Voltage SecureData Integrations for Snowflake

Enabling high-scale, high-performance, and secure data analytics, data science and data sharing.

Acceleration to the Cloud & Privacy Compliance
According to Gartner, the use of cloud-native technologies will be pervasive, not just popular. By 2025, Gartner estimates that over 95% of new digital workloads will be deployed on cloud-native platforms.* Many companies that have invested in expensive on-premises data warehouse systems, Hadoop data lakes, and their surrounding ecosystems have shifted their data into the cloud.

Why? The cost-effectiveness of cloud storage and its ever-increasing array of services enable organizations to get more value from monetizing their rapidly expanding data volumes in these large-scale environments.

However, as demonstrated by an almost continuous stream of reports, cloud-related data breaches are not likely to decline. So, amid these very insistent demands for cloud adoption from the business side, data security nevertheless remains a board-level concern. Voltage SecureData Enterprise can help customers ensure that the adoption of cloud services, like the Snowflake Data Cloud, doesn't result in a breach of sensitive data, (such as PII, PCI, PHI, and intellectual property), corresponding regulatory fines, or damage to brand, reputation, and customer trust.

Key Benefits

Voltage SecureData Enterprise enables secure analytics by applying data-centric protection within the existing schema of a data store. SecureData preserves a protected dataset's referential integrity so that it retains its value for analytics.

Here are three use cases that show how enterprises can benefit from Voltage SecureData Enterprise Integrations for Snowflake:

- **Persistently protect data**
  Protect data before, when or any time after it lands in the Snowflake Data Cloud.
- **Conduct secure analytics in Snowflake**
  Analyze data in its protected form with role-based access enabled directly on specific data elements via SQL function calls or conducted transparently in combination with Snowflake masking policies.
- **Securely and safely share data**
  Import data already protected by Voltage, analyze on other platforms, clouds without removing protection and monetize analytics by third parties external to Snowflake.

* Gartner. (2021, November 10). Gartner Says Cloud Will Be the Centerpiece of New Digital Experiences [Press release]. **www.gartner.com/en/newsroom/press-releases/2021-11-10-gartner-says-cloud-will-be-the-centerpiece-of-new-digital-experiences**

Product Highlights

How Voltage Data Privacy and Protection Can Help

The OpenText™ Cybersecurity Voltage Data Privacy and Protection framework includes critical capabilities from data discovery to disposition. Understanding the flow, use, and storage of data is key to compliance in today's era of global privacy legislation. Organizations need practical tools to find data within the scope of privacy policies, automate tagging and enrichment of metadata, identify data subject information, and to assess risk.

Voltage Data Privacy and Protection provides solutions that discover, analyze, and classify all data, whether structured, semi-structured, or unstructured. Policies covering the entire data lifecycle allow enterprises to act on their data with contextual awareness and deep insights from rich risk profile visualizations. Voltage SecureData Enterprise in particular uses standards-validated, data-centric security innovations to pseudonymize and anonymize sensitive information, to deliver persistent privacy for data wherever it resides, moves, or is used.

Voltage SecureData Integrations for Snowflake Benefits

The Snowflake platform provides several native, layered security options, including network security, identity and access management, transparent disk encryption, TLS, and standard data-at-rest encryption. Voltage SecureData Enterprise adds important data-centric protection options that enhance data security in the Snowflake Data Cloud in several critical ways:

- Format-preservation for the usability of data in its protected form
- Data security controls for data privacy regulation compliance
- Persistent protection enabling multi-cloud and data sharing strategies
- Flexibility of standards-validated and independently assessed techniques
- Safe unicode support for all alphabets

Data protected by Voltage SecureData Enterprise's range of tokenization technologies retains its referential integrity, enabling customers to perform data analytics upon protected data sets. Voltage SecureData Enterprise preserves data formats so that the protected form of the data fits seamlessly into existing table schema. Reversible and irreversible methods for sensitive data types across all languages are provided. These methods include format-preserving encryption (FPE), secure stateless tokenization (SST), and format-preserving hash (FPH) that enable customers to pseudonymize and anonymize data, as required, wherever it is or wherever it must go.

In addition, the mobility of data protected by Voltage SecureData Enterprise is unconstrained: data remains protected while flowing into or out of Snowflake, from or to other cloud services or cloud platforms. This approach supports a multi-cloud strategy and data sharing requirements without requiring organizations to compromise on data security at the boundaries of these services. And with Voltage SecureData Enterprise, you always remain in complete control of your encryption keys and token tables, from the master keys down to the data encryption keys themselves, all in a stateless system that imparts no additional storage or management overhead.

Customer Benefit	Snowflake	Voltage SecureData
Storage-level and file-level encryption (customer managed master keys in a stateful system)	Yes	
Column-level dynamic data masking policies	Yes	
Field-level encryption (compatible with column-level masking policies)	Yes	Yes
Format-preserving field-level data protection (compatible with column-level masking policies)		Yes
Format-Preserving Encryption (FPE) for sensitive PII, PHI		Yes
Secure Stateless Tokenization (SST) for PCI audit scope reduction		Yes
Format-Preserving Hash (FPH) for one-way de-identification of sensitive data		Yes
Data sharing using persistent data protection for increased security and usability		Yes
Transport and use protected data across multi-cloud hybrid IT		Yes
Customer controlled master and data encryption keys in a stateless system		Yes

