snowflake®

SIX COMPELLING REASONS TO MODERNIZE YOUR DATA WAREHOUSE

Start extracting maximum value from your data

CHAMPION GUIDES

TABLE OF CONTENTS

HOW MUCH IS YOUR DATA WORTH?

Today, data can be your most valuable asset. But that's only the case if your organization can extract deep business insights from it. Like crude oil, your data is much more valuable when processed and used to drive your business forward.

This is where conventional data warehouses come up short. They can't handle the volume, complexity, and variety of data companies generate today. Nor can they simultaneously satisfy various departments' need to access and analyze data in real time.

A modern, cloud-built "as a service" data warehouse enables you to do things you just can't do with a legacy on-premises or "cloud-washed" data warehouse, including:

- Perform advanced data analytics across all user profiles and contextual data sources
- Integrate data management and data storage, including governance and quality
- Enable your systems to scale automatically without administrative intervention
- Avoid partitioning hassles and downtime
- Eliminate backup scripts and many of your previous data retention processes
- Forget about tuning and spending time on low level maintenance tasks
- Avoid the need to manage underlying infrastructure

At its most basic, a modern data warehouse must enable your organization to extract maximum business insights and value from all its data. At the same time, it must provide that value in a way that minimizes system configuration costs, ongoing maintenance chores, and continual tweaking to improve performance. The bottom line is that everyone in your organization needs to be able extract value from all your data in a timely and easy-to-use fashion.

WHY CHANGE IF YOUR DATA WAREHOUSE IS WORKING?

Even if you realize there are benefits to modernizing your data warehouse, you might be concerned about moving to a new solution. Perhaps you've already spent time and budget on an on-premises system, and it is working as intended. Why change what isn't broken?

Historically, data warehouses were designed as a central repository for data gathered from multiple sources (OLTP) and blended with a wide variety of business analytics (OLAP). Functional teams such as Marketing, Sales, Product Development, and Finance, as well as users who held technical, line-of-business, or executive roles could extract actionable insight from warehouse-resident data. But infrastructure-wise, even the data warehouses of a few years ago no longer meet the needs of today's organizations.

New data warehouse technology provides a means to use more types of data and data sources. These include structured and semi-structured data, and data sources such as enterprise apps, mobile apps, web apps, and APIs. Tying together all of your information in a uniform way exponentially increases the insight you can gain, taking you far beyond rudimentary analytics. Snowflake, the data warehouse built for the cloud, enables you to derive maximum benefit from your data.

On the pages that follow, we offer six reasons to modernize the way you use your data.

REASON 1: SPEND TIME MAKING THE MOST OF YOUR DATA, NOT MAINTAINING YOUR DATA WAREHOUSE

The ability to focus higher in the value chain on data-driven insights and strategies is the hallmark of a modern data warehouse. Snowflake enables you to concentrate on using your data in a more business-centric fashion instead of worrying about manually creating and maintaining your warehouse infrastructure. Such a transformation changes your priorities from performing systems maintenance to achieving greater organizational success.

A modern data warehouse such as Snowflake delivers agility and economy, and it provides the ability to scale up or down based on need, without manual intervention. It brings processing capabilities online and offline automatically, maximizing the use of costly resources—both computing and human. Any organization, regardless of its size, products, or services, can leverage these benefits.

	Task	Ordinary DW On-premises	Ordinary DW Cloud	Modern DW
Infrastructure	Hardware	You	Cloud vendor	Built-in
	Software	Distribution	Cloud vendor	Built-in
	Hardware cluster setup	You	You	Built-in
	Software provisioning	Tools	Tools	Built-in
Data & service protection	Data protection & retention	Platform or Add-on	Platform or Add-on	Data Warehouse As-a-Service
	Node failure protection/recovery	You	You	
	Disaster recovery	You	You	
	Service monitoring & alerting	You	You	
Security	Physical security	You	Cloud vendor	
	Deployment security	You	You	
	Security monitoring	You	You	
Database management & tuning	Compute scaling	You	You	
	Index management	You	You	
	Data partitioning	You	You	
	Metadata & statistics management	You	You	
	Query optimization	You	You	

Figure 1: A cloud-built data warehouse such as Snowflake eliminates manual infrastructure tasks required by traditional cloud or on-premises solutions, allowing you to focus on cost-effectively extracting maximum value from your data.

