

Snowflake Interview Questions and Answers

Q1) What is Snowflake?

Ans. Snowflake is the Cloud Data Warehouse Platform built for the cloud for storing and retrieving data. It enables data storage, processing, and analytic solutions that are faster, easier to use, and far more flexible than traditional offerings. It eliminates all your data silos and runs your workloads from a single platform.

Q2) What is Snowflake good for?

Ans. Snowflake is good for its architecture and data sharing capabilities. Snowflake architecture enables you to store, compute, and scale independently. Its data sharing functionalities make it easy to share, govern and secure data in real-time across organizations.

Q3) Why is Snowflake so popular?

Ans. There are a lot of reasons for Snowflake gaining momentum these days. Let's look at them one by one

- Snowflake serves a wide range of technology areas, including business intelligence, data integration, security & governance, and advanced analytics.
- It provides cloud infrastructure and supports modern design architectures suitable for agile and dynamic usage trends.
- Supports out-of-the-box features like data sharing, on-the-fly scalable compute, data cloning, separation of storage and compute, third-party tools access, etc.
- Snowflake simplifies data processing.
- Fits for many use cases - ODS with staged data, data lakes with raw data, and data marts/data warehouse with presentable and modeled data.
- Snowflake offers scalable computing power.

Q4) Is Snowflake OLAP or OLTP?

Ans. Snowflake is designed for OLAP (Online Analytical Processing) database system. Depending on the usage, you can use it for OLTP (Online Transaction Processing) purposes as well.

Q5) Does Snowflake support stored procedures?

Ans. Yes, Snowflake supports stored procedures. A stored procedure is similar to a function; as such, it's created once and used many times. Using the CREATE PROCEDURE command, you can make it, and with a CALL command, you can execute it.

Snowflake storage procedures are written in JavaScript API (in the form of JavaScript methods and objects). These APIs allow stored procedures to execute database operations such as CREATE, SELECT, and UPDATE.

Enroll for Best [Snowflake Certification Course from MindMajix](#)

Designed & Certified by Industry Experts - Practical Learning Path - Live Projects - Career Guidance - Job Assistance

Q6) What is Snowflake ETL?

Ans. ETL is an acronym for Extract, Transform, and Load. It's a process used to extract data from one or more sources and load it to a specified data warehouse or database. The sources may include third-party apps, flat files, databases, and more.

Snowflake ETL means applying the ETL process to load data into the Snowflake database/data warehouse. That includes extracting data from data sources, making required transformations and then loading it into Snowflake.

Q7) What database does Snowflake use?

Ans. Snowflake is a cloud-based Data Warehouse solution presented as a SaaS (Software-as-a-Service) with full support for ANSI SQL.

Q8) How do you execute a snowflake procedure?

Ans. Stored Procedures enable you to create modular code containing complex business logic by including different SQL statements with procedural logic.

- To execute Snowflake Procedure, perform the following steps:
- Execute a SQL statement.
- Retrieve the results of a query
- Retrieve result set metadata

Q9) How do you make a Snowflake task?

Ans. To make a Snowflake task, you need to use the CREATE TASK command.

The following needs to be done to use this command:

- CREATE TASK on the schema.
- USAGE on the warehouse in the task definition.
- Execute stored procedure or SQL statement in the task definition.

Q10) How is Snowflake different from redshift?

Ans. Both Snowflake and Redshift offer on-demand pricing but differ in package features. Snowflake separates compute usage from storage in its pricing structure, while Redshift combines both.

