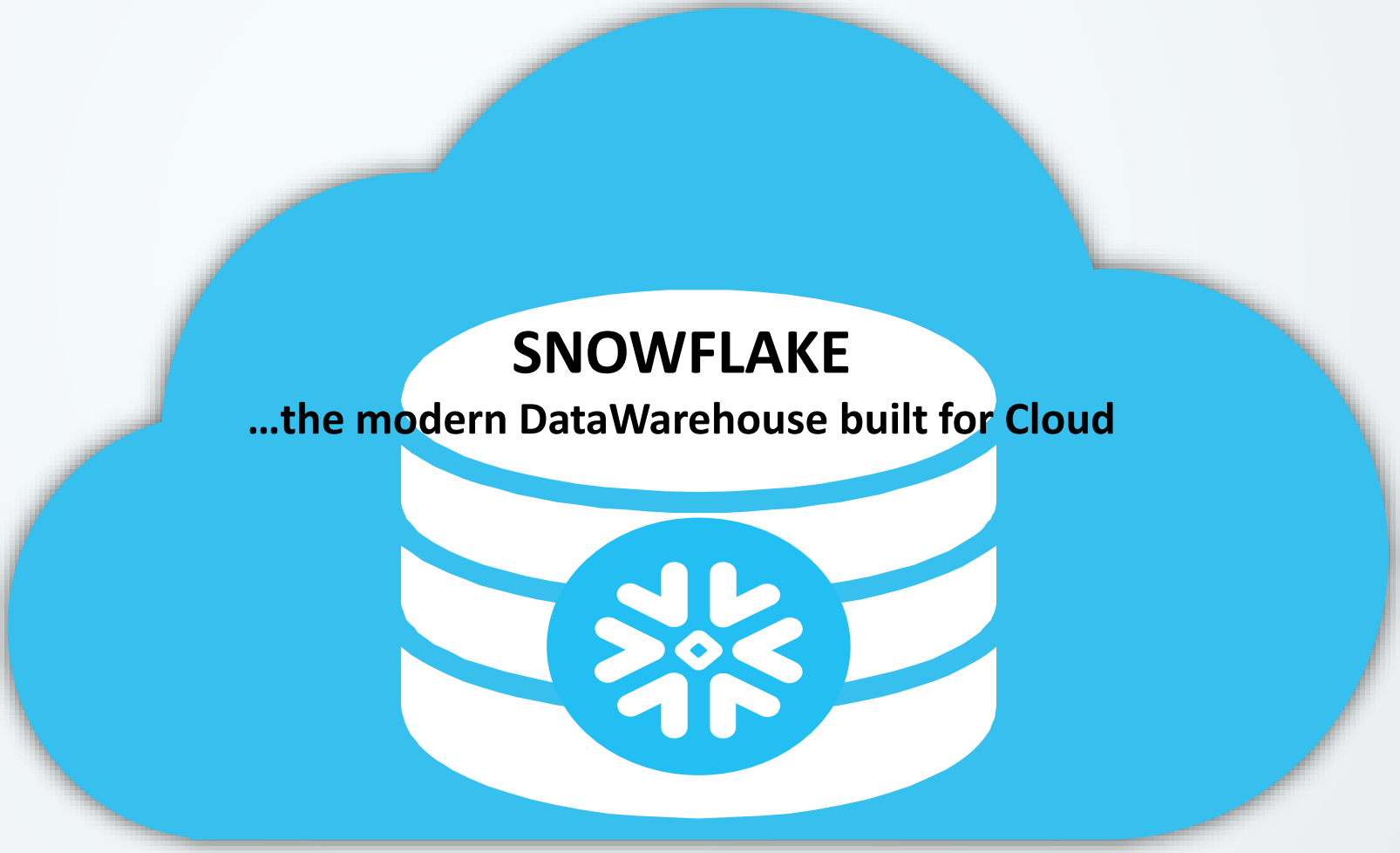


“Simplicity is the ultimate sophistication”

