snowflake®

7 BEST PRACTICES FOR OPTIMIZING YOUR SNOWFLAKE INVESTMENT

CHAMPION GUIDES

EBOOK

TABLE OF CONTENTS

INTRODUCTION

Companies across nearly all industries are striving to digitize their businesses. To do so, they are likely implementing at least one, if not multiple, cloud-based solutions. This also means they have more people working across these systems, provisioning resources, running queries, performing analyses, and implementing new tasks. Regardless of what systems are in use, these companies are creating and accumulating more data than ever before—and their storage needs are expanding at an exponential rate.

All of these factors can contribute to ballooning computing and data storage costs—that is, if the right monitoring and control mechanisms are not put in place.

The problem many companies face, however, is their existing data platform's capacity. Many legacy solutions are being strained to capacity, and they lack some of the basic tools and visibility required to achieve modern cost governance.

For instance, many existing tools are inherently inflexible with fixed cost pricing, making it challenging to implement cost-saving changes. Even when these products do allow for some level of visibility and governance (perhaps because they were designed for a different era with different levels of need in mind), they require complex capacity planning. As a result, **the time it takes to ensure a company is keeping an eye on its data costs can end up proving costly itself.**

As today's cloud-first transformation movement takes hold, cost governance has become a paramount concern. Thankfully, more modern data platforms such as Snowflake are built to help maximize cost transparency and user control. In particular, Snowflake is designed for customers and industries where data collection and usage fluctuates to a large degree. The platform enables flexible scaling and pay-as-you-go pricing, so decision-makers can ensure they pay only for the capacity and usage they need at any given moment. In this way, companies save time and money by eliminating the need to overprovision.

HOW SNOWFLAKE CALCULATES COSTS

Historically, many data cloud companies relied on utilization forecasts and charge-back models to determine costs. Snowflake takes a fundamentally different approach, at the core of which is that pricing is aligned with customers' usage needs, and therefore customers only pay for what they use.

There are several unique elements to Snowflake's pricing model. It starts with credits for compute capacity. This capacity can be scaled up or down as needed, and customers pay based on their consumption, not unlike a utility bill. In addition to this scalability, Snowflake's product is built around the ability to optimize performance while managing costs. Thus, compute resources, called virtual warehouses, are elastic, meaning they can be turned on or off as needed.

A marked advantage Snowflake offers here is that it automatically compresses data by a factor of 3 to 5, which can reduce costs significantly compared to products that charge for raw data storage. In addition, Snowflake helps clients manage data usage and storage through Snowflake Zero-Copy Cloning, enabling employees to access the same live data sets without duplication.

Lastly, Snowflake enables a public cloud infrastructure. This provides customers more choice and even more direct control over potential costs, regardless of which partners they employ.

HOW SNOWFLAKE ENABLES COST GOVERNANCE

As much as Snowflake provides the tools and systems that empower companies to take control of cost management, the practice of cost governance is not a "set it and forget it" proposition. In cost-sensitive organizations, companies often designate a watchdog inside their organization to make cost governance their primary responsibility. That person or team should be tasked with fully understanding the company's usage needs and optimizing usage over time while remaining in control of provisioning resources.

Snowflake features three essential attributes to bolster this form of active cost governance:

- **Visibility:** Companies can fully understand their spending and attribute it accurately.
- **Control:** Companies can set limits and take actions.
- **Optimization:** Companies can identify inefficient spend, and reallocate dollars to where they will most impact performance.

For example, in 2021 **Western Union** was able to achieve resource optimizations of 50% by taking advantage of Snowflake's offerings. Similarly, **Cisco** found it was able to reduce its costs by 15% through optimization.

Still, even as companies such as Cisco and Western Union have found success in getting a handle on cost governance, the challenge is likely to persist as cloud adoption becomes more ubiquitous. As Gartner reported in March 2022, many new cloud adopters' employee bases simply don't understand the factors driving costs of the products they use. That, coupled with many longtime cloud users' tendency to use shortcuts to access cloud resources, are both likely to drive up costs over time.

