



Delivering Data Warehousing as a Cloud Service

People need access to data-driven insights, faster than ever before. But, current data warehousing technology seems designed to maximize roadblocks rather than insight.

That's certainly the case for conventional data warehouse solutions, which are so complex and inflexible that they require their own teams of specialists to plan, deploy, manage, and tune them. By the time the specialists have finished, it's nearly impossible for the actual users to figure out how to get access to the data they need.

The cloud offers some possible solutions for these problems, but many 'cloud data warehouse' offerings are simply legacy data warehouse appliances thrown into the cloud, with the same complexity and inflexibility. Other cloud data stores lack fundamental data warehouse features, like full standard SQL support and transactional flexibility. This makes them challenging to use as an enterprise data warehouse.

'Big data' solutions aren't any more helpful at delivering insight. They require new skills and often new tools as well, making them dependent on hard-to-find operations and data science experts.

The cloud offers the opportunity to create a true SaaS data warehouse that can eliminate many of the problems of existing technology, while retaining the familiar tools of a true data warehouse. Unfortunately, many "cloud data warehouse" offerings stop short of this goal.

A SOFTWARE SERVICE IS MORE THAN JUST CLOUD

The rising popularity of cloud infrastructure has led to a small number of 'cloud data warehouse' offerings. However, these offerings are quite different from a data warehousing service.

Many 'cloud data warehouses' are simply traditional data warehouses hosted in the cloud. Paying a cloud service provider to host the infrastructure removes some burden, but the data warehouse software still needs just as much management and tuning. Even in the cloud, it's just as complex and inflexible as it was

on-premises. It's like the difference between buying and leasing a car (if buying were analogous to on-premises deployments): leasing has its benefits, but it still leaves you responsible for refueling the car, maintaining it, driving it, and making sure it's working properly.

Delivering software as a service requires a more fundamental shift. Staying with the car analogy, SaaS would operate more like a car service than a car purchase. With a car service, you don't spend any time worrying about the car or the driving—you simply tell the service when and where you need to travel, and the service takes responsibility for everything required to get you there.

A data warehouse as a service —a true cloud data warehouse—would seamlessly deliver the benefits of a true SaaS solution, including:

- **Instant availability.** The application is available for use within hours or even seconds of when a customer purchases it, rather than in weeks or months.
- **No infrastructure.** Hardware and software are deployed, configured, and managed by the application provider as part of the service.
- **Minimal management.** Users do not spend time worrying about how to patch, upgrade, scale, and optimize the software—the software service does that automatically and transparently.
- **Security and resiliency.** No need to manage physical security, deployment security or monitoring, as this is all a part of the service. Data is secured and replicated for safety across multiple locations.

REQUIREMENTS FOR DATA WAREHOUSING AS A SERVICE

Transforming data warehousing into a cloud service requires re-imagining the entire technology infrastructure that supports it. A true service allows you to shift away from managing, monitoring, and tuning the infrastructure for your database, and instead focus on obtaining value from your data. To do that, data warehousing as a service will need to be:

- **Ready to go at any time.** There's no setup overhead and no need to spend time procuring and installing hardware, installing and configuring software, or any other prep work.
- **Easy to scale.** As the size of individual use cases expands, and the number of use cases grows, a Data Warehouse as a Service must be able to fluidly respond and scale.
- **Seamlessly supportive of any data.** All types of data, including semi-structured formats like JSON, are easy to support. Instead of users spending time transforming data to get it into a form that the database can handle, the database adapts to the data that's loaded in it.

- **Focused on enablement rather than management.** There's no ongoing management of the system, data protection, or security. The service manages itself. Users can focus on asking questions of their data.
- **Dynamically adaptable.** The service monitors and observes the data warehouse and adapts, identifying and making optimizations based on how the service is being used.
- **Full and transparent data warehouse.** A data warehouse as a service still needs to be a true data warehouse. It should have full support for standard SQL, unlimited joins, as well as seamless updates and deletes.

THE SNOWFLAKE ELASTIC DATA WAREHOUSE™ SERVICE

Snowflake has built a new data warehouse from the ground up as a software service. Because Snowflake started with an innovative new architecture, we were able to reimagine the data warehouse and deliver a true service, rather than a repackaging of existing technology in a cloud wrapper.