Leonardo da Vinci

Day-2

Introduction to Snowflake

Structured + semi-structured data

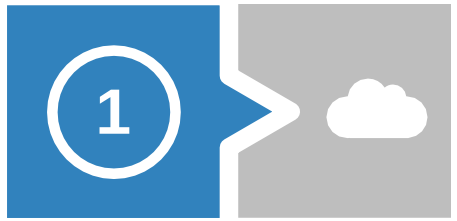
Instant, unlimited scalability

No management burden

Live data sharing

Unique architecture:
Multi-cluster, shared data

Introduction to Snowflake



Built for the Cloud

Built from scratch, optimized for cloud, storage & compute is decoupled

Software as a Service

No software, infrastructure or upgrades to manage

Pay only for used compute & storage

Storage & compute charged independently, only for use

Scalable

Virtual warehouse enable compute scaling

Introduction to Snowflake

Traditional Data Warehouse & Big Data Platforms

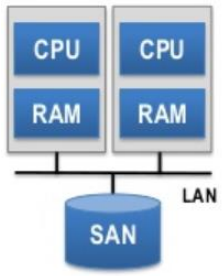
- Capacity planning & migration
- Storage & server management
- Index, partition, and sort keys
- Backup, failover, disaster recovery
- Security management

Snowflake

- No infrastructure management
- No knobs to tune
- Built-in resiliency and data protection
- Built-in, enterprise-grade security

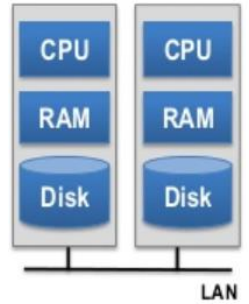
Introduction to Snowflake

Traditional Distributed Architecture



Shared Disk

- In shared disk architecture the nodes share memory as well as the storage
- disks have active nodes which are shared in case of failures.
- The hardware in shared disk is comparatively expensive
- The data is not partitioned
- Major advantage is that it is highly available
- Good for OLTP



Shared Nothing

- In shared nothing architecture the nodes do not share memory or storage
- disks have individual nodes which cannot be shared
- It has cheaper hardware as compared to shared disk architecture.
- The data is strictly partitioned
- Major advantage is that it is highly scalable.
- Good for DSS.