REASON 2:
FREE YOUR DATA TEAM AND YOUR DATA USERS

The ongoing need for users to run specific data queries has been a bottleneck. That's because regardless of who requested the query, whether a line-of-business owner or a top executive, and regardless of the location of the data, whether on premises or in the cloud, these queries required an intermediary data team to provide specific data views and skews. This requirement hindered the ability to make immediate and well-informed decisions, lengthened time to market for new offerings, and delayed time to value for customers.

A modern data warehouse such as Snowflake increases the speed at which your teams can access a wide swath of data and associated analytics.

Snowflake's unique architecture makes it possible to run multiple workloads concurrently without impacting performance.

In addition, Snowflake empowers individual users to perform their own targeted, self-service data manipulation, regardless of their skill level, because there is no need to write scripts or ask the data team for assistance. They can accomplish this task quickly and simply by using familiar languages, such as SQL. You won't completely eliminate your data team, but you will free the team to focus on priority tasks such as setting and enforcing data access policies.

REASON 3:
EASILY INTEGRATE TOOLS THAT ENSURE HIGH DATA QUALITY

High data quality is a fundamental requirement for planning and strategy creation. That's true whether your data warehouse infrastructure is legacy-based or cutting edge. But with a modern, cloud-built data warehouse, high data quality is even more important because more data is used to inform critical decisions.

You must ensure the quality of your data to ensure it's trustworthy. The old adage of "garbage in, garbage out" is still true, even with a modern data warehouse. Given the wider scope and user base of a modern data warehouse, two considerations are paramount:

- **You need data governance to determine what data is truly important to your organization, and to enable you to rely on your individual business units and subject matter experts to prioritize and validate data.**
- **You need the ability to clean your data to ensure everything that goes into your data warehouse is usable, accurate, and delivers true business insight.**

Tools can assist with these tasks. Snowflake closely partners with multiple specialized vendors (for example, Talend and Informatica) to ensure data import tasks are performed simply and with the highest degree of oversight and accuracy. The goal is to ensure you access the widest variety of data possible, ensure data meets your highest organizational standards, and take best advantage of Snowflake capabilities and other downstream processing activities.

REASON 4: EXPAND THE DEFINITION OF PERFORMANCE

As you begin to transition to a modern data warehouse, it's important to think about system performance a little differently, particularly if you're considering Snowflake as your data warehouse provider.

With traditional data warehouses, everyone had to build, manage, and tune clusters. As a result, organizations did not factor in systems management overhead into performance benchmark calculations. However, times have changed. With Snowflake, building and managing clusters is not a requirement, which fundamentally changes the way you evaluate and compare a traditional on-premises or cloud-washed data warehouse to an as-a-service cloud-built system.

Consider a proof of concept between two data warehouses, one a legacy system and one an "as a service" system. At face value, the former delivers a one-second response time and the latter a two-second response time. That seems to indicate a 50% better level of performance from the legacy system and an easy differentiator for selecting a solution. But does this metric truly represent a legitimate comparison that takes all variables into consideration?

The legacy system requires people to boost performance, scale capacity, add nodes, perform reclustering, and manage data. If, for example, the system needs three more nodes to achieve the desired response time, and adding the nodes takes three minutes of administrative time, the total required time is actually 541 seconds. With a modern cloud-built data warehouse, it might take only one second to allocate the extra resources, because the as-a-service data warehouse does it automatically, resulting in a total response time of three seconds. The bottom line is that you need to include management time in any comparison of manual versus automatic data warehouse implementations.