That report's overall recommendation very much aligns with Snowflake's—that companies need to prioritize and take a proactive approach to cost governance in the cloud, and acknowledge that not every product in the market is set up to enable that. As Gartner recommended, "Explore product-specific financial controls and guardrails from your cloud provider to gain finer control over cloud spending and cost overruns."

Of course, the path to understanding and optimizing costs and performance is likely to look different for each company—and may be a moving target over time. Still, some fundamental lessons apply to every business. To help set you on the right path, here are **seven best practices for optimizing your Snowflake investment:**

1) MAKE SURE YOU CAN SEE—AND CONTROL—USAGE ACROSS THE BUSINESS

The combination of access and control requires architecting an overall account structure for the cost governance for the company. Under that scenario, leaders and account managers should be able to not only get straightforward reporting and analytics-ready usage data, but also should be equipped to use that information to make decisions on what to turn on and off. To be most effective, the more granular usage and cost data the better. Ideally, you'll want to be able to create custom reports that allow your team to drill down and identify areas where usage and spending can be reigned in.

2) LEVERAGE AUTOMATION

Snowflake features a number of tools designed to help companies keep tabs on costs while reducing the need for clients to manually monitor or search for savings. For example, by using Resource Monitors, Snowflake administrators can opt in to receive automatic alerts on resource consumption by setting up specific credit quotas. Once these quotas are reached, account managers can receive notifications, or the account can be automatically suspended if needed. If that approach is too blunt, administrators can stagger and escalate these suspension triggers as different data sources or warehouses across the company near predetermined quotas at various intervals.

3) SHUT THINGS DOWN WHEN WORK STOPS

In addition to Resource Monitors, companies can enable the **Auto-Suspend and Auto-Resume functions**. This way, when compute resources are not being used, organizations won't be charged for idle time once a certain period of inactivity is reached. Along those same lines, Auto-Resume allows for companies to restart a warehouse as soon as new activity such as a query kicks in. This strategy helps companies reduce waste without having to compromise performance by asking users to "wait" for warehouses to start up once again.

4) TAKE A TIMEOUT WHEN NEEDED

Another way to ensure that downtime doesn't result in higher costs is to set in place "timeouts" designed to stop work on any long-running queries that exceed preset thresholds. Keep in mind that it's also important to make sure customers are alerted when such queries time out.

5) PURGE REGULARLY

Yet another means of ensuring that companies aren't paying for time or usage unnecessarily is to conduct periodic housekeeping. Start by purging dormant users, such as employees who have gone long periods without using the platform, or people who have simply never logged in. Similarly, it's a good idea to dump unused tables, so you don't pay to store objects that have long since been abandoned or forgotten.

6) SPOT DEVIANT PERFORMANCE AND ELIMINATE IT

Here's where the data visualizations and dashboards available in **Snowsight** can come in handy. Companies can examine seven-day-average usage patterns across various tables and virtual warehouses and look for anomalies. Once anomalies have been identified, they can be investigated and addressed accordingly.

7) ANALYZE USAGE

Ultimately, the best way for companies to govern data costs is to continually perform comprehensive analyses of their usage. As much as your teams might plan for a certain level of production and reliance on a platform over time, in the real world, usage can vary widely from day to day and even from hour to hour.

To date, legacy data warehouses have provided neither the transparency nor the elasticity needed for companies to get a handle on what they are using and what they are paying for. With Snowflake, the structure and tools are in place to deliver the resources needed to all users while removing wasteful spending.

CONCLUSION

The best way for companies to start taking better control of their cost governance is to examine their usage patterns and identify any gaps or inefficiencies. That's where Snowflake's products can pay immediate dividends, from automating usage monitoring to enabling preset triggers to ensure costs stay manageable as usage grows. To get started analyzing your own company's usage patterns, we recommend trying this **Resource Optimization Quick Start Guide.**

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**