Introduction to Snowflake

Traditional Architectures

Shared-disk

Shared storage

Single cluster

Shared-nothing

Decentralised, local storage

Single cluster

Snowflake

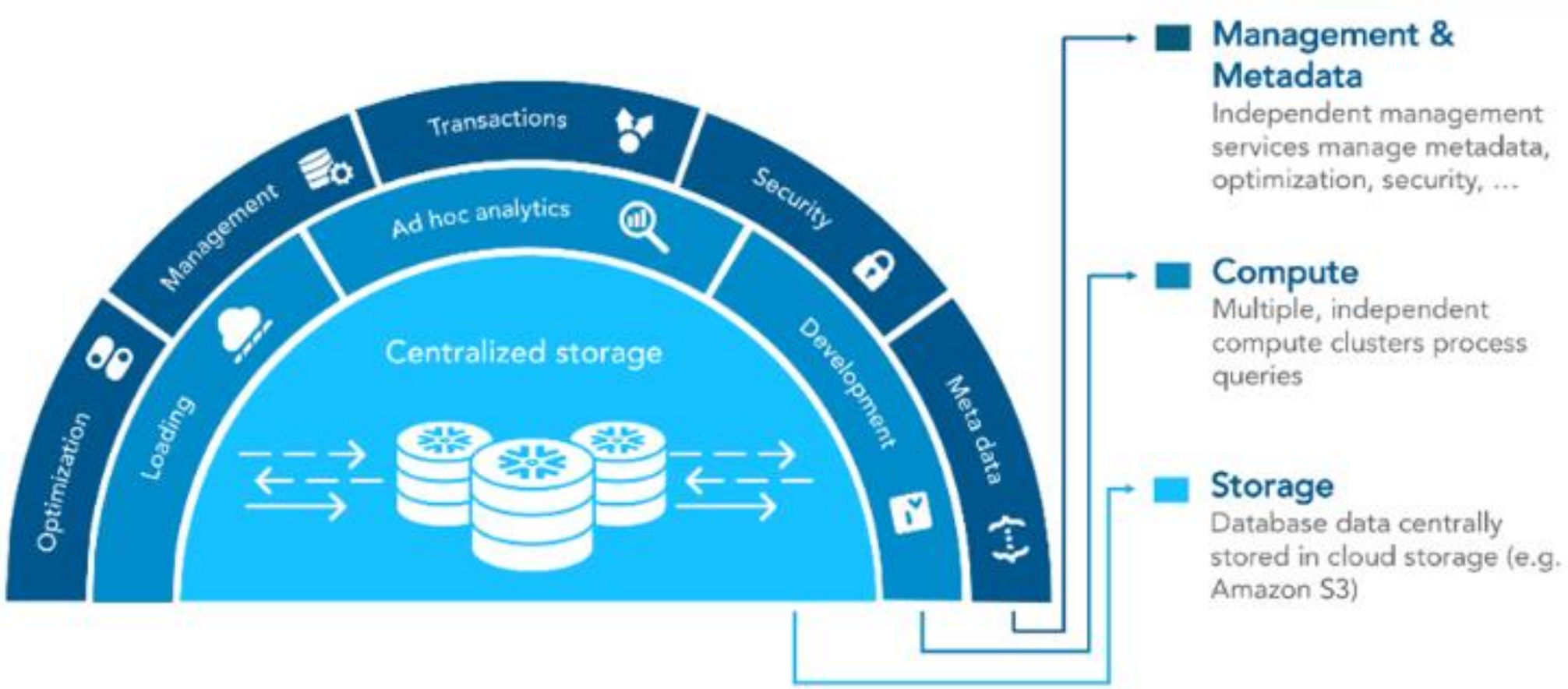
Multi-cluster, shared data

Centralised, scale-out storage

Multiple, independent compute clusters

Snowflake Architecture Overview

The Snowflake Elastic Data Warehouse has a unique architecture that delivers the power of data warehousing, the flexibility of big data platforms and the elasticity of the cloud – at a fraction of the cost of other solutions



Snowflake Architecture

Storage
(cloud storage services)

