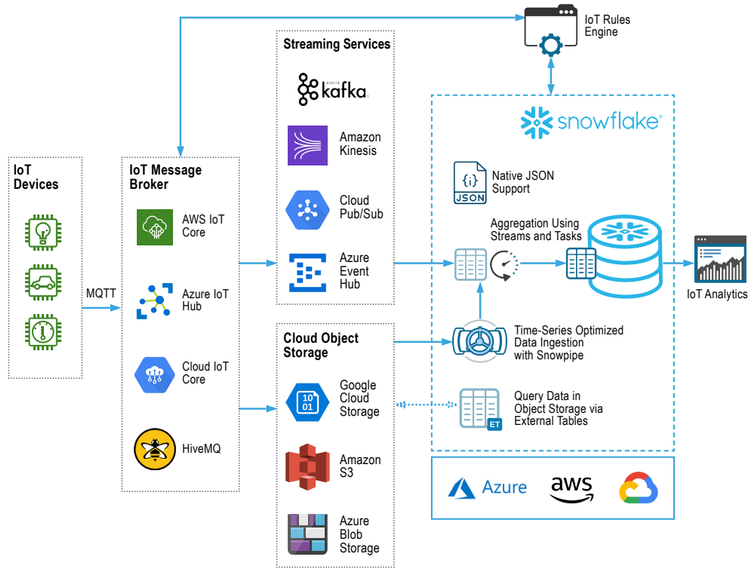
Ingesting, storing, and analyzing large amounts of device data in near real time presents many challenges. The sheer volume of device data often overwhelms traditional data pipelines and is very costly to store for long periods. Because device data is not structured, it requires transformations that disrupt the ingestion process. Additional delays and expenses can also arise because raw device data is often stored in cloud object storage and must be moved to a data warehouse for analysis. Adding to the challenges, data may be transmitted via unreliable internet connectivity, which means it arrives out of order and requires manipulation before it can be analyzed.

The following Snowflake capabilities meet the needs for building IoT apps:

- Integration with a streaming infrastructure, such as Kafka, enables reliable ingestion of device data, such as temperature sensors and fleet telemetry, without delays. Analysis of current sensor data enables apps to control IoT devices in near real time.
- Snowpipe automatically optimizes time-series queries by ingesting data chronologically, simplifying the analysis of out-of-order data, which often results from unreliable device connectivity.
- Streams and tasks automate the workflows required to aggregate incoming data when sensor data is too granular. For example, sensor data can be automatically aggregated into five-minute time increments to simplify its analysis.
- Native support for JSON simplifies the maintenance of the data pipeline and the analysis of event data with historical data. Sensor data can be loaded directly into a Snowflake VARIANT column and easily joined with structured data using standard SQL.
- External tables enable direct access to raw device data in cloud object storage or data lakes without having to move it, saving time and expense.

BUILD APPS THAT:

- Monitor patient health continuously to improve outcomes by analyzing data from fitness bands, glucometers, Holter monitors, and other devices and by delivering insights to patients, providers, and customers.
- Analyze telemetry from fleets such as engine data, speed, acceleration, and location to predict maintenance schedules, route traffic optimally, improve fleet efficiency, and ensure driver safety.
- Analyze shopper traffic patterns and monitor inventory using radio-frequency identification (RFID) sensors to enable retailers to improve cross-selling, optimize shelf space, plan staffing needs, and predict supply to meet anticipated demand.
- Monitor water, electricity, oil, and gas sensors and deliver insights that allow smart cities, utility companies, and consumers to improve energy efficiency and conserve water.
- Monitor and control smart devices such as door locks, security cameras, smoke detectors, lights, thermostats, and sprinklers. Build consumer profiles, personalize devices, and predict anomalies based on device usage and user behavior patterns.



IoT Reference Architecture. **Download high-resolution image and annotations here.**

APPLICATION HEALTH AND SECURITY ANALYTICS DATA APPS

Developers of security and application health analytics apps require access to vast amounts of historical and current data. Historical data helps train ML models to predict anomalies, and current log data and real-time event data is needed to detect threats and respond proactively. In addition, real-time application and infrastructure metrics are needed to help identify application health concerns and take action to prevent outages.

