Introduction to Snowflake

Sunil Gurav

What is Snowflake?

Snowflake is an analytic data warehouse provided as Software-as-Services(SaaS). Snowflake provides data warehouse that is faster, easier to use and more flexible that other traditional data warehouses.

Snowflake data warehouse is not built on existing databases or not on big data software platform as Hadoop.

The snowflake data warehouse uses a new SQL database engine with unique architecture designed for the cloud.

Key Concept and Architecture

Data Warehouse as Cloud Service:

Snowflake data warehouse is true SaaS offering.

- There is no hardware (virtual or physical) for you to select, install, configure and manage.
- There is no software for you install, configure and manage.
- Ongoing maintenance, management and tuning is handled by snowflake.

Snowflake completely runs on cloud infrastructure. All the component of the snowflake service runs on public cloud infrastructure.

Snowflake uses virtual compute instance for its compute need and storage service for storage of data. Snowflake can not be run on private cloud infrastructure(on-primises)

Snowflake Architecture

Snowflake architecture is the combination of shared-disk database architecture and shared-nothing database architecture.

Shared-disk database architecture:

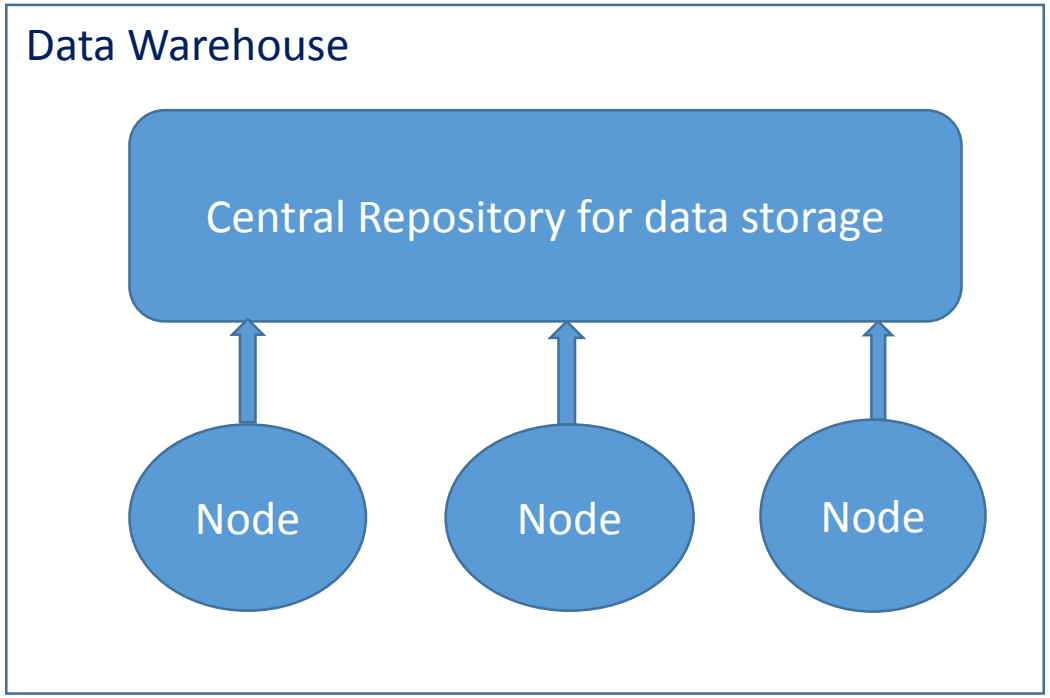
Snowflake uses central repository for data storage that is accessible from all the compute nodes in the data warehouse.

Sharing-nothing architecture:

Snowflake processes queries using Massive Parallel Processing(MMP) compute cluster where each nodes in compute cluster stores a portion of the entire data set locally.

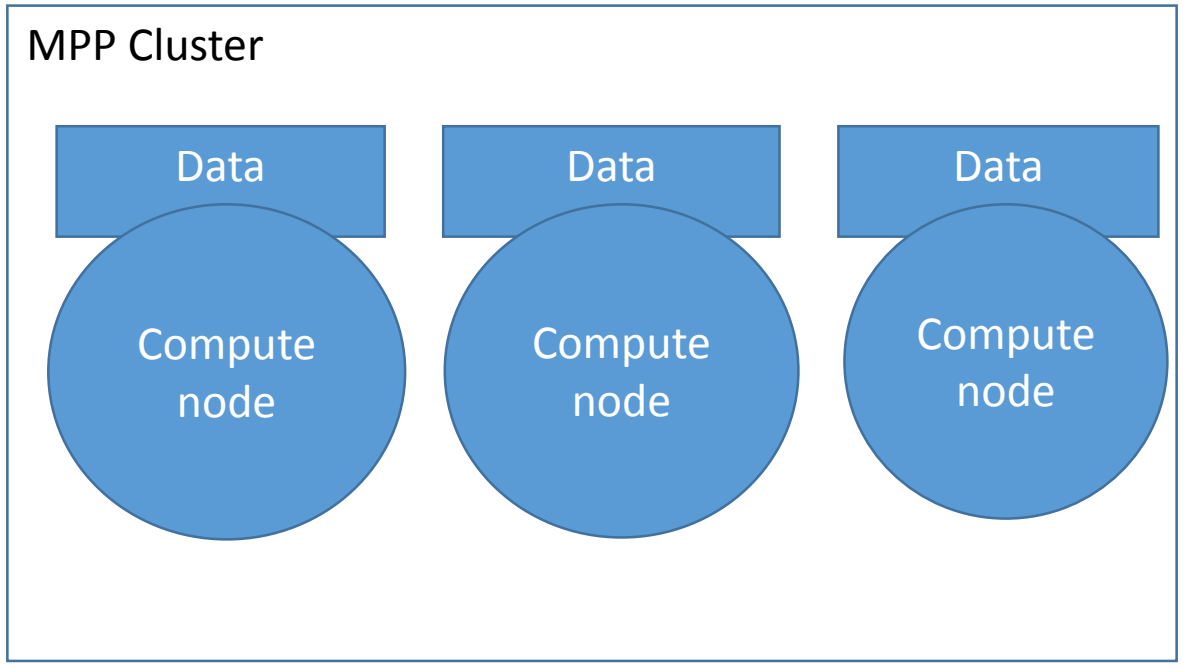
Shared-Disk Database Architecture

In this architecture each node can access the data stored in central repository.

