SNOWFLAKE DATA CLOUD

SHENZHEN ARCHSUMMIT 2021

Jiaqi Yan
Principal Software Engineer

QUICK HISTORY

Founded
August 2012

Founding Team
January 2013

aws

2015 – GA

Azure

2018 – Azure port

2020 – GCP port

3,500+ active customers,
2000+ Employees
"The Data Cloud" launched
and IPO 2020

1B+ queries daily

300+ PB total storage (compressed),
biggest table 68TN rows

1.2k data providers

NPS - 71, industry
average 21
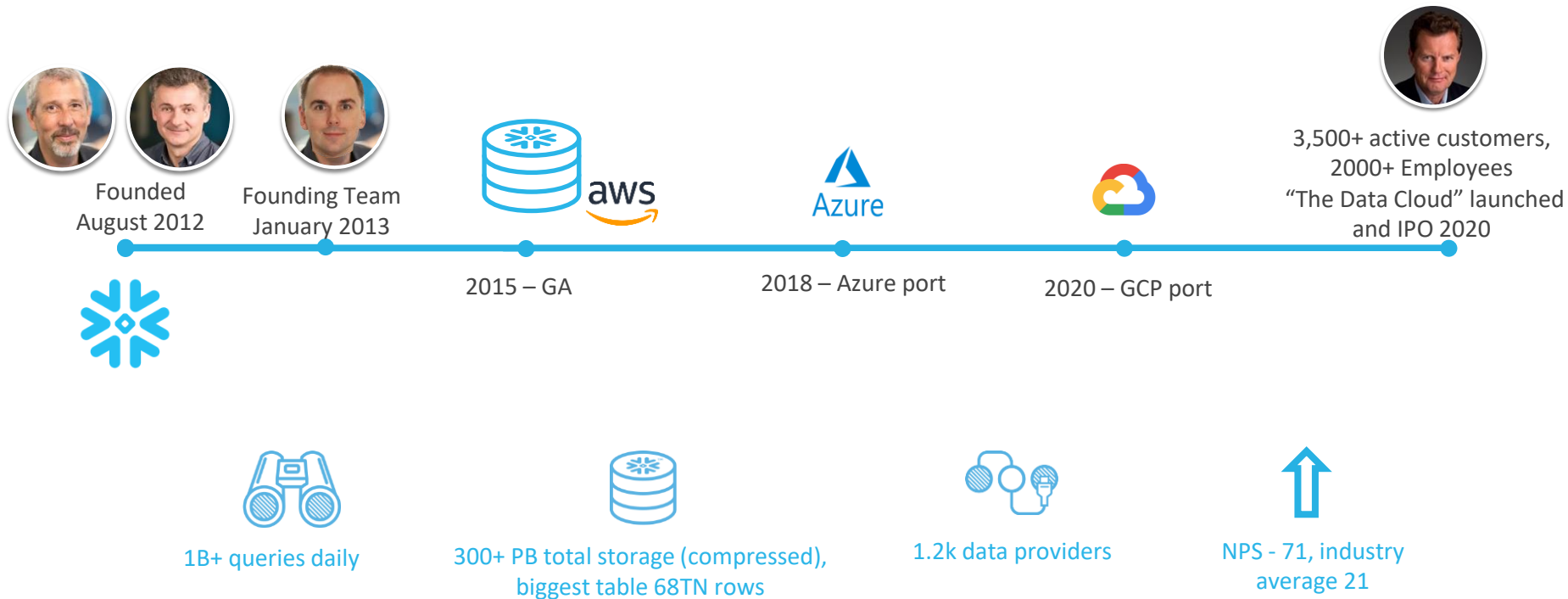
WHY SNOWFLAKE?

No Good Solution to Tackle Modern Data Challenges

More Data

Human Generated
Structured

Machine Generated
Semi-Structured

More Users

Few Analysts

Everyone

Faster Answers

Daily or Less Frequent

Real-Time
Interactive

Diverse Data Sources

Few, mostly internal

Many, both internal and external
(suppliers, customers, competitors, public, …)

SNOWFLAKE DATA CLOUD

Data Analytics Platform Built for the Cloud Era

UNLIMITED SCALE

- All Data
- All Users
- Instant Elasticity
- One System
  - Multi-region
  - Multi-cloud
- No Compromise

COLLABORATIVE

- Data Sharing
- Data Services
- Existing Content

SIMPLICITY

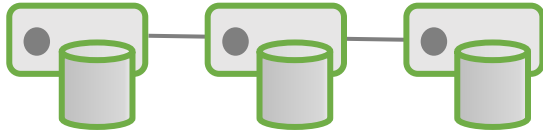- Self-Managed Service
- No Tuning Knobs
- Democratize Data Analytics
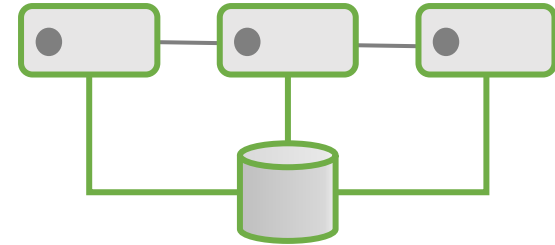
UNLIMITED SCALE

