

## Paper Review

The essay “The Snowflake Elastic Data Warehouse” introduces a data store and processing technique: “snowflake”. Snowflake is a data warehouse that provides pay-as-you-go service to customers on Amazon. In addition to that, Snowflake is also a system that supports features such as extreme flexibility, multi-tenancy, end-to-end security, complete CRUD, transactions, built-in semi-structured, and unstructured data (Dageville, 2016). Therefore, the rest of the paper will analyze each function of Snowflake data warehouse.

With the emergence of cloud computing, the structure of data has changed dramatically. In the past, most of the data came from internal companies. Therefore, those data (from corporate) usually have a certain pattern, the type, scale, and structure of data are predictable. However, with more and more cloud platforms appear, the amount of data increases exponentially. Those data have become semi-structured or unstructured and the content of the data is unpredictable. Therefore, companies need a data warehouse that could resolve those data and Snowflake is a software system that solves those problems. Firstly, Snowflake is a software that is easy to use. The Software-as-a-Service (SaaS) function provides standard interfaces and specifications for companies. In addition to that Snowflake does not require much parameter tuning and physical structure design. Thus, companies could get used to it in a short period of time and start processing the data in hand quickly. Another important advantage is its elastic. Snowflake supports elastic scaling, because Snowflake separates computing and storage, so it supports separate expansion of computing and storage (Dageville, 2016). Also, Snowflake provides complete ANSI SQL support and supports transactions. By having that function, it could keep the company’s old data (before cloud computing) and compatible with new data entering the companies’ system, so that the cost for users to transfer from the traditional data

warehouse becomes very small. The last but not least factor is security, to ensure data is safe, Snowflake adopts an end-to-end data encryption algorithm. The data is encrypted on the hard disk and during network transmission, making it impossible for potential attackers, even cloud service providers, to obtain users' plaintext data (Dageville, 2016).

Nowadays, two mainstream techniques have become the focus of attention in the IT field: big data and cloud computing. One fundamental difference is that big data only involves processing massive amounts of data, while cloud computing involves structure design (like algorithm optimization, security etc.). The combination of the two brings beneficial results to not only companies but also users. For a huge amount of data, cloud computing brings significant service supports like system security, data storage and information processing. In addition to that, as the essay “The Snowflake Elastic Data Warehouse” mentioned, cloud computing can also expand the data collection because cloud computing will integrate a large amount of data from users and industries in the future. Thus, big data and cloud computing are two techniques that are closely connected and in this perspective, Snowflake Elastic Data Warehouse has overcome an important problem in cloud computing.

### **Honor Code Pledge:**

**On my honor, as a University of Colorado Boulder student, I have neither given nor received unauthorized assistance.**

### **References**

1. Benoit Dageville, Thierry Cruanes, Marcin Zukowski, Vadim Antonov, Artin Avanes, Jon Bock, Jonathan Claybaugh, Daniel Engovatov, Martin Hentschel, Jiansheng Huang, Allison W. Lee, Ashish Motivala, Abdul Q. Munir, Steven Pelley, Peter Povinec, Greg Rahn, Spyridon Triantafyllis, and Philipp Unterbrunner. 2016. The Snowflake Elastic Data Warehouse. In Proceedings of the 2016 International Conference on Management of Data. Association for Computing Machinery, New York, NY, USA, 215–226.  
DOI:<https://doi.org/10.1145/2882903.2903741>