Comparison response times for legacy (t1) versus modern data warehouse (t2):

$$Query_{t1} + Management_{t1}: 1s + (3m \times 3\ nodes) = 541s$$

180X more time!

$$Query_{t2} + Management_{t2}: 2s + 1s = 3s$$

REASON 5: FOSTER BETTER COLLABORATION

As discussed earlier, data governance and high data quality are critical needs from a modern data warehouse. These are not just technology issues, although the IT team is heavily involved in ensuring these needs are met. The business people who know the data best must be involved in its governance and quality assurance to ensure the most accurate analytical results. However, the IT and business teams tend to be operationally divided by an invisible, yet impenetrable wall. Those who manage the data and input it into the system (the IT team) typically don't delve into the needs of those who use the data (the business teams), and vice versa. Most teams acknowledge that cooperation is beneficial, but cooperation is not frequently a part of the total process.

The solution to this dilemma requires the integration of the data management functions across all constituencies, including IT and business teams, to facilitate collaboration and ensure that each team has the necessary tools to accomplish mutually dependent tasks.

With a modern data warehouse, you can:

- **Provide controls for data quality, including the ability to correct and change data and ensure it's trustworthy.**
- **Enable master data management, including the ability to identify the unique characteristics of a persona or product using the data.**
- **Perform metadata management, including the ability to use "data about the data" to ensure that data is applied in the right context.**
- **Deliver self-service apps that foster seamless cooperation between teams via cloud enablement technology.**

REASON 6:
REAP THE BENEFITS OF CONCURRENT DATA ACCESS

One constraint of legacy data warehouses is the need to queue up a variety of jobs to ensure both prime-time system availability and high performance, while simultaneously addressing less time-sensitive needs. This means "monster jobs" are scheduled at night and operations such as data loading take place in off hours to ensure they don't affect the business intelligence (BI) environment during normal working hours. But most everyone wants their data faster for near-immediate evaluation and decision-making. That reality makes a "queue and delay" implementation suboptimal for virtually every party involved.

The modern cloud data warehouse eliminates these scheduling requirements. With Snowflake, you can allocate a variety of different workloads to different data warehouses, even if the workloads overlap. This is especially important for industries such as finance and retail, where analytics information must be immediately available. This means:

- You can easily create an unlimited number of data warehouses (independent compute clusters)
- All teams can share the same accurate and consistent data
- The system suffers no degradation in performance or availability
- Users can immediately query live views of the data
- Users can complete their jobs using the right information precisely when it's needed

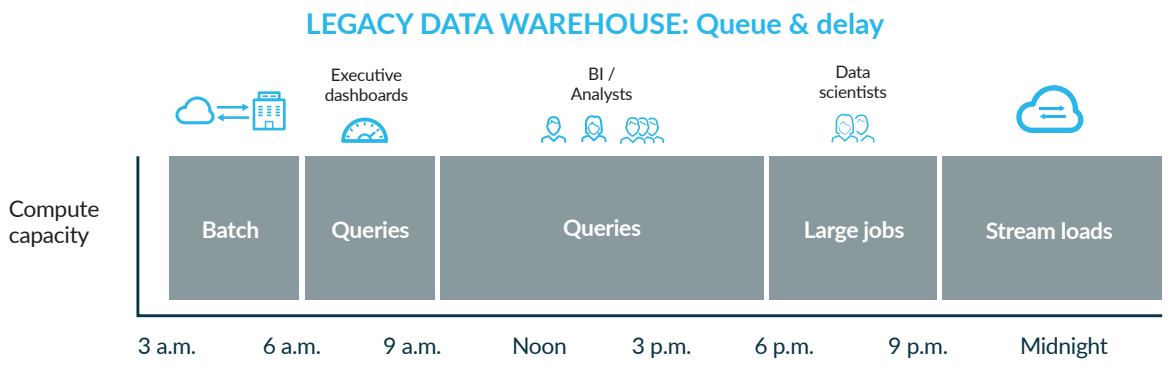


Figure 2: In an ordinary data warehouse environment, batching from a corporate system takes place in off hours, dashboards and BI resources are active during normal business hours, and larger jobs run at night. Data loading also occurs during off hours to ensure it doesn't affect production queries. This "queue and delay" process slows data access for a variety of critical stakeholders.