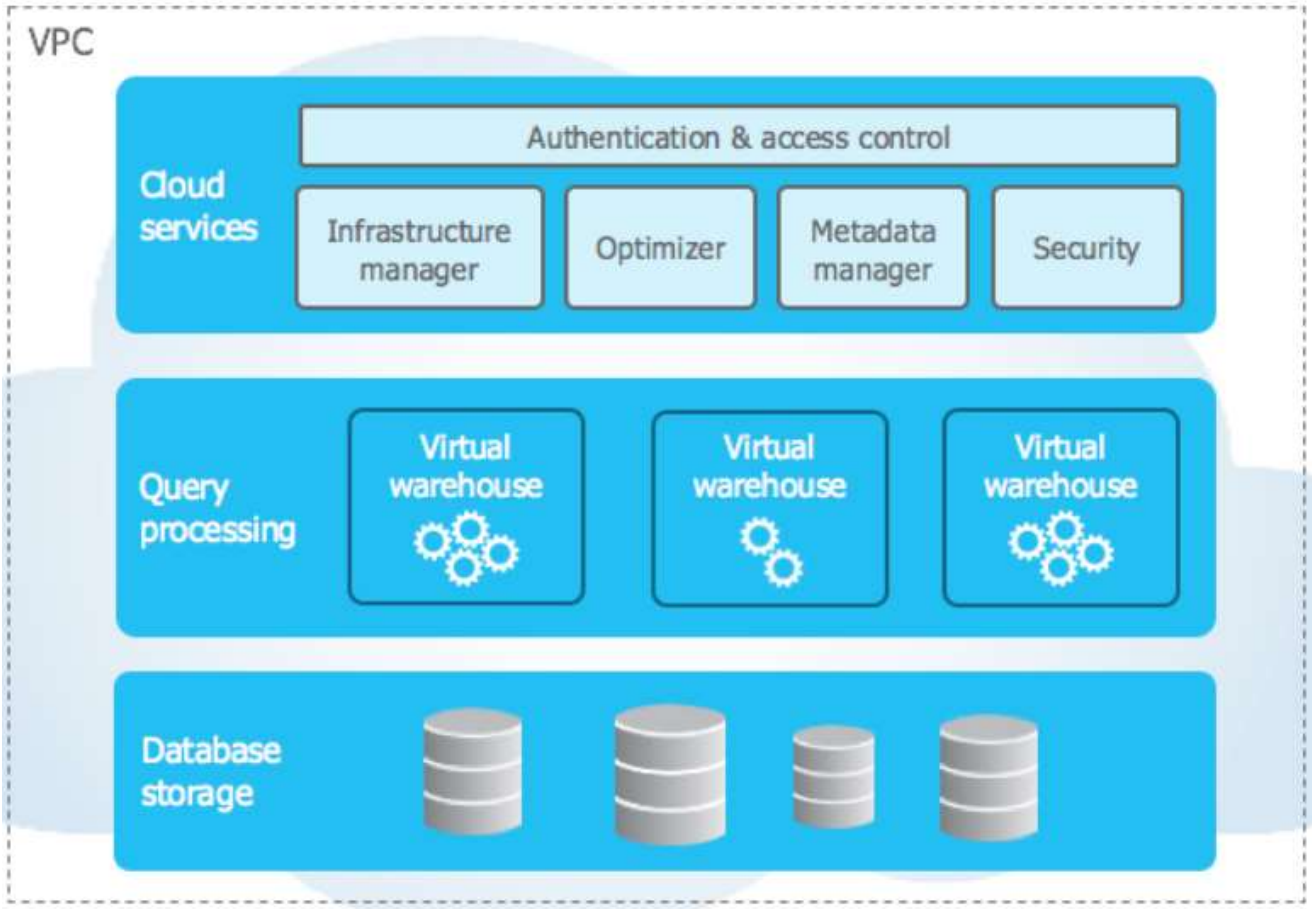


Shared-nothing Database Architecture

- Snowflake processes the queries using MMP compute cluster where each node in the cluster stores a portion of entire data set locally.



Snowflake Architecture



Snowflake architecture mainly consist of three layers:

- Database Layer
- Query Processing Layer
- Cloud Services

Database Layer

When you load the data into snowflake, Snowflake reorganizes that data into its internal optimized, compressed, columnar format.

Snowflake store this optimized data in cloud storage.

Snowflake manages all aspects of how this data is stored — the organization, file size, structure, compression, metadata, statistics, and other aspects of data storage are handled by Snowflake. The data objects stored by Snowflake are not directly visible nor accessible by customers; they are only accessible through SQL query operations run using Snowflake.

Query Processing Layer

- Query execution is performed in the processing layer. Snowflake processes queries using “virtual warehouses”. Each virtual warehouse is an MPP compute cluster composed of multiple compute nodes allocated by Snowflake from a cloud provider.
- Each virtual warehouse is an independent compute cluster that does not share compute resources with other virtual warehouses. As a result, each virtual warehouse has no impact on the performance of other virtual warehouses.

Cloud Services

The cloud services layer is a collection of services that coordinate activities across Snowflake. These services tie together all of the different components of Snowflake in order to process user requests, from login to query dispatch. The cloud services layer also runs on compute instances provisioned by Snowflake from the cloud provider.

Among the services in this layer:

- Authentication
- Infrastructure management
- Metadata management
- Query parsing and optimization
- Access control

Connecting to Snowflake

Snowflake supports multiple ways of connecting to the service:

- A web-based user interface from which all aspects of managing and using Snowflake can be accessed.
- Command line clients (e.g. SnowSQL) which can also access all aspects of managing and using Snowflake.
- ODBC and JDBC drivers that can be used by other applications (e.g. Tableau) to connect to Snowflake.
- Native connectors (e.g. Python) that can be used to develop applications for connecting to Snowflake.
- Third-party connectors that can be used to connect applications such as ETL tools (e.g. Informatica) and BI tools to Snowflake.

