Snowflake Inc. is a cloud-based data-warehousing startup that was founded in 2012.

Based in San Mateo, California

Snowflake offers a cloud-based data storage and analytics service, generally termed "data warehouse-as-a-service"

Runs on Amazon S3 since 2014

On Microsoft Azure since 2018

Rolled out on Google Cloud Platform in 2019

# SNOWFLAKE RECOGNIZED AS A LEADER BY GARTNER: THIRD CONSECUTIVE YEAR POSITIONED IN THE MAGIC QUADRANT REPORT

# What makes Snowflake unique ?

### Break free from the past

We designed and implemented a new, unique architecture to stay clear from the limitations of existing architecture and software offerings.

### Design for the cloud

On-premise solutions are too complex, too expensive and getting more so each day. We developed Snowflake as a service from the ground up for the cloud, capitalizing on the leading architecture and technology of the public cloud.

### Support modern data and applications

People want to focus on deriving insight from data – not configuring, tuning, and managing a data platform. Rather than just deliver traditional software hosted in the cloud, we deliver a data platform as a service.


# Snowflake Architecture: Data Storage Processing & Cloud services

# Snowflake Architecture

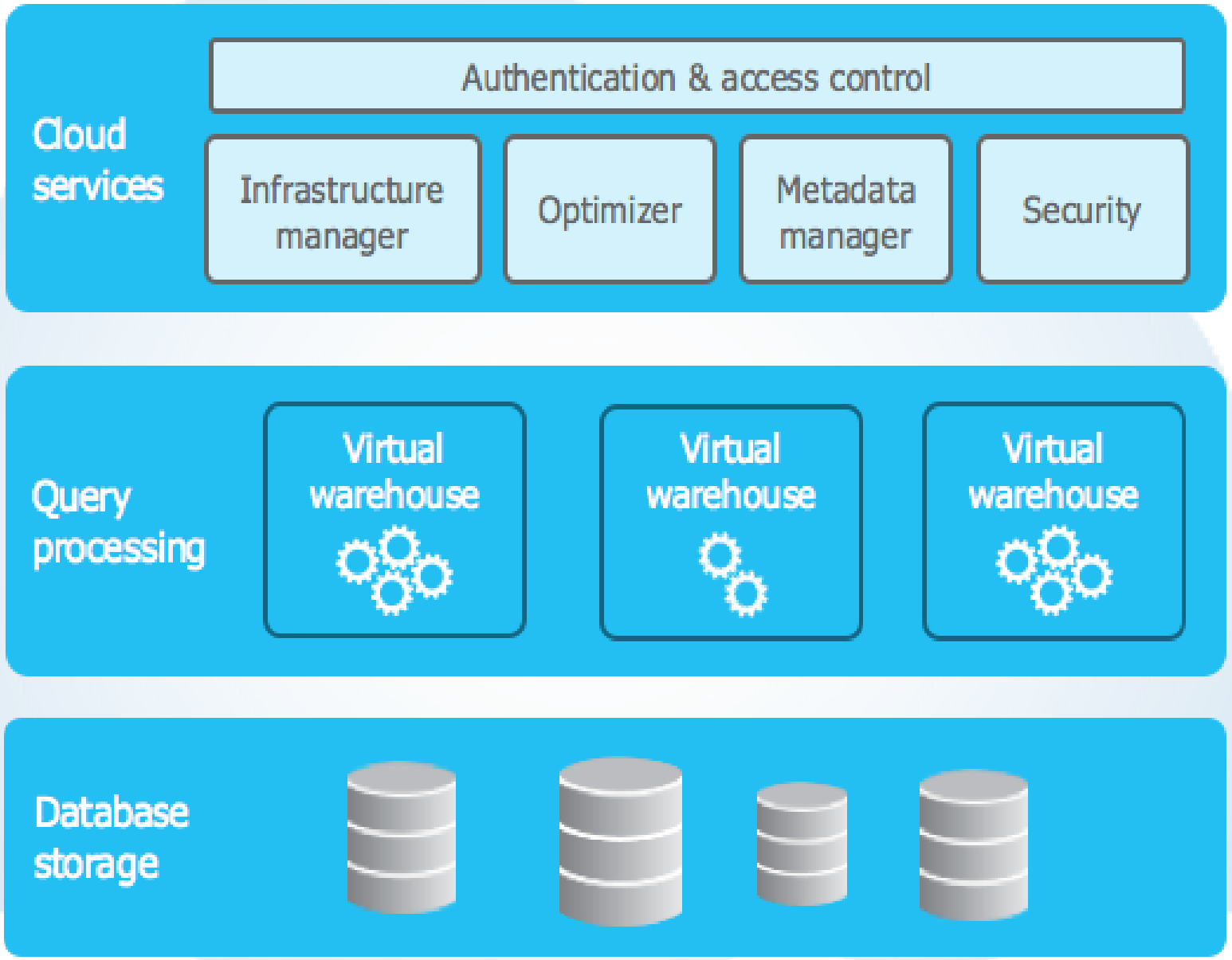
Hybrid of

- **Shared-disk database architectures**

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

- **Shared-nothing database architectures**

Snowflake processes queries using MPP (massively parallel processing) compute clusters where each node in the cluster stores a portion of the entire data set locally