The Snowflake Elastic Data Warehouse is built to eliminate hardware and software infrastructure management, along with the painstaking tuning and security considerations of existing technologies.

Infrastructure

Snowflake takes full advantage of the cloud to automatically provision and manage the database infrastructure so that you don't need to. Upgrades are performed automatically over time, with no interruption of service or manual intervention. Unlike traditional data warehouses moved to the cloud, you don't need to worry about how to adapt as your usage evolves. Because of Snowflake's unique multidimensional elasticity, you can scale up and down on-the-fly to support whatever scale of data, compute, or users you need.

Comparing traditional databases to DWaaS ● customer ● vendor

MANAGEMENT TASKS		ON-PREMISES	DATA WAREHOUSE IN THE CLOUD	DATA WAREHOUSE AS A SERVICE
Infrastructure	Datcenter operations	●	●	●
	Hardware deployment	●	●	●
	Upgrades	●	●	●
	Scaling	●	●	●
Data management	Data partitioning	●	●	●
	File management	●	●	●
	Data protection	●	●	●
Data management & tuning	Index management	●	●	●
	Sort keys	●	●	●
	Metadata maintenance	●	●	●
	Query optimization	●	●	●
Security	Physical security	●	●	●
	Deployment security	●	●	●
	Monitoring	●	●	●

Data Loading and Management

Snowflake is built to make it as easy as possible to load and manage your data. Because Snowflake can directly load both structured and semi-structured data (like JSON and Avro), you don't need to spend time transforming and converting semi-structured data to load. Unlike other cloud data warehouses, there's no need to partition across separate clusters. You also don't need to worry about the traditional problem of finding an idle window in which to do loading—Snowflake's unique architecture allows you to load data whenever you need, without any performance impact on other users and workloads.

“Snowflake makes it possible for us to focus on making use of our data without the complexity and resources required by traditional data warehousing and big data solutions. The native support for semi-structured JSON data in particular has simplified our data pipeline while delivering great performance.”

— Ethan Erchinger, Director of Operations, Chime

Database Management and Tuning

Snowflake eliminates most of the knobs and management that are common in other databases. It can do this because it was designed to monitor how the data warehouse is being used and dynamically adapt, all without user intervention. Snowflake automatically manages data distribution based on how you use and query data, without requiring you to manually choose and tune distribution keys. Snowflake also dynamically adapts as you scale up and down so that scaling does not require downtime or disruption.

Snowflake delivers a data warehouse that manages and automatically optimizes itself based on usage, eliminating the overhead associated with tuning and managing a conventional data warehouse in the cloud. However, it retains the useful components of the traditional data warehouse, like full ANSI-standard SQL and transactional support.

Security and Resiliency

Snowflake uses the benefits of the cloud to provide unparalleled data security and resiliency. Out of the box, it includes automated data protection, availability, security upgrades, and more. All of the data that is hosted in Snowflake is automatically replicated across multiple physical locations the moment it's loaded in Snowflake. Data in transit and at rest is fully encrypted. What's more, Snowflake monitors the underlying infrastructure of the database to ensure seamless performance and reliability.

This unique combination of capabilities can't be found in other cloud data warehouses, and it makes it possible for you to focus on getting answers from your data, not on managing and tuning your data warehouse.

CONCLUSION

Traditional data warehousing technologies, whether in the cloud or not, fail to deliver the seamless ease of use and limitless scalability of a true DWaaS. Snowflake's groundbreaking technology allows anyone to instantly provision, load, use and secure their data warehouse with far less time and effort than traditional tools.

GET STARTED

To find out how Snowflake can help you get more value from your data, go to <https://www.snowflake.net/free-trial/> and try Snowflake for free today.

Snowflake Computing, the cloud data warehousing company, has reinvented the data warehouse for the cloud and today's data. The Snowflake Elastic Data Warehouse is built from the cloud up with a patent-pending new architecture that delivers 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 is headquartered in Silicon Valley and can be found online at snowflake.net.