Cloud Platform

Snowflake provides data warehouse as Software-as-Services(SaaS) that completely runs on cloud platform.

This means that all three layers of snowflake's architecture(Database, Query Processing, Cloud Services) are deployed and managed entirely on selected cloud platform.

Snowflake account can be hosted on the following cloud platform:

- Amazon Web Services(AWS)
- Microsoft Azure

Pricing

Differences in unit costs for credits and data storage are calculated by Snowflake Region and not by cloud platform. For more information about pricing as it pertains to a specific region, see the pricing page (in the Snowflake website).

Data Loading

Snowflake supports data loading from files staged in any of the following locations:

- Internal(snowflake) stage
- Amazon S3
- Microsoft Azure Blob Storage

Snowflake supports both bulk data loading and continuous data loading(snowpipe). Likewise snowflake supports unloading data from tables into any of above staging locations.

Current Limitations for accounts on Azure

The following services and features are currently unavailable for snowflake accounts hosted on Azure:

Area	Limitations
Security & Data Encryption	• No support for encryption with customer-managed keys (aka Tri-Secret Secure). • No support for secure connectivity to customer-owned virtual networks (similar to AWS PrivateLink).
Data Loading	• No support for accessing an external AWS S3 stage using policies attached to an IAM role.
3rd-party Applications	• All native Snowflake clients support Azure; however, not all 3rd-partner applications provide the same level of support. The following Snowflake partners are known to support Snowflake hosted on Azure: ◦ Alooma ◦ Attunity ◦ Databricks ◦ Fivetran ◦ Informatica ◦ Looker ◦ Matillion ◦ MicroStrategy ◦ Periscope ◦ Power BI ◦ Qubole ◦ Sigma Computing ◦ Stitch ◦ Tableau ◦ Talend ◦ Wherescape

Snowflake Regions

Snowflake Regions let your organization choose where your data is geographically stored across your global operations. They also determine where your compute resources are provisioned. You can choose to have multiple accounts in separate regions to address global latency concerns and/or provide additional backup and disaster recovery support beyond the standard support provided by Snowflake.

Note that regions do not limit user access to Snowflake; they only dictate the geographic location where data is stored and compute resources are provisioned.

Important:

- Each Snowflake account is located in a single region (i.e. multi-region accounts are not supported).
- In addition, Snowflake does not yet support accessing or sharing data between regions. If you wish to use Snowflake across multiple regions, you must have a separate Snowflake account in each region.

List of Regions by cloud platform

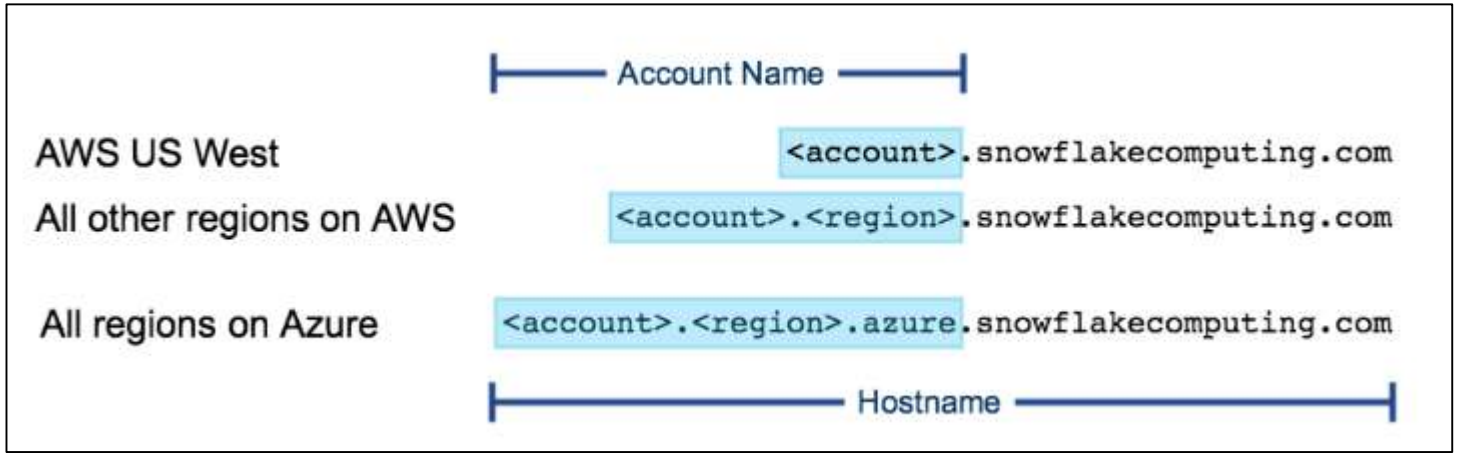
Snowflake maintains the following regions:

Snowflake Region	Region ID	Notes
AWS US West	`us-west-2`	This is the default region; it is not utilized/required unless AWS PrivateLink is enabled for your account.
AWS US East	`us-east-1`	
AWS EU (Frankfurt)	`eu-west-1`	
AWS EU (Dublin)	`eu-central-1`	
AWS Asia Pacific (Sydney)	`ap-southeast-2`	
Azure East US 2	`east-us-2.azure`	
Azure West Europe	`west-europe.azure`	

Region and Cloud Platform in Account Names

A hostname for a snowflake account starts with a unique account name and ends with snowflakecomputing.com.

However, depending on the cloud platform (AWS or Azure) and Snowflake Region where your account is hosted, the full ***account name*** may require ***additional*** segments, as shown in the following diagram:



United States (US) Regions

Snowflake maintains the following US regions:

Cloud Platform	Snowflake Region	Geographic Location	Support Hours for Standard Edition
AWS	US West	Oregon	6:00am - 6:00pm Pacific Time (PST or PDT)
AWS	US East	Virginia	6:00am - 6:00pm Pacific Time (PST or PDT)
Azure	East US 2	Virginia	6:00am - 6:00pm Pacific Time (PST or PDT)

If your account is hosted on AWS and latency is a concern, you should choose the available US region with the closest geographic proximity to your end users.