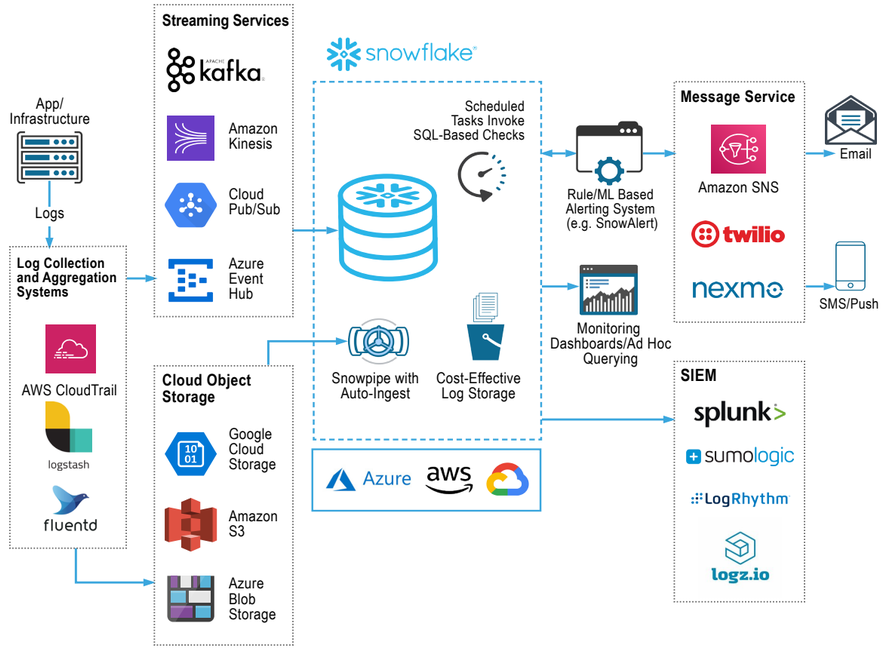
Combining sizable amounts of historical and current data for analysis presents several challenges. First and foremost, storing large volumes of data for long periods is costly. Then there's the problem of ingesting real-time event data, which is difficult to do reliably. And, if you stage data in cloud object storage rather than stream it, the result is delayed time to insight and additional storage costs.

The following Snowflake capabilities meet the needs for building application health and security analytics data apps:

- Commodity storage prices deliver cost-effective storage for very large amounts of data over longer periods.
- Integration with a streaming infrastructure, such as Kafka, enables reliable ingestion of log data from applications servers. Server health indicators such as memory and CPU utilization are ingested without delay and analyzed to detect anomalies in near real time.
- Access to historical data allows developers to train ML models that more accurately detect unwanted anomalies.
- Snowflake's Kafka connector enables reliable streaming of event data so threats can be answered without delay.
- Schedules, tasks, and SnowAlert permit developers to implement rules to detect anomalies and to complement ML models.

BUILD APPS THAT:

- Garner insights from application data and infrastructure metrics, such as CPU utilization, disk I/O, and memory usage. Deliver proactive alerts before outages occur.
- Analyze log and behavior data and train ML models to predict threats. Respond proactively by detecting attacks based on model predictions and suspicious changes in behavior.
- Reduce security operations center (SOC) response times by automating detection and response. Analyze fresh data and deliver alerts via dashboards, SMS, email, and push notifications.
- Become compliant with SOC 2, PCI DSS, HIPAA, and other requirements by identifying compliance misconfigurations, vulnerabilities, anomalies, and hidden threats.
- Analyze logs for open-ended exploration of app health data and potential security vulnerabilities and embed visualization dashboards within DevOps workflows.



Application Health and Security Analytics Reference Architecture. **Download high-resolution image and annotations here.**

MACHINE LEARNING AND DATA SCIENCE APPS

Feature engineering requires data scientists to brainstorm, test, and select the most relevant attributes for their ML models. This process requires extensive experimentation. Once models are deployed, a stream of fresh data is needed to ensure their ongoing accuracy as external conditions change. Models also need to support both batch predictions (for example, an email marketing campaign) and near real-time predictions (for example, the detection of cyberthreats or the delivery of abandoned-cart offers).