TRADITIONAL DATABASE ARCHITECTURES

Limited Scalability, Not Elastic

Shared-nothing

- Distributed Storage
- Single Cluster
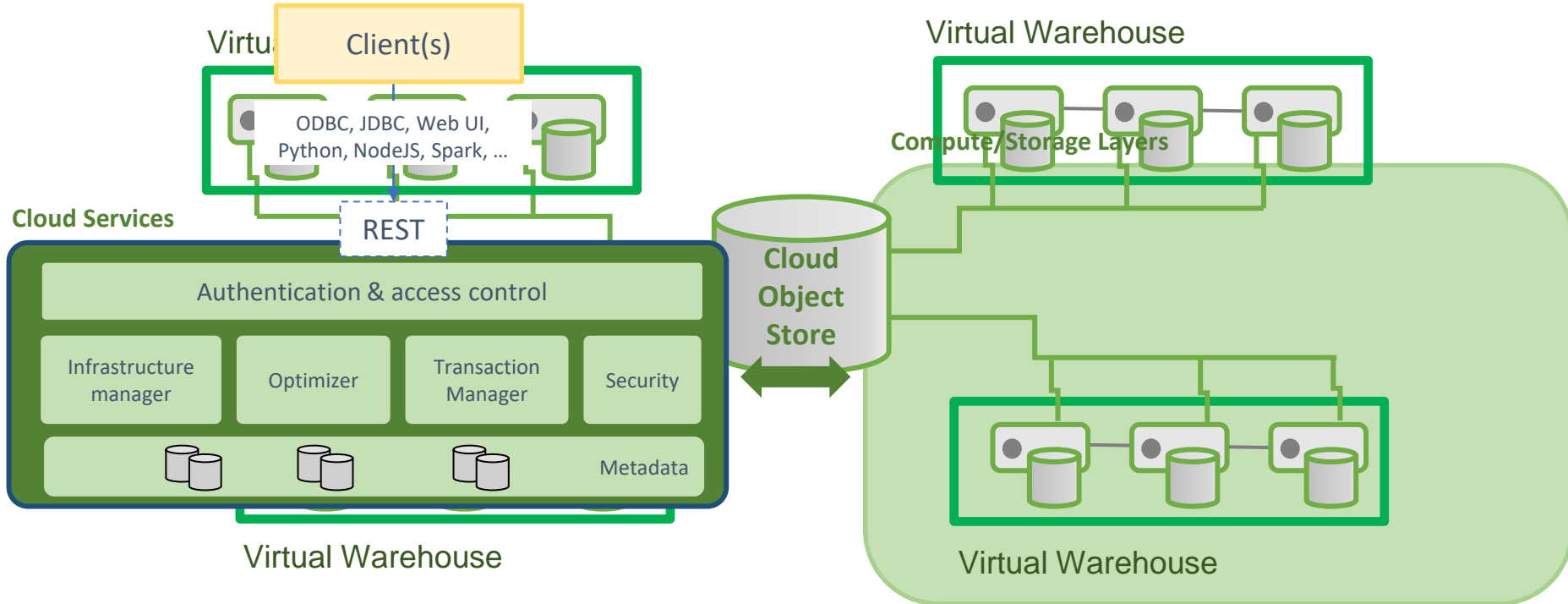- Adopted by Gamma, Teradata, Redshift, Vertica, Netezza, …

Shared-disk

- Centralized Storage
- Single Cluster
- Adopted by Oracle, Hadoop

SNOWFLAKE REGION ARCHITECTURE
## Multi-cluster, Shared-data

STORAGE TIER

- **Immutable Storage**
  - Each table is automatically partitioned horizontally
  - Partition size is kept very small, generally 16MB
  - Each partition is backed by an immutable file
  - Partitions are columnar organized, compressed, encrypted
  - Partitions are the unit of change for transactions

- **Semi-structured**
  - Variant data type used to store schemaless semi-structured data
  - Automatic columnarization of semi-structured attributes

- **Partition Metadata**
  - Out-of-box, metadata is automatically stored for all columns/sub-columns in a partition
  - Leverage that metadata to perform partition pruning
  - Re-clustering service to improve pruning
  - Track all table mutations to provide full ACID support

Cloud Object Store

COMPUTE TIER
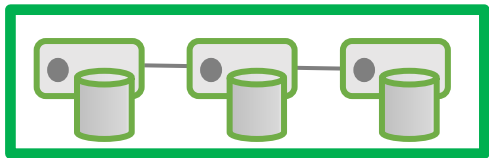
- **Virtual warehouse**
  - Snowflake Entity used to manage the set of compute resources used by a workload
  - Made of one or more compute clusters
  - Cluster size range from one to several hundred nodes
 - Workloads are fully isolated from each other