European Union(EU) Regions

Snowflake maintains following EU regions:

Cloud Platform	Snowflake Region	Geographic Location	Support Hours for Standard Edition
AWS	EU (Frankfurt)	Germany	6:00am - 6:00pm Central Europe Time (CEST or CEDT)
AWS	EU (Dublin)	Ireland	6:00am - 6:00pm Central Europe Time (CEST or CEDT)
Azure	West Europe	Netherlands	6:00am - 6:00pm Central Europe Time (CEST or CEDT)

These regions are provided for organizations that prefer or require their data to be stored in the European Union. Multiple regions are provided to allow organizations to meet their individual compliance and data sovereignty requirements. Snowflake does not move data between accounts, so any data in an account in an EU region will be maintained in this region only, unless explicitly copied or moved by users.

Asia Pacific(AP) Regions

Snowflake currently maintains a single AP region:

Cloud Platform	Snowflake Region	Geographic Location	Support Hours for Standard Edition
AWS	AP (Sydney)	Australia	6:00am - 6:00pm Australian Eastern Time (AEST or AEDT)

This region is provided for organizations that prefer or require their data to be stored in the Asia Pacific part of the world. Snowflake does not move data between accounts, so any data in an account in an AP region will be maintained in this region only, unless explicitly copied or moved by users.

Snowflake Edition

Snowflake provides several alternatives to ensure that your usage of the service fits your organization's specific requirements. This includes offering multiple *Snowflake Editions* to chose from, with each successive edition building on the previous edition through additional features and/or higher levels of service. And, as your organization's needs change and grow, changing editions is easy.

Note:

The Snowflake Edition that your organization chooses determines the unit costs for the credits and the data storage you use. Another factor that impacts unit costs is whether you have a Snowflake *On Demand* or *Capacity* account:

- On Demand: Usage-based pricing with no long-term licensing requirements.
- Capacity: Discounted pricing based on an up-front Capacity commitment.

Standard Edition

Standard edition is introductive level, providing full, unlimited access to all of snowflake standard feature.

Functional Area	Features/Services
Security and Data Protection	• SOC 2 Type II certification. • Network policies for limiting/controlling site access by user IP address. • Automatic encryption of all data. • Support for multi-factor authentication. • Object-level access control. • Access to all modified and deleted data (up to 1 day) through Time Travel. • Disaster recovery of modified and deleted data (up to 7 days beyond Time Travel) through Fail-safe.
SQL	• Standard SQL, including most DDL and DML defined in SQL:1999. • Advanced DML such as multi-table INSERT, MERGE, and multi-merge. • Broad data type support. • Multi-statement transactions. • User-defined functions (UDFs) with support for both SQL and Javascript. • Native support for semi-structured data (JSON, Avro, ORC, Parquet, and XML).

Standard Edition

Tools and Interfaces	• Full-featured web-based interface and command line client. • Programmatic interfaces for Python, Node.js, and Spark. • Native support for JDBC and ODBC. • Extensive ecosystem for connecting to ETL, BI, and other third-party systems.
Data Import and Export	• Support for bulk loading from delimited flat files (CSV, TSV, etc.) and semi-structured data files (JSON, Avro, ORC, Parquet, and XML). • Support for bulk unloading to delimited flat files and JSON files.

Support	• Access to the Snowflake Community Lodge, Snowflake's online community forum and support portal.
	• Standard support, including:
	  ◦ Availability Monday through Friday during business hours (6:00am - 6:00pm):
	  ◦ For accounts in the US West, US East, and East US 2 Snowflake Regions, business hours are based on Pacific Time (PST or PDT).
	  ◦ For accounts in the EU (Frankfurt) and EU (Dublin) Snowflake Regions, business hours are based on Central Europe Time (CEST or CEDT).
	  ◦ For accounts in the AP (Sydney) Snowflake Region, business hours are based on Australian Eastern Time (AEST or AEDT).
	  ◦ 4-hour response window for Severity 1 issues.

Premier Edition

Snowflake offers Premier Edition as a cost-effective option for organizations that do not need additional features, but would benefit from expanded access to Snowflake Support.

It includes all the features and services of Standard Edition, with an upgrade to the next level of support:

Functional Area	Features/Services
Support	• Premier support, including: ◦ 24x365 availability for Severity 1 issues. ◦ 1-hour response window for Severity 1 issues.

Enterprise Edition

Enterprise Edition provides all the features and services of Premier Edition, with the following additional features designed specifically for the needs of large-scale enterprises and organizations:

Functional Area	Features/Services
Security and Data Protection	• Federated authentication and SSO for centralizing and streamlining user authentication. • Periodic rekeying of encrypted data for increased protection. • Extended data retention for Time Travel (up to 90 days).
Automated Resource Management	• Multi-cluster warehouses for scaling compute resources to meet fluctuating concurrency needs.
SQL	• Materialized views, including automatic maintenance of results.

Enterprise Edition for Sensitive Data(ESD)

Enterprise for Sensitive Data offers even higher levels of data protection to support the needs of organizations with extremely sensitive data, particularly PHI data that must comply with HIPAA regulations.

It includes all the feature and services of Enterprise Edition, with the addition of the following enhanced security and data protection:

Functional Area	Features/Services
Security and Data Protection	• Additional levels of security and data protection, as covered by the Snowflake Security Policy. • Encryption of all communication between Snowflake data storage (i.e. databases) and compute resources (i.e. virtual warehouses). • Tri-Secret Secure using customer-managed keys. • Support for secure, direct proxy (through AWS PrivateLink) to your other VPCs or on-premises data centers. • Support for PHI data (in accordance with HIPAA regulations). • Support for PCI DSS.

Virtual Private Snowflake(VPS)

Virtual Private Snowflake offers our highest level of security for organizations that have the strictest security requirements, such as financial institutions.