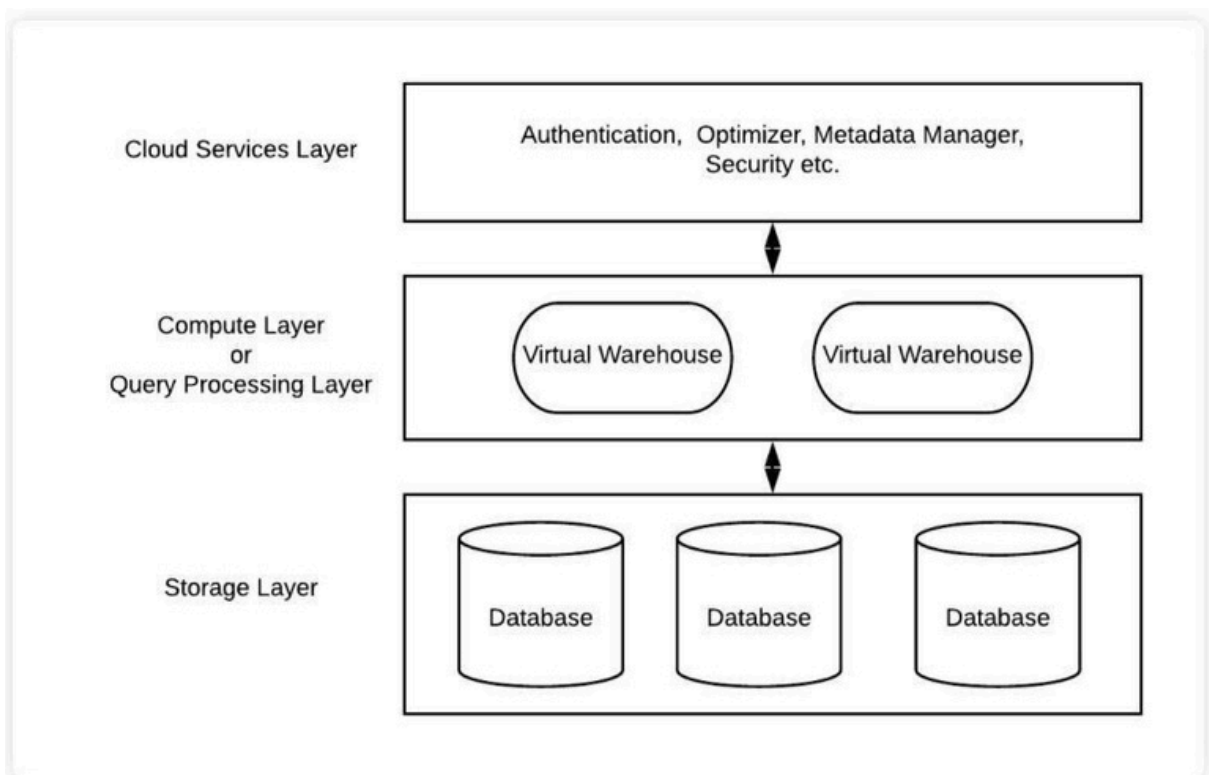
Q11) How do you check Snowflake history?

Ans. To retrieve the task history details for runs in a scheduled or executing state, query the TASK_HISTORY table function in the Information Schema.

Q12) Explain Snowflake Architecture

Ans. Snowflake supports a high-level architecture, as depicted in the diagram below.

- Snowflake's architecture is a hybrid of traditional shared-nothing and shared-disk database architectures. Like shared-nothing architecture, Snowflake processes the queries with massively parallel processing (MPP) compute clusters, and each cluster's node stores locally a part of the complete data set. Like shared-disk architecture, it uses a central data repository to create data accessible in the platform from all compute nodes.
- On the whole, snowflake architecture offers the scale-out and performance of a shared-nothing architecture, along with the data management of a shared-disk architecture.
- Snowflake architecture consists of three layers - Database Storage, Cloud Services, and Query processing.



Q13) What are the three layers of Snowflake architecture?

Ans. The three layers of Snowflake architecture include:

Database storage - In Snowflake, it reorganizes into its internal optimized, columnar, and compressed format when data is loaded. This data is stored in cloud storage.

Query Processing - Queries are executed in the processing layer and are processed using “virtual warehouses.”

Cloud Services - It's a collection of services that coordinate activities across Snowflake. Services included in this are authentication, metadata management, infrastructure management, access control, and Query parsing & optimization.

Q14) Who are the competitors to Snowflake?

Ans. Top Snowflake competitors are AWS, Google, Microsoft, Cloudera, IBM, SAP, and Teradata.

Snowflake Advanced Interview Questions

Snowflake Developer Interview Questions

Q15) Can you explain how Snowflake is different from AWS?