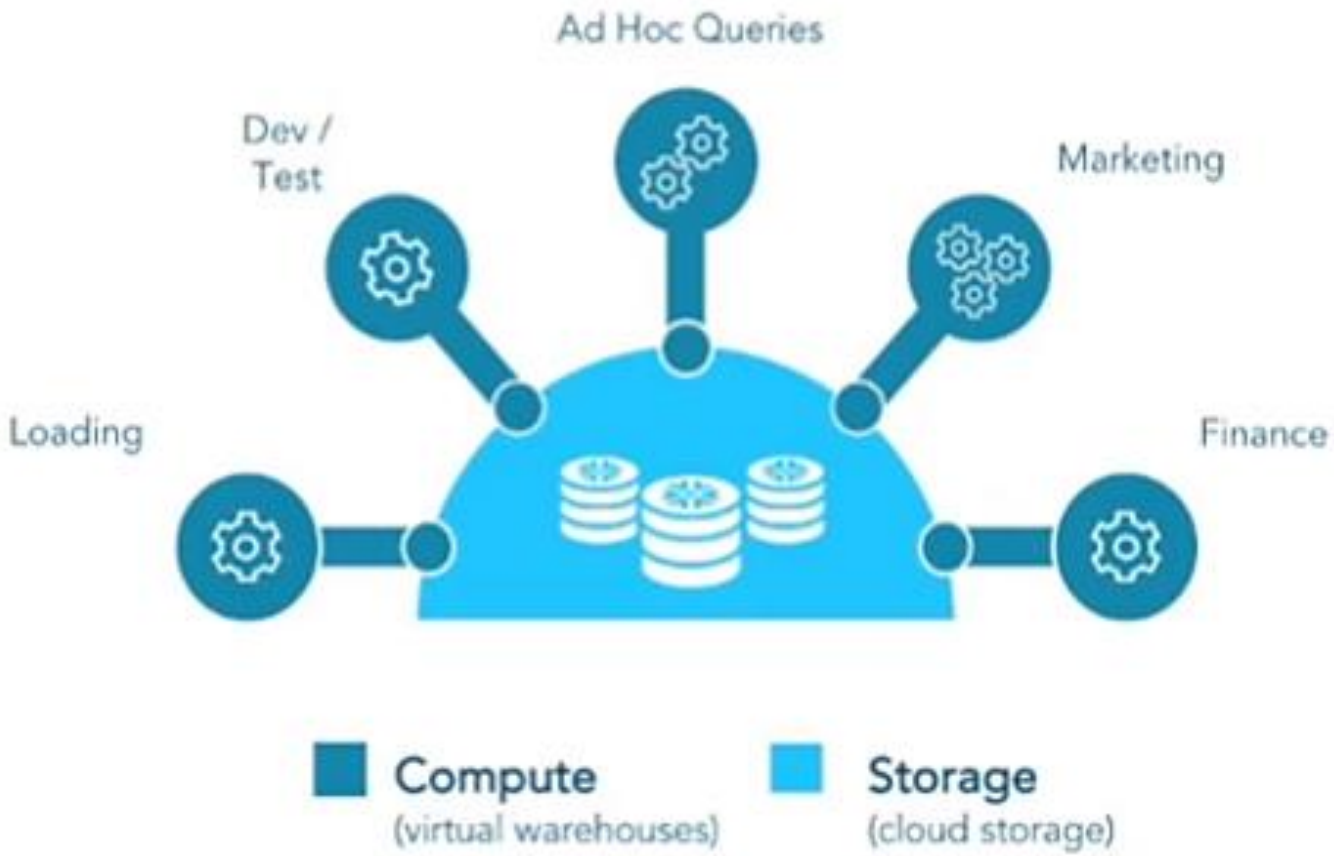
Decouple compute from storage

- One place for data
- Storage scales without scaling compute
- Only pay for compute when used

Centralize data in cloud storage

- Near-infinite capacity
- Automatic, transparent scaling
- Built-in replication & redundancy
- Very low cost: <$25 / compressed TB / mo

Snowflake Architecture



Compute Cluster AKA Virtual Warehouse (VW) – Compute Resources (RAM, CPU, SSD)