It includes all the feature and services of Enterprise Edition for Sensitive Data (ESD), but in an separate Snowflake environment, isolated from all other Snowflake accounts (i.e. VPS accounts do not share any resources with other accounts). VPS delivers this level of extreme security through the use of:

Functional Area	Features/Services
Security and Data Protection	• Dedicated pool of virtual servers used in virtual warehouses. • Dedicated metadata store. • Full encryption of all data transmissions, including internal transmissions between the components isolated within the VPS environment.

Overview of key Feature

This topic lists the notable/significant features supported in the current release. Note that it does not list every feature provided by Snowflake.

- Security and Data Protection
- Standard and Extended SQL Support
- Tools and Interfaces
- Connectivity
- Data Import and Export
- Data Sharing

Security and Data Protection

- Choose the level of security you require for your Snowflake account, based on your Snowflake Edition.
- Choose the geographical location where your data is stored, based on your Snowflake Region.
- User authentication through standard user/password credentials.
- Enhanced authentication:
  - Multi-factor authentication (MFA).
  - Federated authentication and single sign-on (SSO) — requires Snowflake Enterprise Edition.
- All communication between clients and the server protected through TLS.
- Deployment inside a cloud platform VPC.
- Isolation of data via Amazon S3 policy controls.
- Support for PHI data (in compliance with HIPAA regulations) — requires Snowflake Enterprise for Sensitive Data (ESD).
- Automatic data encryption by Snowflake using Snowflake-managed keys.
- Object-level access control.
- Snowflake Time Travel (1 day standard for all accounts; additional days, up to 90, allowed with Snowflake Enterprise) for:
  - Querying historical data in tables.
  - Restoring and cloning historical data in databases, schemas, and tables.
- Snowflake Fail-safe (7 days standard for all accounts) for disaster recovery of historical data.

Standard and Extended SQL Support

- Most DDL and DML defined in SQL:1999, including:
  - Database and schema DDL.
  - Table and view DDL.
  - Standard DML such as UPDATE, DELETE, and INSERT.
  - DML for bulk data loading/unloading.
  - Core data types.
  - SET operations.
  - CAST functions.
- Advanced DML such as multi-table INSERT, MERGE, and multi-merge.
- Transactions.
- Temporary and transient tables for transitory data.
- Lateral views.
- Statistical aggregate functions.
- Analytical aggregates (Group by cube, rollup, and grouping sets).
- Parts of the SQL:2003 analytic extensions:
  - Windowing functions.
  - Grouping sets.
- Scalar and tabular user-defined functions (UDFs), with support for both SQL and JavaScript.
- Information Schema for querying object and account metadata, as well as query and warehouse usage history data.

Tools and Interface

- Web-based GUI for account and general management, monitoring of resources and system usage, and querying data.
- SnowSQL (Python-based command line client).
- Virtual warehouse management from the GUI or command line, including creating, resizing (with zero downtime), suspending, and dropping warehouses.

Connectivity

- Broad ecosystem of supported 3rd-party partners and technologies.
- Support for using free trials to connect to selected partners.
- Extensive set of client connectors and drivers provided by Snowflake:
  - Python connector
  - Spark connector
  - Node.js driver
  - Go Snowflake driver
  - .NET driver
  - JDBC client driver
  - ODBC client driver
  - dplyr-snowflakedb(open source dplyr package extension maintained on GitHub)

Data Import and Export

- Support for bulk loading and unloading data into/out of tables, including:
  - Load any data that uses a supported character encoding.
  - Load data from compressed files.
  - Load most flat, delimited data files (CSV, TSV, etc.).
  - Load data files in JSON, Avro, ORC, Parquet, and XML format.
  - Load from S3 data sources and local files using Snowflake web interface or command line client.
- Support for continuous bulk loading data from files:
  - Use Snowpipe to load data in micro-batches from internal stages (i.e. within Snowflake) or external stages (i.e. in S3 or Azure).

Data Sharing

- Support for sharing data with other Snowflake accounts:
- Provide data to other accounts to consume.
- Consume data provided by other accounts.