- **Just-in-time Compute**
  - Sub-second auto-resume when associated workload starts
  - Online resize up and down possible while workload runs
  - Auto-suspend when workload is no longer running
  - Snowflake charges usage by second of compute resource used
    ➔ **FAST is free!**

- **Partition Cache**
  - Node local memory and SSD storage used to cache partitions
  - Only columns/sub-columns which are accessed are cached
  - Highly available, fully stateless

CLOUD SERVICES
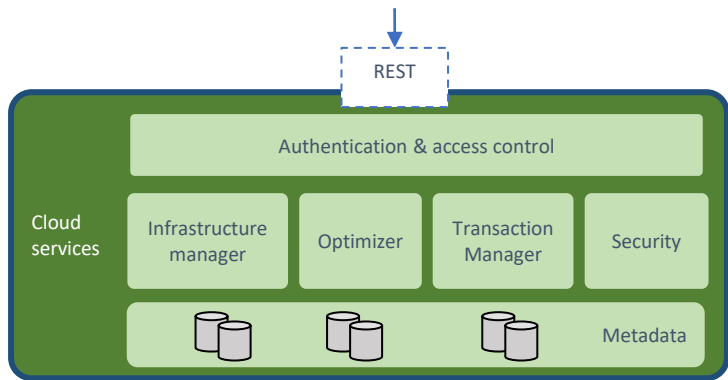
- **Control Plane of a Snowflake Region**
  - Connection Management
  - Infrastructure Provisioning and Management
  - Metadata storage (use FDB) & management
  - Query planning and optimization
  - Transaction management
 - Security management

- **Self-managed**
  - Self-upgrade of both software and hardware
  - Self-healing: replacement of failed servers and transparent re-execution of any failed queries
  - Highly available over multiple availability zone
  - Stateless: persistent sessions for load-balancing and transparent fail-over

SNOWFLAKE DATA CLOUD

One Integrated Platform Supporting Multiple Workload Types

DATA WAREHOUSE

Complete SQL
ACID
Low-latency
High-concurrency
UDFs, UDTs
Data Governance
Stored Procedures

DATA ENGINEERING
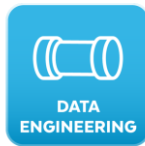
Streaming Ingest
Tasks
Table Streams
External Functions
Data Pipelines

DATA LAKE
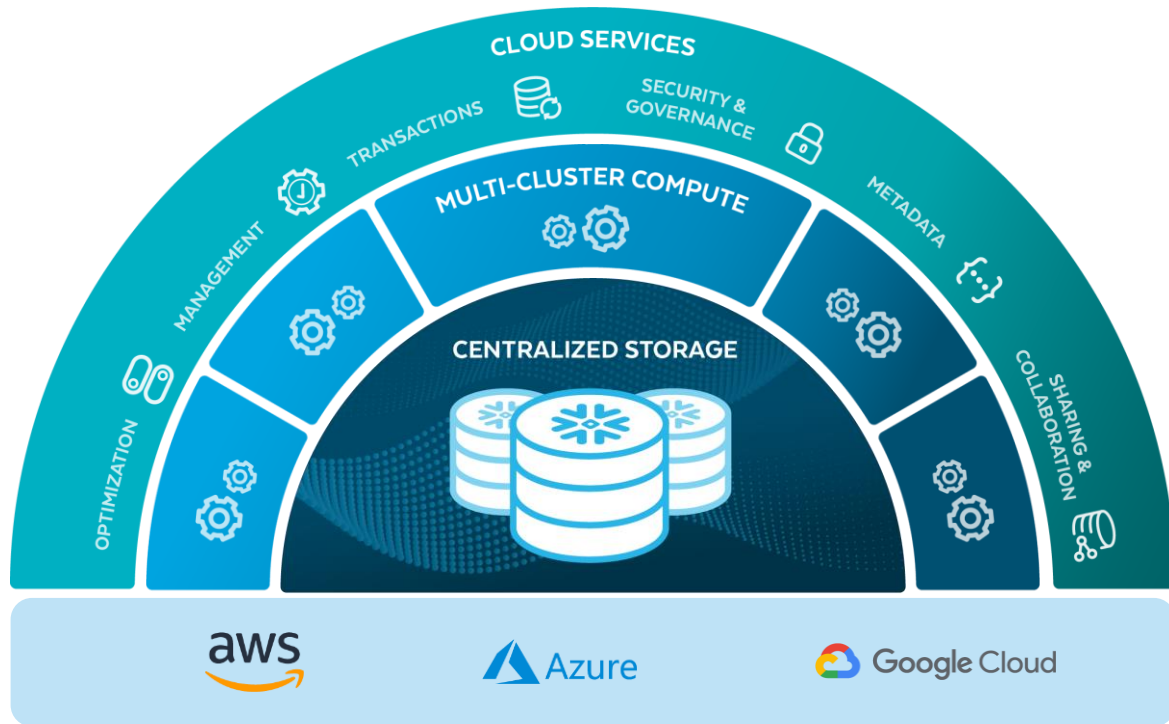
Semi-structured Data
Unstructured Data
External Tables