- Scale Out and In on the Fly
- Scale Up and Down in Seconds

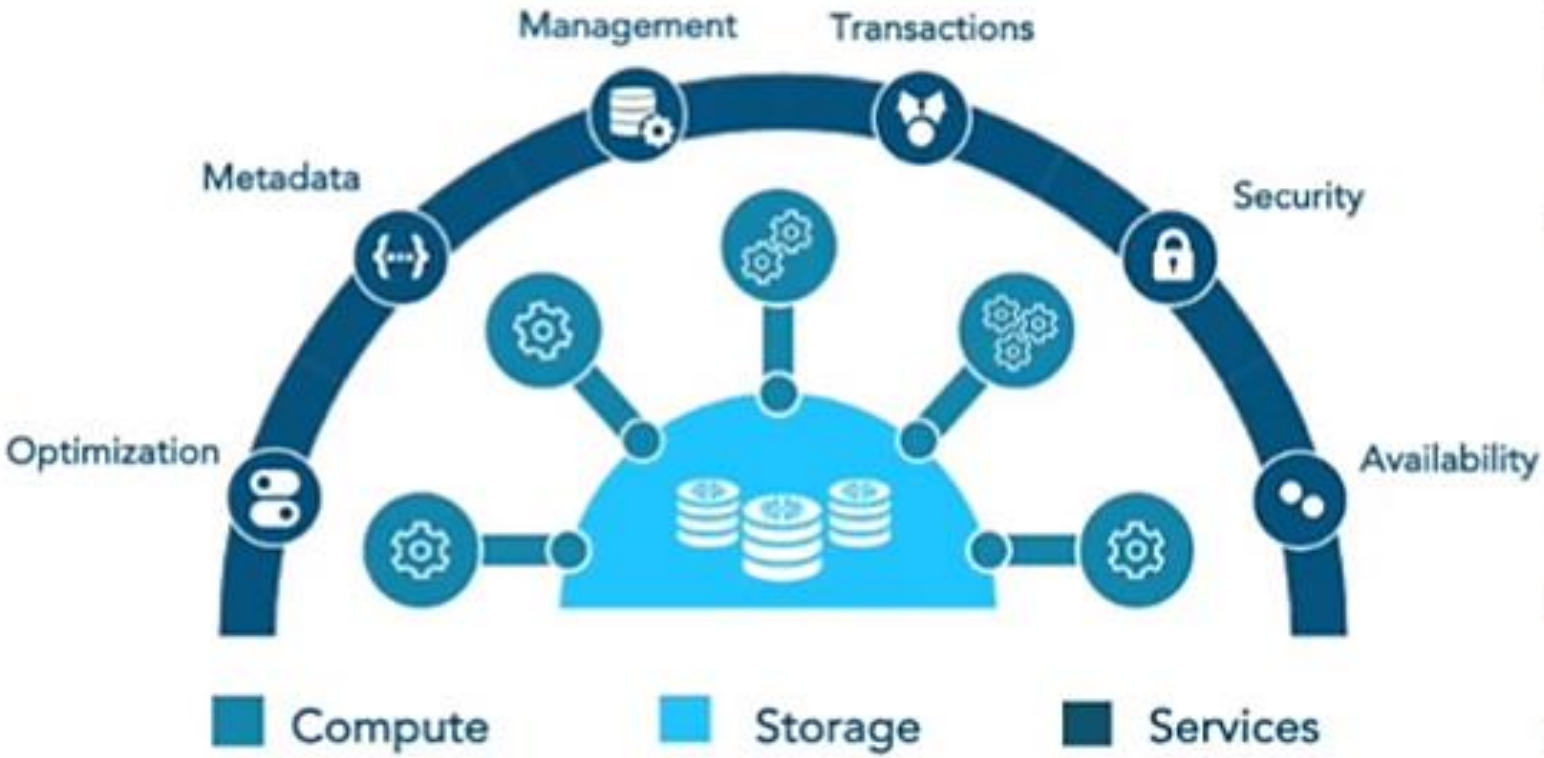
Multiple Independent Cluster

- No Resource Contention
- Load & Query Simultaneously

3 Types of Compute

- Compute Provisioned by you
  - VW
- Compute Provisioned by Snowflake
  - For Serverless features
  - Cloud Services Compute

Snowflake Architecture



Single service

- Simple: all data & workloads in one system
- Complete: centrally controlled infrastructure, monitoring, security, availability

Built-in intelligence

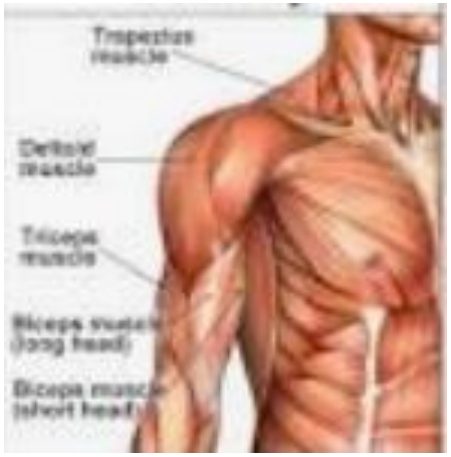
- Consistent: ACID transactions across all users and workloads
- Optimized: system-wide metadata and optimization