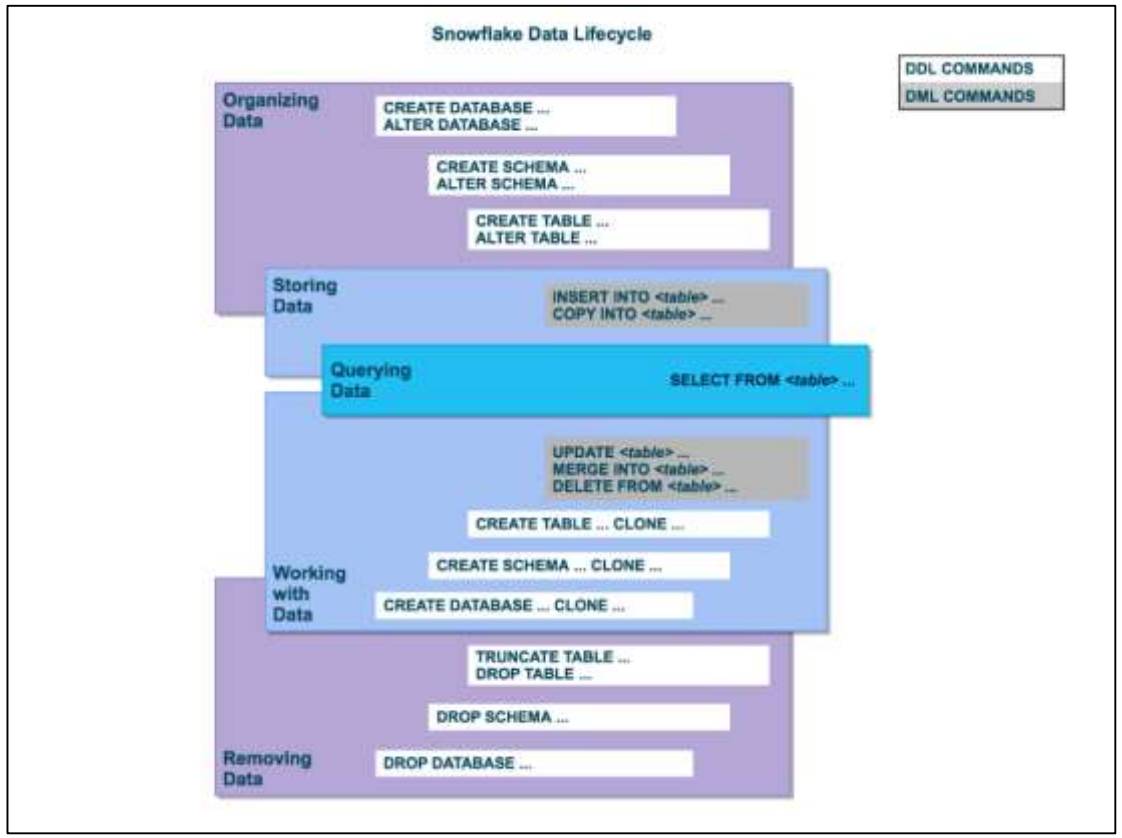
Overview of the Data Lifecycle

Snowflake provides support for all standard SELECT, DDL, and DML operations across the lifecycle of data in the system, from organizing and storing data to querying and working with data, as well as removing data from the system.

- Lifecycle Diagram
- Organizing Data
- Storing Data
- Querying Data
- Working with Data
- Removing Data

Lifecycle Diagram

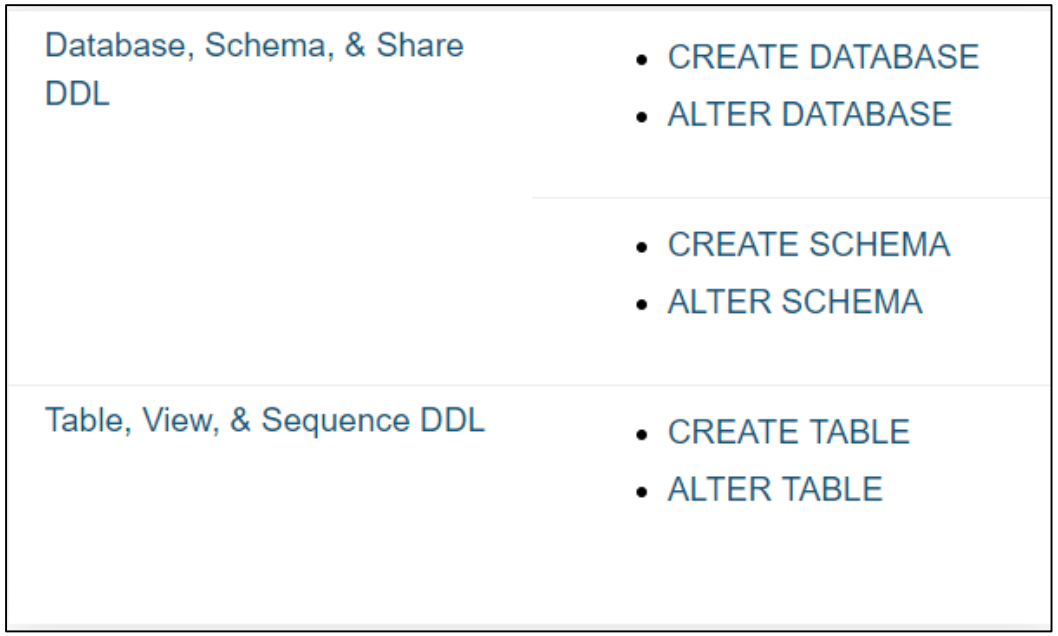
All user data in snowflake is logical represented as tables that can be queried and modified through standard SQL interface. Each table belongs to schema which in turn belongs to database.



Organizing Data

You can organize your data into databases, schemas, and tables. Snowflake does not limit the number of databases you can create or the number of schemas you can create within a database. Snowflake also does not limit the number of tables you can create in a schema.

For more information, see:



Storing Data

You can insert data directly into tables. Snowflake provides DML for loading data into snowflake tables from external, formatted files.

For more information, see:

DML Commands (by Category) (General)	• INSERT
DML Commands (by Category) (Data Loading)	• COPY INTO <table>

Querying Data

Once the data stored in table, you can issue SELECT statement to query data.

For more Information, see:

Working With Data

Once the data is stored in table, All standard DML operations can be performed on the data. Snowflake support DDL actions such as cloning entire databases, schemas and tables.

For more information, see:

DML Commands (by Category) (General)	• UPDATE • MERGE • DELETE
Table, View, & Sequence DDL	• CREATE <object> … CLONE
Database, Schema, & Share DDL	• CREATE <object> … CLONE

Removing Data

In addition to using DML command, DELETE, to remove data from table. You can truncate or drop an entire table. You can also drop entire schemas and databases.

For more information, see:

Table, View, & Sequence DDL	• TRUNCATE TABLE • DROP TABLE
Database, Schema, & Share DDL	• DROP SCHEMA • DROP DATABASE

Continuous Data Protection

Continuous Data Protection (CDP) encompasses a comprehensive set of features that help protect data stored in Snowflake against human error, malicious acts, and software or hardware failure. At every stage within the data lifecycle, Snowflake enables your data to be accessible and recoverable in the event of accidental or intentional modification, removal, or corruption.