**Cloud Services**

The cloud services layer is a collection of services that coordinate activities across Snowflake

**Query Processing**

Query execution is performed in the processing layer. Snowflake processes queries using "virtual warehouses"

**Database Storage**

When data is loaded into Snowflake, Snowflake reorganizes that data into its internal optimized, compressed, columnar format. Snowflake stores this optimized data in cloud storage

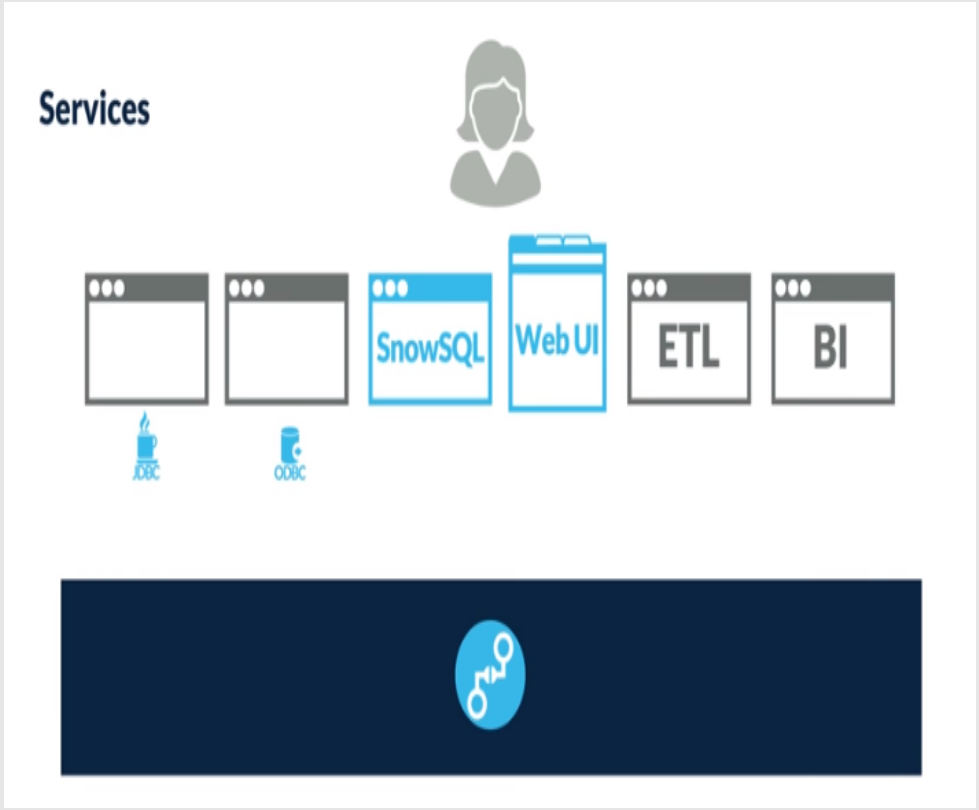
# Storage

- Structured Relational data store in Datatypes:
  - Varchar
  - Number
  - Boolean
  - Timestamp etc.
- Semi Structured Data (Avro, JSON, Parquet) stored in
  - Variant

# Compute

# Services

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




# Supported cloud platforms

Amazon Web Services (AWS)

Google Cloud Platform (GCP)

Microsoft Azure (Azure)

# QUIZ

# Data Storage is independent of compute

- True
- False

Answer : True

# DataBase Storage Layer manages

a. VWH
b. Metadata
c. Query Processing
d. MPP

Answer : Metadata

# NoSQL is example of

a. Share Disk Architecture
b. Share Nothing Architecture

Answer : Share Nothing Architecture

Which of the following is true about SF Architecture

a. It is Share Disk Architecture
b. It is Share Nothing Architecture
c. Uses VWH for MPP
d. All of the above

Answer : all of the above

# Recap

**Cloud Database :** Database delivered in a public cloud as a managed service that is optimized for analytics, scale and ease of use

Snowflake Inc. is a **cloud-based data-warehousing** startup that was founded in 2012.

Snowflake recognized as a leader by **Gartner**

**Snowflake Uniqueness** :
Break from Past
Designed for cloud
Support Modern Data and Applications

**Snowflake Architecture:**
Hybrid of Shared disk database architecture and Shared nothing database architecture
3 Layers: Database Storage, Query Processing, Cloud Services