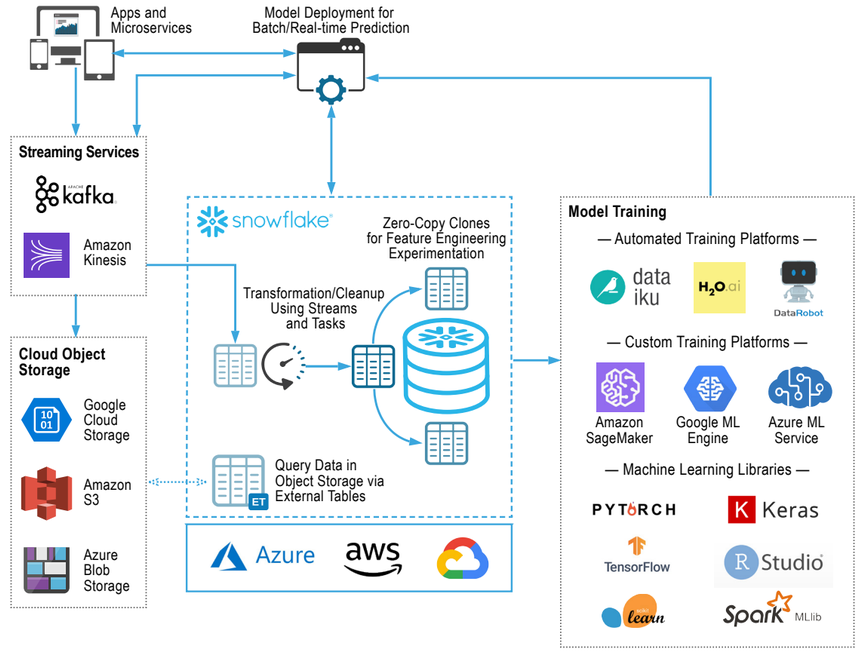
Data scientists face a plethora of challenges to bring together the data needed for modeling. Constructing accurate ML models requires a large volume of clean, historical data and an iterative process of continuous improvement. Copies of entire data sets must be made to support each experiment—a costly and lengthy process that can limit the number of experiments conducted and hurt model accuracy. To support near real-time predictions, application developers need to query fresh data with very fast response times, which may require expensive overprovisioning of compute resources.

The following Snowflake capabilities meet the needs for building ML and data science apps:

- Snowflake Zero-Copy Cloning allows data scientists to instantly create virtual copies of entire data sets for each experiment without the need for data movement, which dramatically simplifies and accelerates feature engineering and model construction.
- Streams and tasks support complex transformations and the cleanup of data pipelines.
- External tables support the querying of raw data in cloud object storage without requiring data ingestion.
- Snowflake's Kafka connector enables reliable ingestion of streaming data.
- A native connector for Python simplifies integrations with training platforms and ML libraries.
- Autoscaling enables fast, near real-time predictions without overprovisioning compute resources, which protects product margins.

BUILD APPS THAT:

- Make product recommendations based on purchase history, browser cookies, and clickstream activity to help retailers increase revenue.
- Deliver cashback, cross-sell, abandoned-cart, and other offers that increase conversion rates based on customer profiles and customer history.
- Predict wind, temperature, precipitation, and other atmospheric conditions based on weather models and sensor readings.
- Build ML models that predict opportunities and schedules for vehicle preventative maintenance based on engine diagnostic telemetry.
- Analyze logs and user activity to identify bad actors, detect malicious behavior, and respond to prevent cyberattacks.



ML and Data Science Reference Architecture. **Download high-resolution image and annotations here.**

EMBEDDED ANALYTICS DATA APPS

Developers who want to embed analytics within their applications must deliver strong user experiences, regardless of the number of users. Customers expect data to be fresh and query-ready 24/7. Even when large queries or ingestion jobs are run in parallel, performance SLAs need to be met.