Connecting to Snowflake

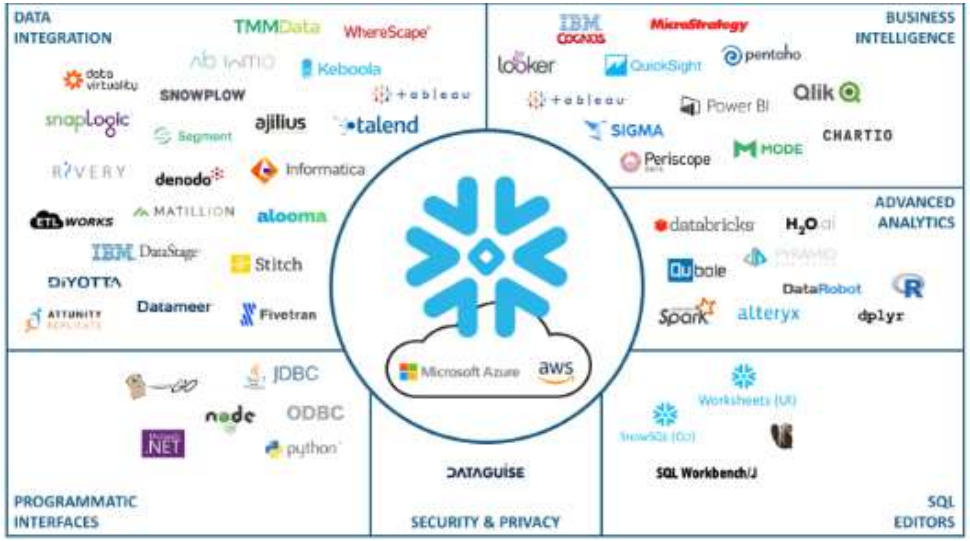
These topics provide an overview of the Snowflake-provided and 3rd-party tools and technologies that form the ecosystem for connecting to Snowflake. They also provide detailed installation and usage instructions for using the Snowflake-provided clients, connectors, and drivers.

1. Overview of the Ecosystem
2. Snowflake Partner connect
3. SnowSQL(CLI Client)
4. Snowflake connector for Python
5. Snowflake connector for Spark
6. Node.js Driver
7. Go Snowflake Driver
8. .NET Driver
9. JDBC Driver
10. ODBC Driver
11. Client Considerations

Overview of Ecosystem

Snowflake works with a wide array of industry-leading tools and technologies, enabling you to access Snowflake through an extensive network of connectors, drivers, programming languages, and utilities, including:

- Snowflake-provided client software: SnowSQL (CLI), Python, Node.js, JDBC, ODBC, etc.
- Certified partners who have developed cloud-based and on-premises solutions for connecting to Snowflake through our drivers and connectors.
- Other 3rd-party tools and technologies that are known to work with Snowflake.



Data Integration

Commonly referred to as ETL, data integration encompasses the following primary operations:

Extract: Exporting data from specified source.

Transform: Modifying the source data as needed using rules, merges, lookup tables or other conversion methods to match the target.

Load: Importing the transformed data into target database.

More recent usage references the term ELT, emphasizing that the transformation part of the process does not necessarily need to be performed before loading, particularly in systems such as Snowflake that support transformation during or after loading.

Business Intelligence(BI)

Business intelligence (BI) tools enable analyzing, discovering, and reporting on data to help executives and managers make more informed business decisions. A key component of any BI tool is the ability to deliver data visualization through dashboards, charts, and other graphical output.

Business intelligence also sometimes overlaps with technologies such as data integration/transformation and advanced analytics; however, we've chosen to list these technologies separately in their own categories.

Advanced Analytics

Also referred to as data science, machine learning (ML), artificial intelligence (AI), and “Big Data”, advanced analytics covers a broad category of vendors, tools, and technologies that provide advanced capabilities for statistical and predictive modeling.

These tools and technologies often share some overlapping features and functionality with BI tools; however, they are less focused on analyzing/reporting on past data. Instead, they focus on examining large data sets to discover patterns and uncover useful business information that can be used to predict future trends.

Security and Privacy

Security and privacy tools ensure sensitive data maintained by an organization is protected from inappropriate access and tampering. These tools support a wide range of operations, including risk assessment, intrusion detection/monitoring/notification, data masking, and more.

Snowflake is known to inter-operate with the following security and privacy tools:

Solution	Connection Type	Version / Installation Requirements	Additional Information

DATAGUISE	JDBC	• **Dataguise:** DgSecure
		• **Snowflake:** No requirements
			• Additional reading:
			  ◦ Snowflake Computing (Dataguise website)
			  ◦ Sensitive Data Governance in Snowflake (ebook - Dataguise website)

Programmatic Interfaces

The Snowflake ecosystem supports developing applications using many popular programming languages and development platforms.

Using our client drivers and connectors, Snowflake supports connecting natively through the following languages and platforms:

Solution	Connection Type	Version / Installation Requirements	Additional Information
GO	Go	• **Go Language:** 1.8 (or higher) • **Snowflake:** Go Snowflake Driver — gosnowflake (GitHub)	
JDBC	JDBC	• **Java:** 1.7 (or higher) • **Snowflake:** JDBC Driver — snowflake-jdbc (Maven Central Repository)	
Microsoft .NET	.NET	• **.NET:** Visual Studio 2017 • **Snowflake:** .NET Driver — snowflake-connector-net (GitHub)	
node JS	Node.js	• **Node.js:** 4.0.0 (or higher) • **Snowflake:** Node.js Driver — snowflakedb (install via `npm`)	
ODBC	ODBC	• **C Language:** Requirements are OS-specific • **Snowflake:** ODBC Driver (download from the Snowflake web interface)	
python	Python	• **Python:** 2.x (2.7.9 or higher) or 3.x (3.4.3 or later) • **Snowflake:** Connector for Python — snowflake-connector-python (install via `pip`)	

SQL Editing/Querying Tools

Snowflake provides native SQL editing and querying solutions:

Solution		Additional Information
snowflake	Worksheets >_	• SQL editor directly integrated into the Snowflake web interface with support for multiple, independent working environments. • No installation or configuration required.
snowflake	SnowSQL	• Python-based command line interface (CLI) for performing all tasks in Snowflake, including querying, executing DDL/DML commands, and bulk loading/unloading of data. • Download from the Snowflake web interface and install using provided installer.