DATA SCIENCE

Java/Scala/Python
Data Frames

DATA APPLICATIONS

Rest APIs
Real-time

SNOWFLAKE DATA CLOUD REGION

SNOWFLAKE DATA CLOUD (2015)

Single Data Cloud Region (AWS)

First Snowflake Region
AWS-US-WEST

Snowflake Region (AWS)
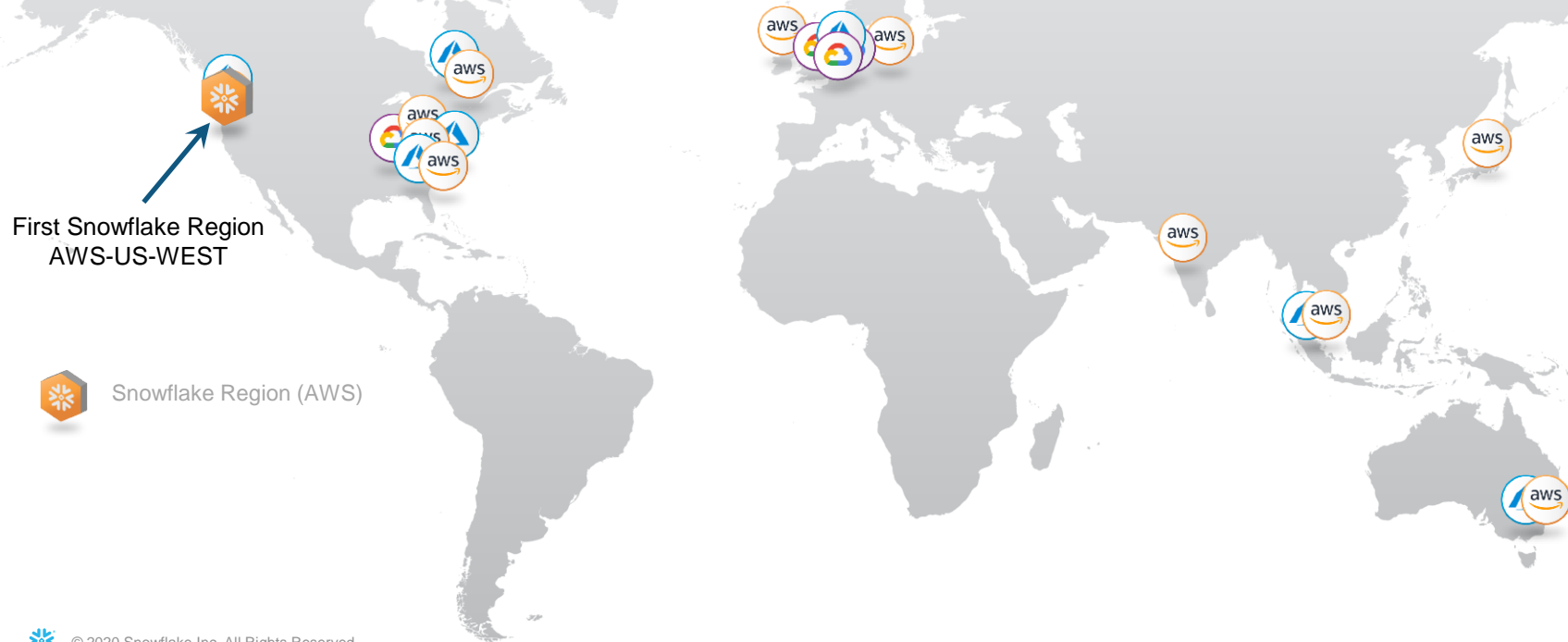
SNOWFLAKE DATA CLOUD (2021)

22 Data Cloud Regions (10 countries, 3 clouds)

GLOBAL
DATA MESH

Snowflake Region (AWS)

Snowflake Region (Azure)

Snowflake Region (GCP)