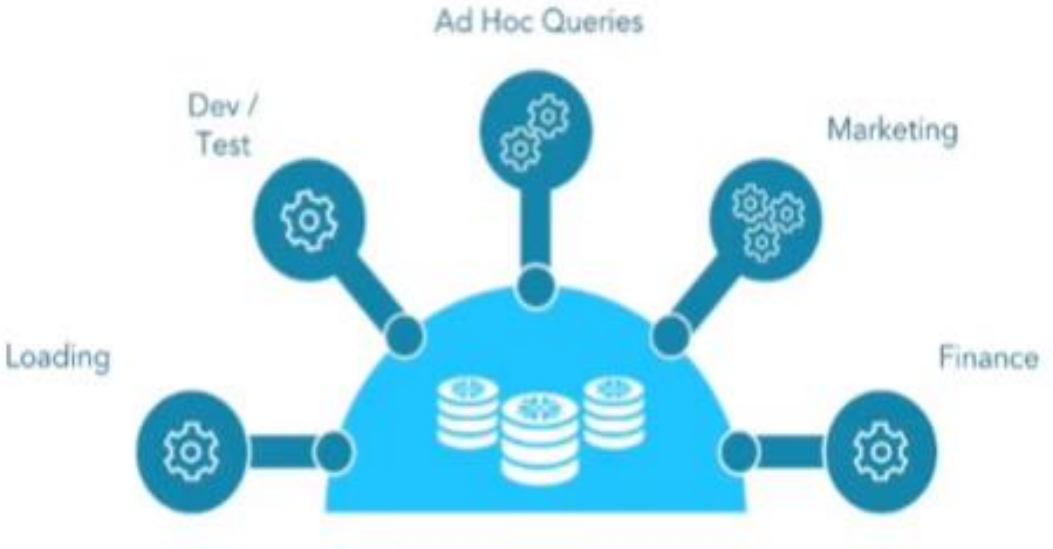
Snowflake Architecture- Storage

- Decouple storage from compute
- One place for all enterprise data
- Data stored in compressed format of choice e.g. gzip, bz2, Brotli, Zstd, Deflate etc.
- No need for provisioning of storage in advance
- Near infinite capacity
- Automatic and transparent scaling
- Built-in replication and redundancy
- Utilizes cheap storage options i.e. AWS S3, Azure BLOB, GCP storage bucket
- Storage cost based on chosen platform and region
- Encryption at rest by default. Annual rekey of all encrypted data

