comforte

COMFORTE DATA PROTECTION FOR SNOWFLAKE

TAKE COMPLETE CONTROL OF YOUR DATA

SOLUTION AT A GLANCE

- Consistent protection of sensitive data (PII, PCI or PHI) at rest, in motion and in use, inside and outside of Snowflake
- Field-level format preserving encryption (FPE) and tokenization
- Granular access control
- Analytics on protected data sets without performance issues
- Hold your own key to keep full control of your data
- Sustainable privacy and regulatory compliance

INTRODUCTION

Organizations all over the world are using Snowflake's Data Cloud to unlock value for their businesses. However, moving the data from on-premises stores and applications to the cloud is often challenging due to security and regulatory compliance concerns. Comforte Data Protection for Snowflake provides the ability to properly protect sensitive data to meet regulatory obligations while keeping it usable for business processes, applications and analytics.

HOW SNOWFLAKE PROTECTS DATA

Snowflake provides Dynamic Data Masking to keep sensitive data hidden from unauthorized users. It is a column-level feature using masking policies to selectively mask plain-text data in table and view columns at query time, at every location where the column appears. So, depending on masking policy for their particular role, users may see the plain-text value ("John Smith"), partially masked value ("Jxxx Sxxxx") or fully masked value ("xxxx xxxxx"). However, the data must be loaded in plain-text into Snowflake, and a database and schema must exist before a masking policy can be applied to a column. This means that there is still a high risk of misconfiguration and data exposure as the data still remains unprotected at the database level and when it is used in external applications.

MULTICLOUD DATA PROTECTION TAILORED TO YOUR NEEDS - END-TO-END

Comforte provides robust data protection that exceeds built-in capabilities of cloud-based data stores. It enables organizations to protect structured and semi-structured data not only inside but also outside of Snowflake, and keep it protected as it flows in hybrid and multicloud data ecosystems. This helps customers to provide self-service access to data, enabling processing and analytics that were previously impossible because of privacy or security concerns. Cloud-native architecture and integration capabilities help customers to rapidly implement data-centric security, protect data as early as possible and consistently apply security policies to keep data secure throughout its lifecycle.

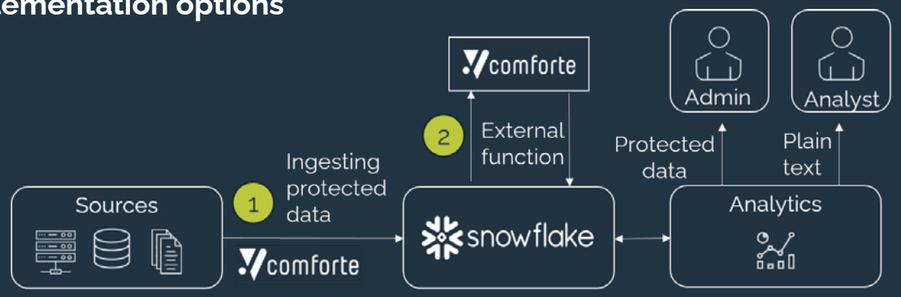
Comforte provides multiple data protection methods to address specific needs and use cases. Techniques such as data masking, format preserving hashing or tokenization and format-preserving encryption (FPE) enable pseudonymization or full anonymization of data. While anonymization completely removes sensitivity from data and makes it unusable (e.g. for advanced analytics), pseudonymization methods—such as FPE—protect the data in a way that it is still usable in analytics or BI tools; preserving the format of the data, referential integrity and more.

HOW THE SOLUTION WORKS

Comforte Data Protection for Snowflake pseudonymizes sensitive data (PII, PHI, PCI) using tokenization or format-preserving encryption (FPE) on a field level. Instead of just masking the values, the data element is completely replaced by a token in the database. The token itself can be mapped back to the original data element, but doesn't expose any sensitive information. The biggest advantage of tokenization over classic encryption is its format preserving property. Tokens preserve the format and length of original data elements which keeps them usable for business applications and analytics. This allows for use of any query or BI tool with no need to change the SQL syntax. Furthermore, tokens do not require a lot of computational resources to process, allowing for high performance and low latency. Since the data store and protection engine are strictly separated, this method helps achieve compliance with privacy and data protection regulations and immensely reduces risks related with data breaches, because tokenized data has no value for potential misuse.

Comforte`s advanced protection technologies have the ability to integrate at any point in data environment from sources to analytics tools. This enables transformation of data based on fine granular policies, making it possible to easily define which user roles can see which data element, in which form and, for example, provide access to the data in clear to data analysts only.

Implementation options



1. Protecting the data before it reaches Snowflake. This option is recommended if no sensitive data should be stored on Snowflake. Data will be protected either at the cloud ingestion stage – before it touches storages such as Amazon S3 - or interception stage when data is loaded into Snowflake itself. This is done by leveraging comforte`s capability to transparently intercept data and apply tokenization on the fly, working as a proxy between the data source and the database (see diagram above).

2. External tokenization function from Snowflake leveraging a Rest API. This option is recommended if you are looking for a solution to add advanced protection to your sensitive data and control access via Snowflake`s role based dynamic data masking policies. That enables certain roles such as data analysts to retrieve data in the clear while others will only see protected data.

BENEFITS

- Accelerate data-driven business initiatives
  - Move to the cloud while keeping data secure
- Run analytics on protected data sets
  - Enable secure usage of data for analytics
- Protect privacy and achieve compliance
  - Pseudonymized data is fully compliant with privacy regulations
- Simplify data security and management
  - Consistent protection of data across different tools, complex architectures, environments and even separate cloud providers

Sample architecture