14
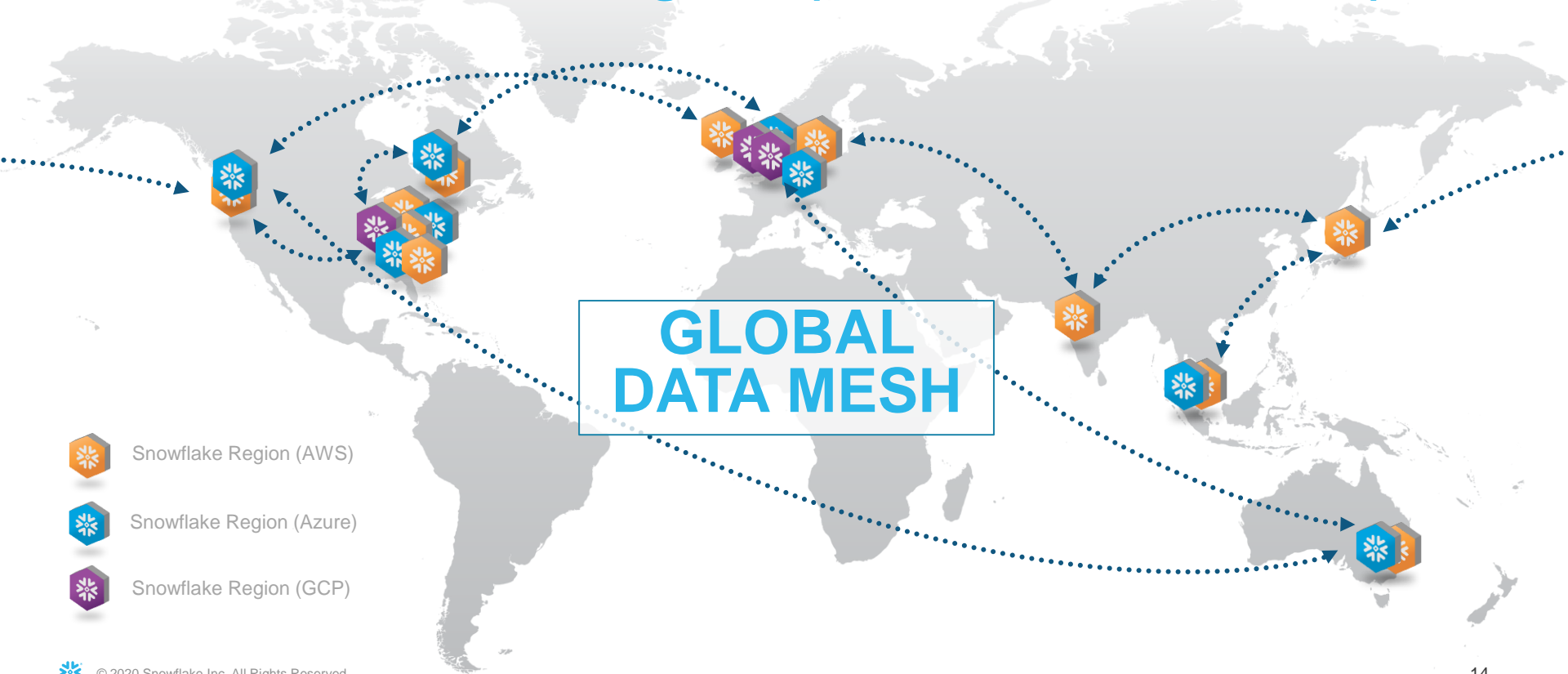
BUILDING A LARGE-SCALE GLOBAL SERVICE
Lessons Learned

Way harder than anticipated…

- Customers expect at least 3+ 9's of availability, 24x7
- At large scale, anything will happen. Hence we need to proactively anticipate and defend
- Everything needs to be fully automated and fully adaptive
- As much as possible self-managed versus dev-ops automation
- Keeping up with exponential growth ➜ scale cloud services and removing bottlenecks
- Weekly release without introducing (visible) regressions

… but so much faster development cycles

- We have built a top-notch and feature rich platform in only few years!
- Weekly release worldwide with single version to maintain
- Virtuous cycle – data driven development to identify and prioritize feature development
  - For example, use focus on improving DMLs and transaction processing since dominates
- Snowflake platform is extensively instrumented ➜ we generate many terabytes of service data daily
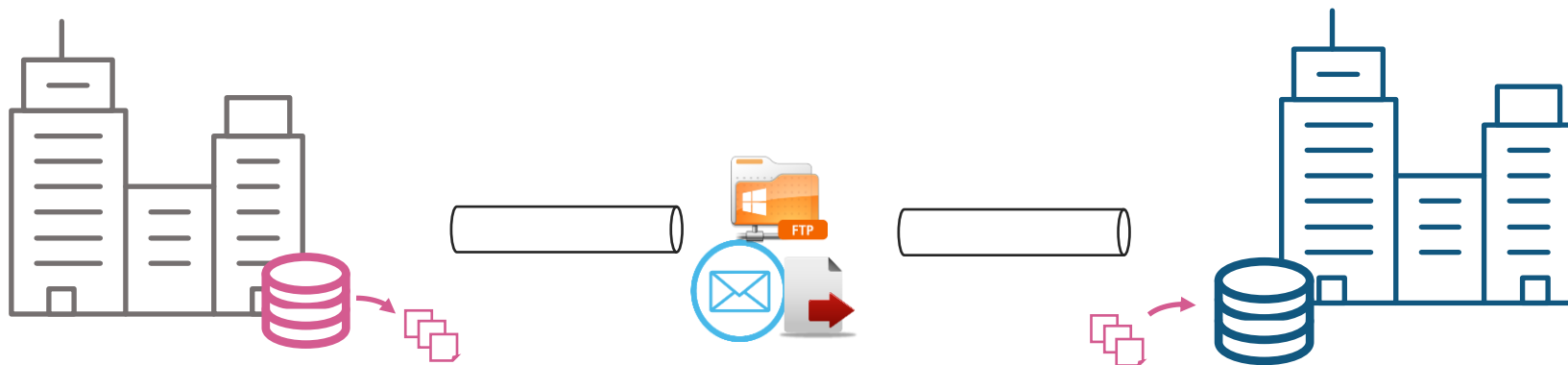
COLLABORATIVE

