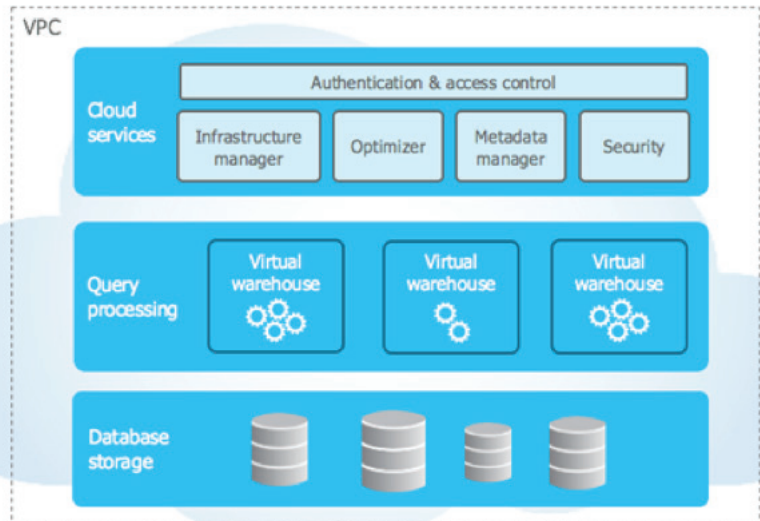


THE BENEFITS OF SNOWFLAKE OVER OTHER DATA WAREHOUSE PROVIDERS

Snowflake's unique architecture is primarily what sets it apart and allows it to compete with the technological giants of Google and Amazon. There are 3 layers to the Snowflake data warehouse architecture: the Data Storage Layer (storage layer), the Virtual Warehouse Layer (compute layer) and the Cloud Services Layer (services layer). The importance of this is that the storage and compute layers are decoupled, that is, they are entirely independent of one another. This allows the compute layer to scale as much as necessary, without friction involved with the storage layer. Although Bigquery has a similar separation of storage and compute, Redshift does not.



This architecture is effective because it was specifically designed by Snowflake for a Cloud environment. Before the advent of cloud data providers such as Redshift and Bigquery, data warehouses were on premise for companies. These on-premise data warehouses required a complex system of servers and other infrastructure elements maintained by data engineers and data architects in order for companies to be able to operate on the data. This old architecture was carried over to providers such as Redshift, which still requires much maintenance on the users' part. However, Snowflake completely changed course and developed this aforementioned architecture which uses storage and compute resources from the cloud to enable many advantages which will be covered in the following sections.

Snowflake is extremely versatile in its ability to scale up and out. Snowflake is designed to allow many users to analyze large amounts of data from the same data source at the same time. To do so, Snowflake must offer a way for users to execute many queries and complex queries at once. The problem of executing increasingly complex queries can be resolved by increasing the size of a compute cluster (scaling up). In order to execute many queries at the same time – an issue known as concurrency – one could use multiple compute clusters, each one working on a different query (scaling out). These reconfigurations can take several hours for providers such as Redshift. This is especially problematic if a customer's workload size fluctuates over the course of a short period of time: Redshift users must either always keep their compute clusters at the maximum size, which takes up many resources, or face the cost of several hours of waiting time. For Snowflake users, they can scale up or out instantly and even while queries are running, through the use of Snowflake's virtual warehouses.



In 2019, Gigaom, a technology blog, released a report on a series of tests to determine the performance of metrics of several data warehouse providers. This series of tests were “compliant with the standards set out by the TPC Benchmark™ DS (TPC-DS) specifications” (McKnight, 2019) and consisted of a series of complex queries designed to simulate a real business use case. In this test, McKnight found that Snowflake outperformed both Redshift and Bigquery in terms of speed for a number of queries, such as Query 14a (a complex and long running query which performs a lookup of sales by items over several tables), Query 80 (reporting on sales and their means)

and Query 94 which calculates order counts. This is a good indicator that any business which is looking to perform complex business intelligence queries should rely on Snowflake for fast query times. This same conclusion was reached by iSmile Technologies, who reported that “In a head-to-head test, Snowflake edged out BigQuery in terms of raw speed, with queries taking, on average, 10.74 seconds (geometric mean). Meanwhile, BigQuery clocked in at 14.32 seconds per query, on average” (Devin, 2021).





Snowflake is a cross-region and cross-cloud platform. Snowflake supports cloud regions all over the world, in order to offer businesses from around the globe access to their services. This is not something new, as both Redshift and Bigquery have clouds over the world, from their creators (Amazon and Google, respectively). Snowflake does compete with these two companies, nevertheless, it is also an active partner with them and hosts its own services on their clouds. This means that Snowflake supports data on Amazon, Google and even Microsoft Azure Clouds, while each of these individual cloud providers are limited to their own clouds. This means that companies that are looking to transition to Snowflake need not worry about which cloud providers they are already using and whether they are compatible with their other data sources or not as they are all supported by Snowflake.



Conclusion

In conclusion, Snowflake is an excellent cloud data warehouse provider with lots of potential to aid customers in their business needs. With a modern architecture, excellent scaling capabilities, the ability to outperform its competitors in most cases and a cloud-agnostic availability, Snowflake is perhaps the most effective data warehouse provider.

Get in touch with our experts at KCM Solutions to find out more about Snowflake.



K C M S O L U T I O N S

ABOUT KCM SOLUTIONS

Most organizations struggle to leverage data and unlock the insights they need to make informed business decisions. At KCM Solutions, we help you harness the power of your data, giving you control of all the information you need to make confident data-driven decisions.

Get in touch with our experts and get started on your data driven journey today.

Contact Information:

Email: info@kcm-solutions.com

Canada: (647) 361-1701

US: (661) 615-2122