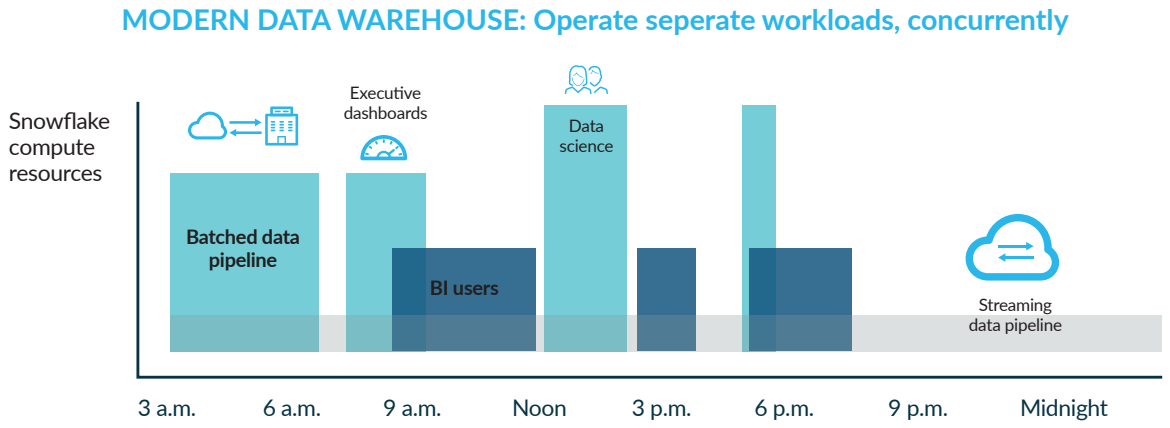


Figure 3: A modern, automated data warehouse can support a variety of different workloads and you can allocate them to separate warehouses. The system can process these workloads concurrently without draining resources or degrading performance, so decision-critical information is delivered without delay.

CASE STUDY: PDX CREATES NEXT GENERATION OF PHARMACY SERVICES USING SNOWFLAKE

PDX is a leading healthcare technology provider to pharmacies and pharmacy chains across the United States. Its products and services enable pharmacies to process prescriptions, store and protect healthcare records, improve patient outcomes, and remain compliant and competitive in a highly regulated industry. PDX wanted to better monetize its systems while providing customers with optimal patient profiles.

PDX chose Snowflake to implement a modern "as a service" data warehouse that incorporates the best practices discussed in this eBook. The new system provides many benefits the legacy PDX system simply could not deliver, including:

- Eliminating infrastructure acquisition costs, management investment, and manual tuning, which means data resources can be scaled dynamically without added hardware or staff.
- Extracting maximum value from its data by offering its clients customer-focused products that address their needs.
- Viewing and analyzing sales data across multiple pharmacies, even across pharmacy chains, in a simple and comprehensive fashion.
- Processing billions of records in minutes without the computational lags that occurred with its legacy system.
- Enabling customers to generate reports without affecting its production systems, and process daily BI workloads concurrently with data loading tasks.

"We feel very confident that whatever we run into, we can scale the Snowflake solution to meet the performance requirements of our pharmacy customers."