Ans. Snowflake provides compute and storage separately, and storage cost is the same as storing the data. AWS addresses this issue by adding Redshift Spectrum, which allows data querying directly on S3, but not as seamless as Snowflake.

Q16) What are the unique features of Snowflake?

Ans. Unique features of Snowflake Cloud Data Warehouse

- Unique Cloud Architecture
- Allows you to query semi-structured JSON/XML data
- Database and Object Cloning
- Support for XML
- Data Sharing
- Hive metastore integration
- Supports geospatial data
- Recovery table using undrop
- Security and data protection
- Result Caching
- Auto optimization
- Snowpipe
- Change Data Capture (CDC) using Streams and Tasks.
- Continuous data protection

Q17) What does Snowflake computing do?

Ans. Snowflake gives a cloud-based data storage and analytics service. It is termed "data warehouse-as-a-service," which allows you to analyze and store data using cloud-based software and hardware.

Snowflake Computing provides secure and governed access to the complete data network and a core architecture to allow various types of data workloads, including a single platform to develop modern data applications.

Q18) What is a snowflake in AWS?

Ans. Snowflake is a cloud data warehouse built on top of the AWS (Amazon Web Services) cloud infrastructure and is a true SaaS offering. There is no software or hardware (physical or virtual) for you to select, install, configure, or manage.

Q19) What is the schema in Snowflake?

Ans. Schema is used to organize data in Snowflake. It's a logical grouping of database objects (views, tables, etc.), and each schema belongs to a single database.

Q20) What is the difference between star schema and snowflake schema?

Ans. Both Star and snowflake schemas are similar, but the difference is in dimensions. In a snowflake, at least a few dimensions are normalized, while in a star schema, logical dimensions are denormalized into tables.

Q21) What are the cloud platforms currently supported by Snowflake?

Ans. Supported data platforms for Snowflake

- Microsoft Azure (Azure)
- Amazon Web Services (AWS)
- Google Cloud Platform (GCP)

Q22) What are the advantages of snowflake schema?

Ans.

Less disk space is used than in a denormalized model.

Better data quality (data is more structured, so data integrity problems are reduced)

Q23) What are the best ETL tools to use with Snowflake?

Ans. List of Best ETL tools for Snowflake

- Blendo
- StreamSets
- Apache Airflow
- Matillion
- Hevo Data
- Etleap

Q24) Do snowflakes index?

Ans. No, Snowflake does not use indexes. This is one of the things that makes Snowflake scale so well for arbitrary queries.

Q25) Where is metadata stored in Snowflake?

Ans. Snowflake automatically generates metadata for files in internal stages or external stages. It is stored in virtual columns and can be queried using a standard SELECT statement.

Q26) What are the benefits of using Snowflake?

Ans. Benefits of Using Snowflake

- It is an encrypted and secure interface.
- The data available on Snowflake is durable and reliable.
- The Snowflake interface offers high flexibility, elasticity, accessibility, and value.
- This platform's elastic nature allows you to scale up the virtual warehouse to load the data faster.
- Snowflake allows you to query against large data sets.
- A fully automated platform.
- Cost-effective

Q27) What are the different types of caching available in Snowflake?

Ans. Different types of caching in Snowflake

- Query Results Caching
- Virtual Warehouse Local Disk Caching
- Metadata Cache

Q28) What is Time Travel in Snowflake?

Ans. Time travel enables you to access data of past times. For example, if you have a Student table and delete it accidentally, you can retrieve it using time travel.

Q29) Is there a cost associated with Time Travel in Snowflake?

Ans. Yes, Storage charges are incurred to maintain historical data during both the Time Travel and Fail-safe periods.

Q30) What are the different ways to access the Snowflake Cloud Datawarehouse?

Ans. Various ways to access Snowflake Cloud Data Warehouse

- Web User Interface

- JDBC Drivers
- ODBC Drivers
- Python Libraries
- SnowSQL Command-line Client

Q31) What is Snowflake Caching?

Snowflake caching is one of the advanced techniques in data management system. That means if a new query published, this platform also checked the previous query for effective data driven results. In the database system, if any database query is matched, it cached and stored for future predictions & process