Figure 1. Voltage SecureData Enterprise enhances Snowflake security

Key Features

Reference Architecture

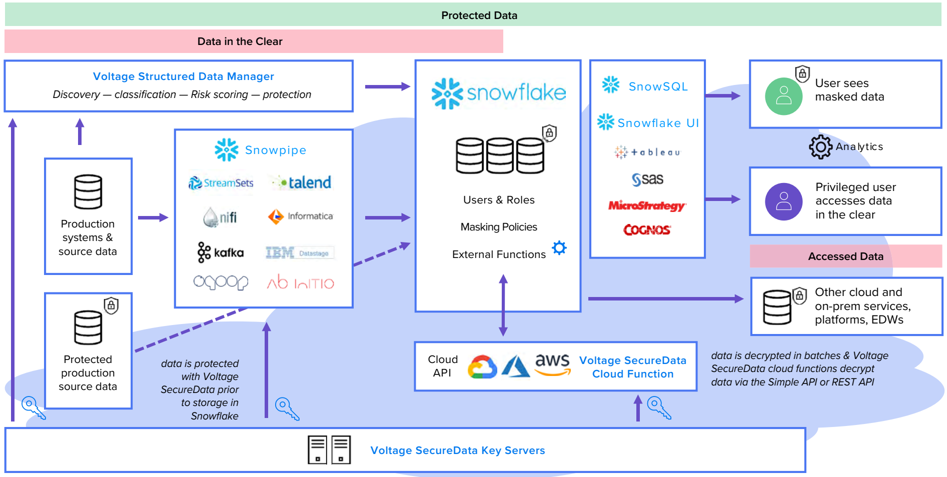
Protected Data
Data in the Clear
Voltage Structured Data Manager
Discovery — classification — Risk scoring — protection
Production systems & source data
Protected production source data
Snowpipe
StreamSets
talend
nifi
Informatica
kafka
IBM Datastage
sqoop
Ab Initio
data is protected with Voltage SecureData prior to storage in Snowflake
snowflake
Users & Roles
Masking Policies
External Functions
SnowSQL
Snowflake UI
tableau
sas
MicroStrategy
COGNOS
User sees masked data
Analytics
Privileged user accesses data in the clear
Accessed Data
Other cloud and on-prem services, platforms, EDWs
Cloud API
aws
Voltage SecureData Cloud Function
data is decrypted in batches & Voltage SecureData cloud functions decrypt data via the Simple API or REST API
Voltage SecureData Key Servers

Figure 2. Share and shift protected data to and from Snowflake

The high-level reference architecture demonstrated in Figure 2 shows the data flow from left to right, from where data may originate, through where it is stored and in use in the Snowflake Data Cloud, to how users might view and access their data.

Commonly, data protection would be done on-premises before uploading to the cloud, or it could be done as it is moved into or lands in the cloud. A wide variety of tools could be used to move this data into Snowflake, from traditional ETL tools and Hadoop-based utilities to tools like Snowpipe. These are all existing or possible integration points for Voltage SecureData Enterprise to invoke format-preserving tokenization technologies if Voltage SecureData Enterprise has not already protected data in on-premises systems and business applications. As the data protection that Voltage SecureData Enterprise provides is both persistent and platform-agnostic, it supports a data protection strategy that is enterprise-wide and multi-cloud, with the ability to protect data in one place and then unprotect that data in a different place, such as in a completely different cloud altogether.

Once moved into Snowflake, data that has already been protected by Voltage SecureData Enterprise is now safe in the Snowflake Data Cloud and its analytic value has been preserved. However, if for any reason, sensitive data must land in Snowflake in the clear, or if a Snowflake customer wants to transform data that is already in Snowflake with Voltage format-preserving security, protection methods can also be invoked as it arrives.

Product Features

Voltage SecureData Enterprise Integrations for Snowflake uses the Snowflake External Functions interface to enable access to the Voltage SecureData Enterprise range of advanced data protection options. The product includes a set of robust scripts that automatically set up all the required cloud infrastructure components, and detailed step-by-step instructions for manual installation and configuration are also provided.

Voltage SecureData Enterprise Integrations for Snowflake includes:

- Build scripts that automate the installation and configuration of the integration and its supporting cloud services.
- Manual installation and configuration guides that explain how the integration is built, and how it can extend to include additional options.
- Support for all data protection types available in Voltage SecureData Enterprise, including FPE, eFPE, FPH, and SST, for pseudonymization or anonymization of PII, PHI, and PCI data.
- Access to all of Voltage SecureData Enterprise's built-in formats, including AUTO, the Japanese formats, and language support via 'Safe Unicode FPE.'

In summary, the data protection that Voltage SecureData Enterprise provides is both persistent and platform-agnostic and allows customers to protect data before, when, or any time after it lands in Snowflake. Customers can conduct analytics tasks in Snowflake on data protected with Voltage SecureData Enterprise while it is in its protected format-preserved form, integrate with Snowflake's role-based access controls and masking policies for transparent access to result sets, and share data securely and safely both inside and outside of Snowflake

> **“Our clients want to lock down their sensitive data, but still be able to unlock the value in that data at scale.”**
>
> **Head of Cybersecurity Practice**
> Global Systems Integrator

Connect with Us
www.opentext.com

for extended first-party, second-party, and third-party analytics. In other words, Voltage SecureData Enterprise Integrations for Snowflake enables you to lock down your sensitive data but still unlock the value in that data at scale.

Learn more at
www.microfocus.com/en-us/cyberres/partners/snowflake

opentext™ | Cybersecurity

OpenText Cybersecurity provides comprehensive security solutions for companies and partners of all sizes. From prevention, detection and response to recovery, investigation and compliance, our unified end-to-end platform helps customers build cyber resilience via a holistic security portfolio. Powered by actionable insights from our real-time and contextual threat intelligence, OpenText Cybersecurity customers benefit from high efficacy products, a compliant experience and simplified security to help manage business risk.