Snowflake Architecture- Compute

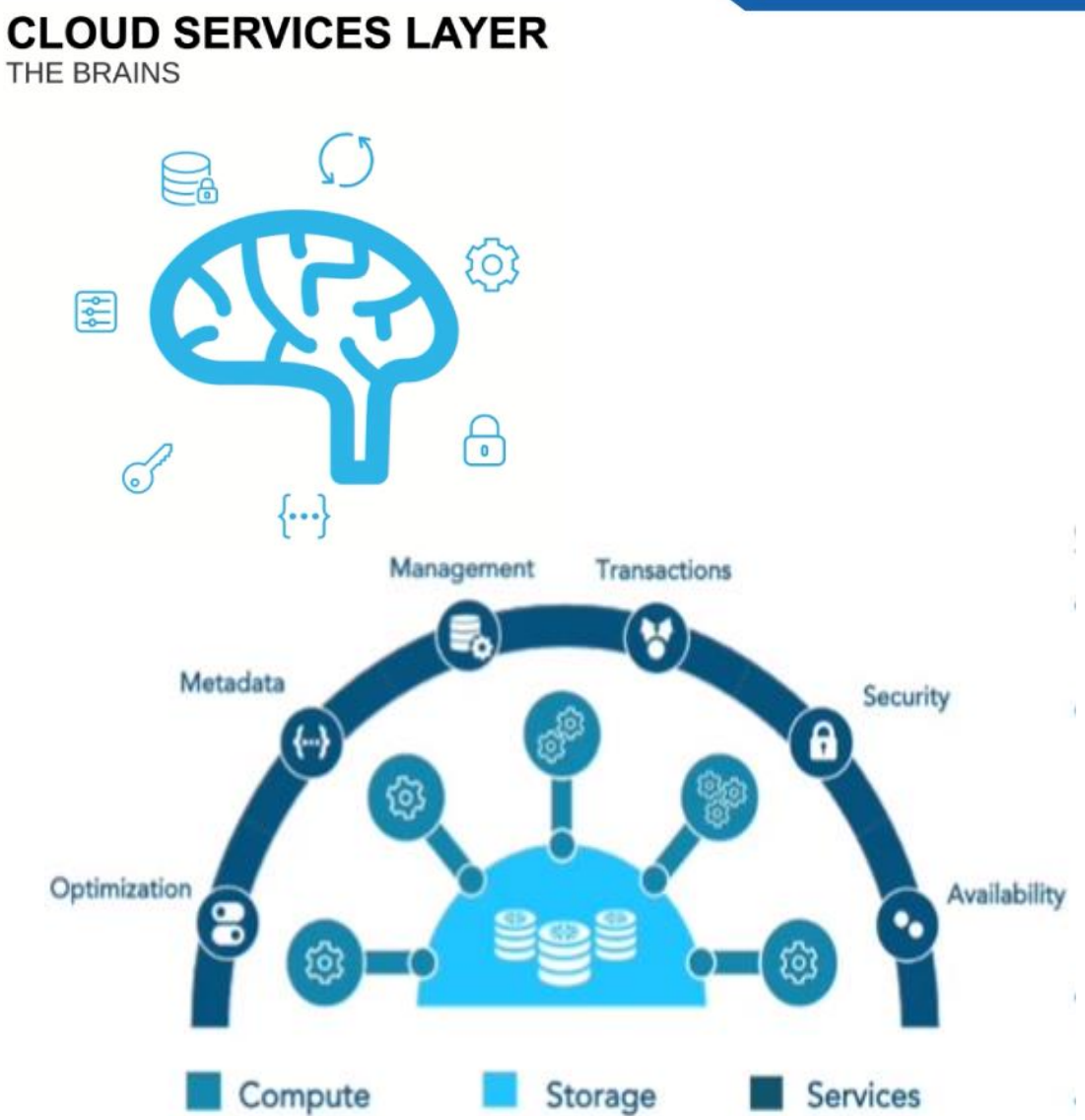
Compute Layer - Muscle





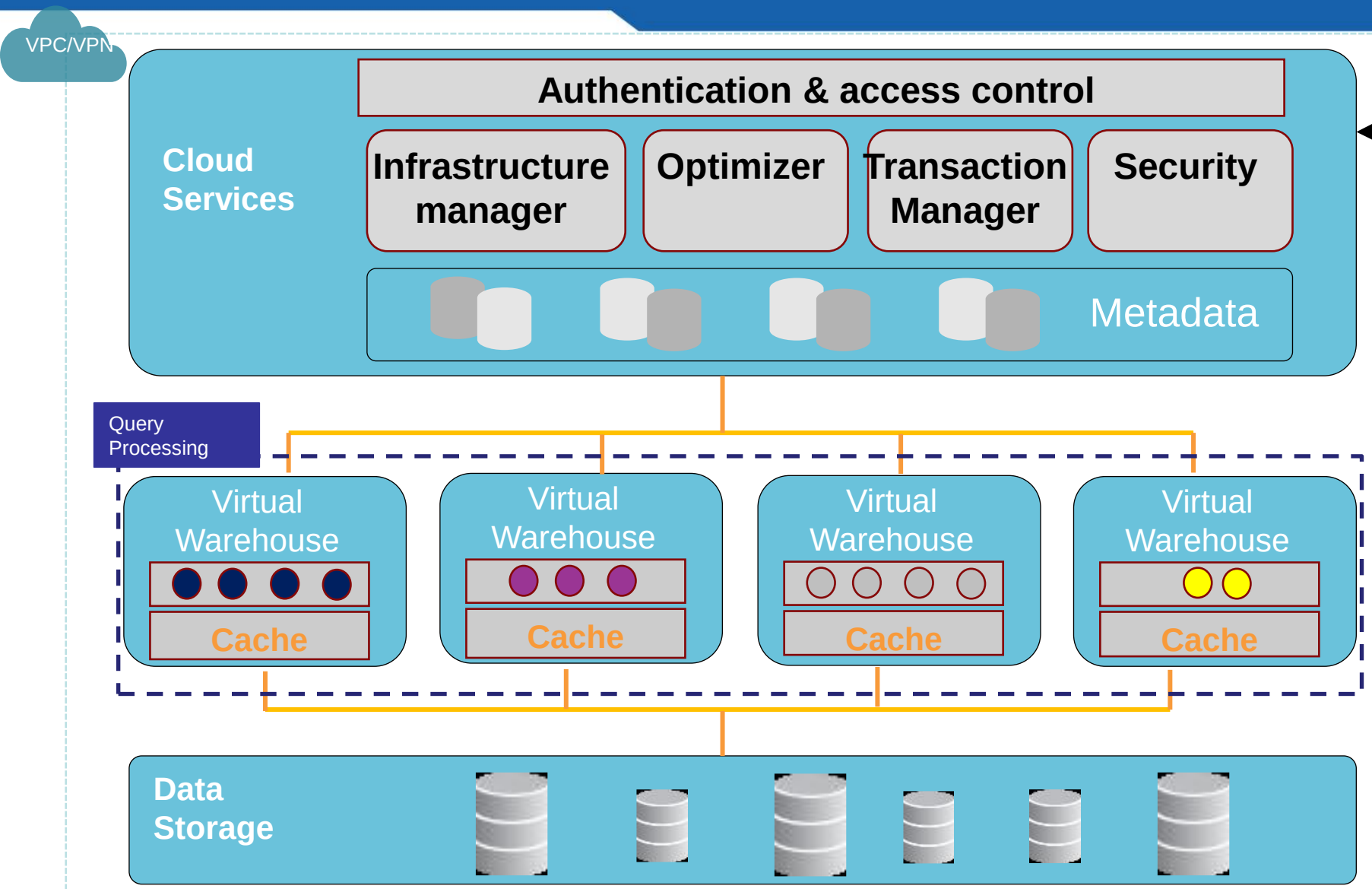
- Muscle of snowflake aka virtual warehouses (VW)
- Scale OUT and IN automatically- horizontal scaling
- Scale UP and DOWN on the fly- vertical scaling
- Auto resume in need, auto suspend when not-in-use
- No resource contention between VWs
- Pay only when in use, per second billing
- Available in standard sizes XS,S,M,L,XL,2XL,3XL,4XL, 5XL, 6XL
- Query performance scales linearly with VW size
- Compute cost based on chosen edition, platform and region
- Virtually unlimited number of VWs can be created