John Foss
Director of Business Intelligence & Manufacturer Reporting,
PDX

[9] Harvard Business Review Analytic Services. An Inflection Point for the Data-Driven Enterprise. https://www.snowflake.com/resource/an-inflection-point-for-the-data-driven-enterprise/

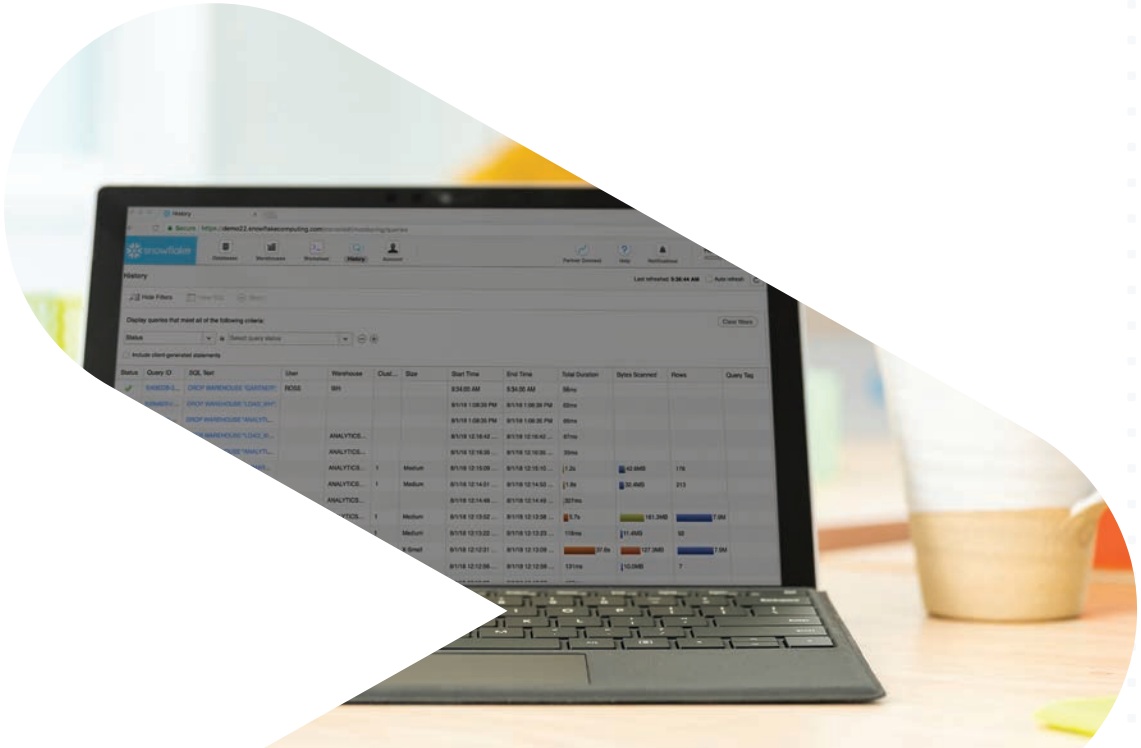
THE ENTERPRISE-WIDE IMPACT OF YOUR DATA WAREHOUSE

With a modern data warehouse such as Snowflake, you can eliminate low-level data warehouse infrastructure tasks, which saves time and reduces overall system complexity and cost. You gain system performance and availability improvements without tedious manual intervention and overhead.

In addition, the system offers your teams a broader view of what they can accomplish enterprise-wide, while retaining a focus on data quality, encouraging improved collaboration across functional teams, and incorporating an emphasis on self-service data access.

And, most importantly, you'll elevate your data to a place in the organizational value chain where it delivers actionable results through more thorough and creative data interpretation and use, in ways you might not currently imagine. With enhanced data strategies, you can conceive and deliver new products and services by relying on accurate, up-to-date, and immediately available information that propels your teams forward in their roles, from the technology team, to business managers, and all the way to the executive suite.

It's time to upgrade your data warehouse and dramatically increase the associated enterprise-wide impact with a modern, "as a service" cloud-built implementation.



ABOUT SNOWFLAKE

Snowflake is the only data warehouse built for the cloud, enabling the data-driven enterprise with instant elasticity, secure data sharing and per-second pricing, across multiple clouds. Snowflake combines the power of data warehousing, the flexibility of big data platforms and the elasticity of the cloud at a fraction of the cost of traditional solutions. Snowflake: Your data, no limits. Find out more at **snowflake.com**