DATA COLLABORATION
Traditional Way

Data providers
1. Export data to files
2. Publish schema
3. Stage files for transport

- Redundant
- Inflexible
- Inefficient
- Insecure
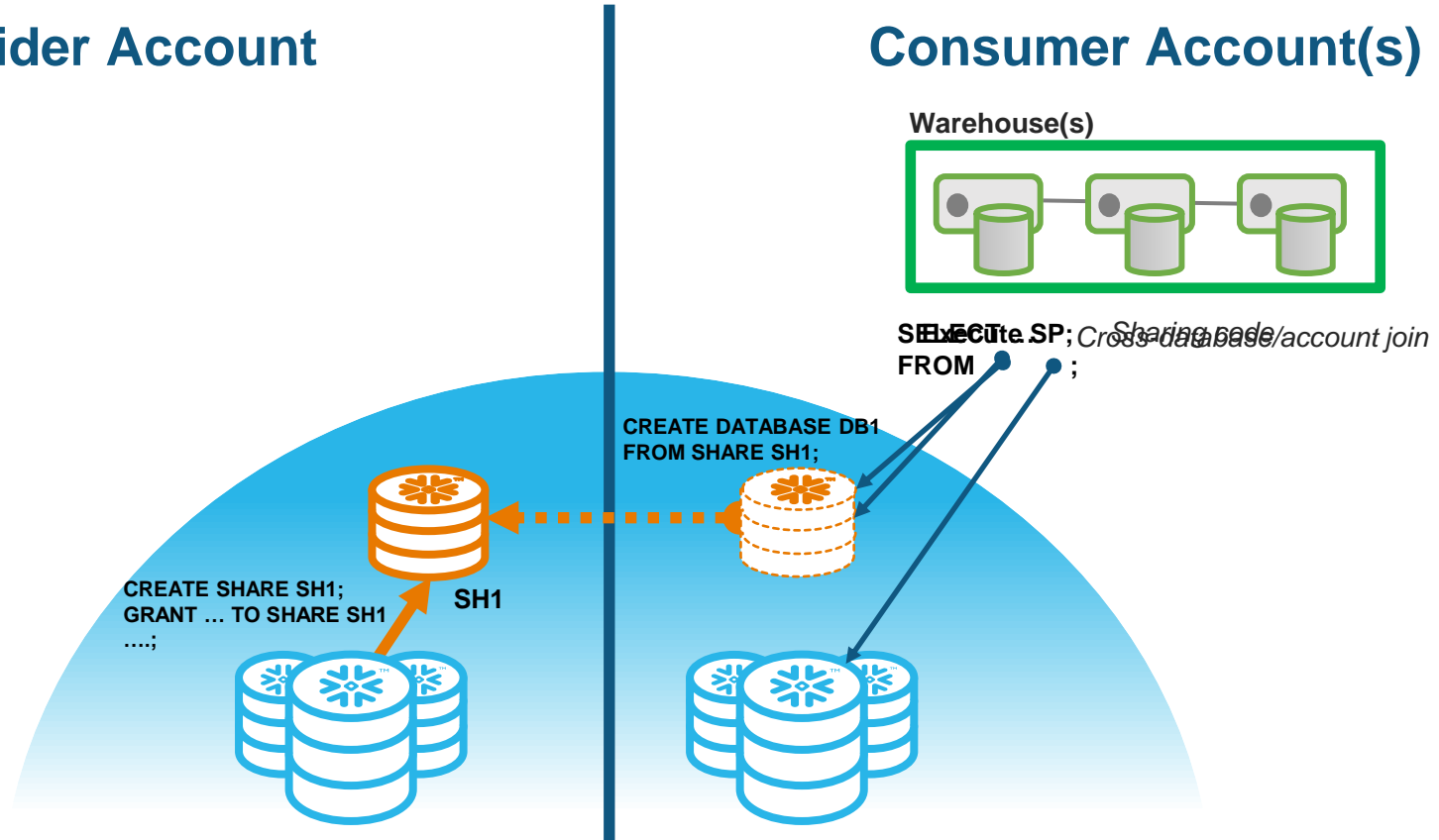- Expensive