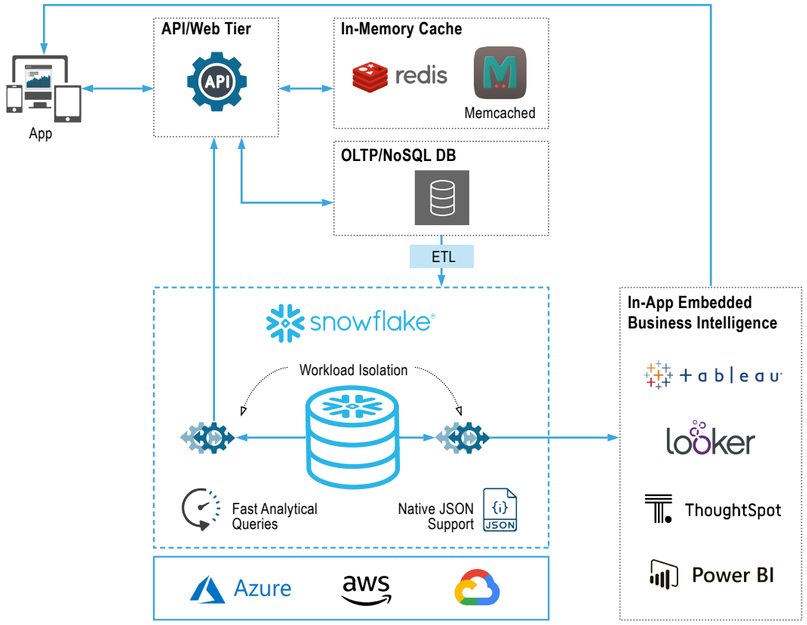
Embedding analytics in data-intensive applications presents key challenges for developers. Performance suffers and concurrency limits are encountered as the number of users or query complexity increases, which results in poor user experiences and possible downtime. To alleviate the performance impact, developers often overprovision compute, which hurts profit margins by paying for idle resources. And ingestion of semi-structured data requires dedicated engineers who can maintain the pipeline to avoid disruptions when schemas change.

The following Snowflake capabilities meet the needs for building embedded analytics data apps:

- Automatic scaling up and down of compute resources delivers consistent performance regardless of load.
- Pay per second pricing improves margins and unit economics by eliminating the need to pay for idle resources.
- Virtual warehouses isolate all workloads to deliver fast queries on demand without contention for resources. There are no limits on the number of concurrent queries that can be executed or on when workloads can run.
- Native support for JSON and other semi-structured formats simplifies the data pipeline without the need for ongoing maintenance, even as schemas change.
- Support for standard SQL enables a broad choice of commercial or open-source BI tools and connectors for Tableau and ThoughtSpot simplify embedding those partner products.

BUILD APPS THAT:

- Deliver insights natively within your branded application and offer rich visualization dashboards without the need to switch applications.
- Enable analysis of app data and insights that are contextual and within the application rather than moving data to a separate environment.
- Create dynamic reports based on live application data without the need to extract data for analysis, which becomes dated quickly.
- Offer more-granular views of your app data and provide flexibility over how data is presented instead of relying on predetermined views.



Embedded Analytics Reference Architecture. **Download high-resolution image and annotations here.**

DELIVER THE MODERN APPLICATIONS YOUR CUSTOMERS WANT

Regardless of the type of data application you're designing, building, or supporting, the best way to differentiate your offering is to select the right underlying architecture. After all, modern data applications that deliver real-time value at massive scale require a modern data platform that's designed for highly performant customer experiences.

With near-unlimited and automatic scalability, concurrency, instant elasticity, and support for structured, semi-structured, and unstructured data, Snowflake provides exactly what your application needs. Data applications never experience resource contention, and every customer is ensured continuous access to data and analysis.

Snowflake's fully managed data platform also removes the operational burden from engineering teams and enables them to focus on improving applications and building new features.

If you're ready for faster innovation, improved engineering efficiency, and a data platform that supports your long-term product vision, **try Snowflake**.

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. Snowflake's platform is the engine that powers and provides access to the Data Cloud, creating a solution for data warehousing, data lakes, data engineering, data science, data application development, and data sharing. Join Snowflake customers, partners, and data providers already taking their businesses to new frontiers in the Data Cloud. **Snowflake.com.**