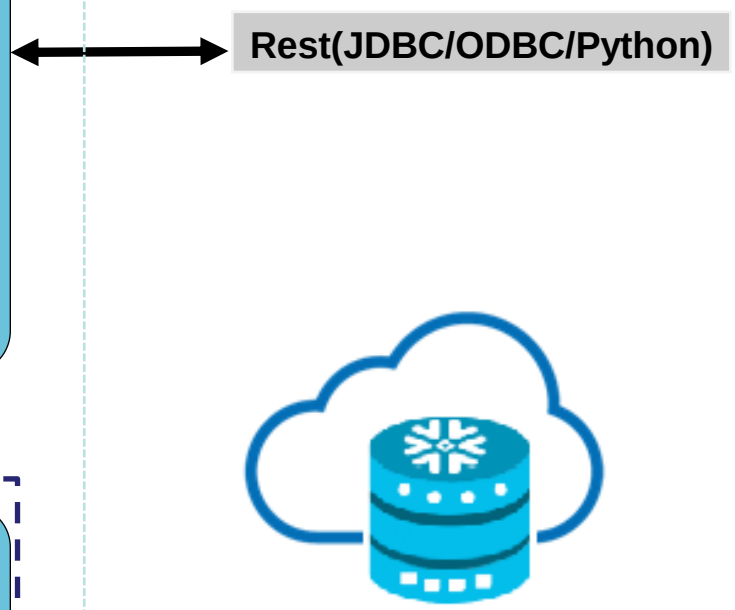
Snowflake Architecture- Service



- Brain of snowflake aka Cloud Services Layer
- Centrally controlled infrastructure
- Metadata storage and optimization
- Query parsing and optimization
- Authentication & Access control
- Enforces ACID transactions across all workloads
- Billing for only excess of 10% of your compute credit usage each day
- Management (Cache and Replication Management)

Snowflake – Three Layers





- All data in one place
- Independently scale every layer
- Every virtual warehouse can access all data