Data customers
1. Additional infrastructure
2. Forced to recreate data structure
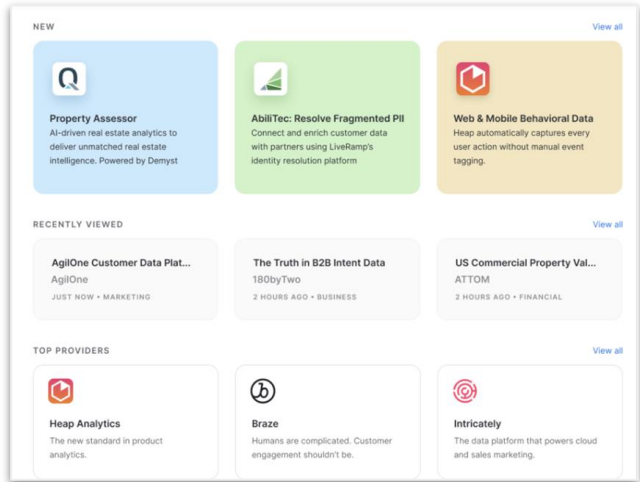3. Delayed updates to data

SNOWFLAKE DATABASE SHARING

Provider Account

Consumer Account(s)

Warehouse(s)

SELECT … **Execute SP;** *Sharing node*
FROM … ; *Cross-database/account join*

CREATE DATABASE DB1
FROM SHARE SH1;

CREATE SHARE SH1;
GRANT … TO SHARE SH1
….;

SH1

SNOWFLAKE DATA MARKETPLACE

READY TO USE DATABASES FROM MULTIPLE PROVIDERS

✔ Live, ready-to-query data; no copying or moving

✔ Only data marketplace with personalized data

✔ Globally available, across clouds

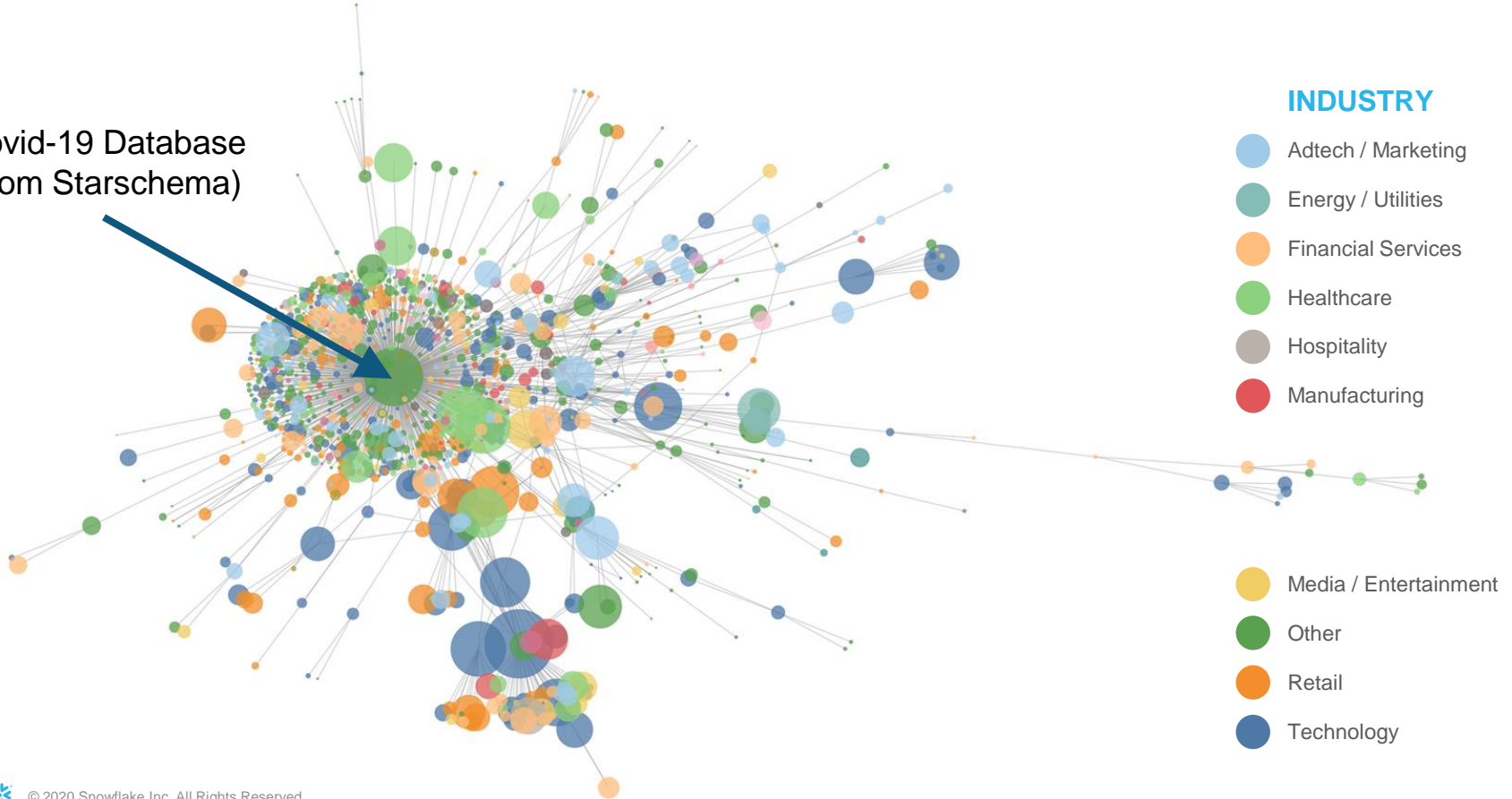Financial   Marketing   Demographic   Macroeconomic   Government   Healthcare   Business