Snowflake also works with various 3rd-party SQL editors, including:

Solution	Connection Type	Version / Installation Requirements	Additional Information
	JDBC	• **DBeaver:** 4.3.4 (or higher) • **Snowflake:** JDBC Driver — automatically downloaded and installed by DBeaver	• JDBC driver is also available via snowflake-jdbc (Maven Central Repository) for manual download
SQL Workbench/J	JDBC	• **SQL Workbench:** No requirements • **Snowflake:** JDBC Driver — snowflake-jdbc (Maven Central Repository)	• Additional reading: ◦ Configuring SQL Workbench/J to Use Snowflake (Snowflake Community)

Snowflake Partner Connect

Partner Connect lets you easily create trial accounts with selected Snowflake business partners and integrate these accounts with Snowflake. This feature provides a convenient option for trying additional tools and services, and then adopting the ones that best meet your business needs.

1. Supported Partners
2. Security Requirements
3. Connecting with a Snowflake Partner
4. Security Roles, Users, and Objects Created by the Partner
5. Whitelisting Partner IP Addresses
6. Launching Partner Applications
7. Disconnecting from a Partner Account
8. Troubleshooting a Connection
9. Connection Already Exists

Supported Partners

Currently, Partner Connect supports the following partner services:

Data Integration:

- Alooma
- Fivetran
- Stitch

Business Intelligence(BI):

- Periscope Data
- Sigma Computing

Security Requirements

Partner Connect is limited to account administrators (i.e. users with the ACCOUNTADMIN role) who have a verified email address in Snowflake:

- To use Partner Connect, you must switch to the ACCOUNTADMIN role or contact someone in your organization who has that role.
- To verify your email address:

1. In the Snowflake web interface, select **Preferences** » **General** from the dropdown menu next to your login name. The **Preferences** page appears.
2. In the **User Information** area, add or verify your email address by clicking the appropriate link(s) in the **Email Address** field.

Connecting with Snowflake Partner

To initiate a trial account with any Snowflake partner currently included in Partner Connect:

- Log into the Snowflake web interface.
- Click the dropdown menu next to your login name » **Switch Role** » **ACCOUNTADMIN** to change to the account administrator role.
- Click the **Partner Connect** tab. The snowflake partner connect page opens.
- To read a description of an application and its requirements, click on the corresponding partner button.
- Click the **Connect** button below the partner description to initiate creating a trial account with the partner and connecting the partner application to Snowflake.

Security Roles, Users and Object created by the partner

During the connection process, the partner application creates the following objects:

Object	Name	Notes
User	PC_<*partner_name*>_USER	
Role	PC_<*partner_name*>_ROLE	The PUBLIC role is granted to this role, which enables the role to access any objects owned/granted to the PUBLIC role. In addition, this role is granted to the SYSADMIN role, which enables users with the SYSADMIN or ACCOUNTADMIN role to also access any Snowflake objects created for partner application access.
Database	PC_<*partner_name*>_DB	
Warehouse	PC_<*partner_name*>_WH	

Whitelisting Partner IP Address

If you use Network Policies to restrict access to your Snowflake account based on user IP address, you must whitelist the partner IP addresses (i.e. add them to the list of allowed IP addresses for the active network policy for your account).

The following table lists the IP addresses for each partner or provides links to pages on the partner sites for this information:

Snowflake Partner	IP Addresses	Notes
Alooma	Various	See the Alooma Support site for the addresses.
Fivetran	`52.0.2.4`	For more setup details, see the Fivetran Documentation.
Periscope Data	Various	See the Periscope Data Documentation for addresses.
Sigma Computing	`104.197.169.18` , `104.197.193.23`	
Stitch	Various	See Stitch Documentation for addresses.

Disconnecting from Partner Account

If you decide to discontinue a trial account initiated through Partner Connect for any reason, complete the following steps:

- Log into the Snowflake web interface.
- Click the user menu » **Switch Role** » **ACCOUNTADMIN** to change to the account administrator role.
- **Click the Partner Connect tab. The snowflake partner connect page opens.**
- Click on the button for the partner application you are disconnecting from. Note the names of the database, warehouse, system user, and role objects created for the partner application.
- Use the corresponding DROP <object> command to drop each of the objects created by the partner connection.
- Contact Snowflake Support to stop the partner account trial and remove the checkmark from the **Partner Connect** button.

SnowSQL(CLI Client)

SnowSQL is the next-generation command line client for connecting to Snowflake to execute SQL queries and perform all DDL and DML operations, including loading data into and unloading data out of database tables.

Snowflake provides platform-specific versions of SnowSQL for download for the following platforms:

Platform	Supported Versions
Microsoft Windows (64-bit)	• Windows 7 or higher • Windows Server 2008 R2 or higher
Mac OS	• OS X v10.12 or higher
Linux (64-bit)	• CentOS 6 or higher • Ubuntu 14 or higher

Loading Data into Snowflake

This topic describes concepts related to loading data into Snowflake tables.

- Bulk Loading Using COPY
  - Data Loading Process
  - Tasks for Loading Data
- Continuous Loading Using Snowpipe

Bulk Loading Using COPY

This section describes bulk data loading into Snowflake tables using the COPY INTO <table> command. The information is similar regardless if you are loading from data files on your local file system or in Amazon S3 buckets or Microsoft Azure containers.

Data Loading Process:

- Upload (i.e. stage) one or more data files into either an internal stage (i.e. within Snowflake) or an external location:

Internal:	Use the PUT command to stage the files.
External:	Currently, Amazon S3 and Microsoft Azure are the only services supported for staging external data. Snowflake assumes the files have already been staged in one of these locations. If they haven't been staged already, use the upload interfaces/utilities provided by the service that hosts the location.

- Use the COPY INTO <table> command to load the contents of the staged file(s) into a Snowflake database table.

This step requires a running virtual warehouse that is also the current warehouse (i.e. in use) for the session. The warehouse provides the compute resources to perform the actual insertion of rows into the table.

Task for Loading Data

- Bulk Loading from a Local File System:

This set of topics describes how to use the COPY command to bulk load from a local file system Into tables.

As illustrated in the diagram below, loading data from a local file system is performed in two, separate steps:

Step 1:	Upload (i.e. stage) one or more data files to a Snowflake stage (named internal stage or table/user stage) using the PUT command.
Step 2:	Use the COPY INTO <table> command to load the contents of the staged file(s) into a Snowflake database table.

Bulk Data Loading Diagram