CONNECTED THROUGH DATA

Covid-19 Database
(from Starschema)

INDUSTRY

- Adtech / Marketing
- Energy / Utilities
- Financial Services
- Healthcare
- Hospitality
- Manufacturing

- Media / Entertainment
- Other
- Retail
- Technology

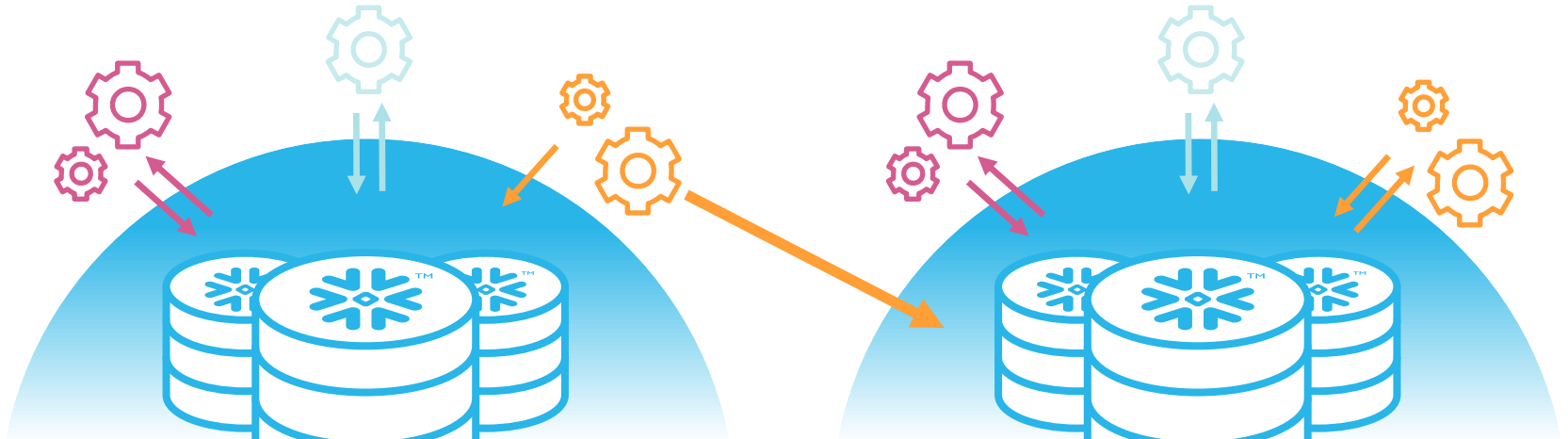# SNOWFLAKE DATABASE SHARING
Conclusion

Secure   Live   Frictionless   Personalized   Global

SIMPLICITY

WHY SIMPLICITY MATTERS

Manage Data, Not Infrastructure!

Infrastructure	**Physical Design**	**Data Collaboration**	**Query Tuning**	**Availability**
Initial Setup	Partitioning	Loading	Statistic Collection	Setup High availability
Upgrading	Indexing	Moving	Memory Management	Handle Hardware Faults
Patching	Ordering	Transforming	Parallelism	Manage Backups
Capacity Planning	Vacuuming	Copying	Query Plan Hinting	
Storage		Securing	Workload Management	
Security				

SNOWFLAKE CLOUD DATA PLATFORM

Minimal Administration

Infrastructure	Physical Design	Data Collaboration	Query Tuning	Availability
Initial Setup	Partitioning	Loading	Statistic Collection	Replication
Upgrading	Indexing	Providing	Memory Management	Backups
Patching	Ordering	Transforming	Parallelism	Re-Clustering
Capacity Planning	Vacuuming	Copying	Query Plan Hinting	Account Management
Storage		Securing	Workload Management	
Security				

Simply load/share data and run queries

Infrastructure

|		
Initial Setup	→	NA - Service is always on
Upgrading Patching	→	Automatic – Performed weekly by Cloud Services
Capacity Planning	→	NA - Just-in-time compute
Storage	→	NA - unlimited storage, spill to blob storage
Security	→	Automatic – Encryption, Monitoring, ...

Physical Design

Partitioning	→	NA – automatic at load time
Ordering	→	Default/Automatic Clustering
Indexing	→	Search optimization service
Vacuuming	→	NA – Immutable partitions

Data Collaboration

Loading

Moving

Transforming

Copying

Securing

Live and Secure Data Sharing
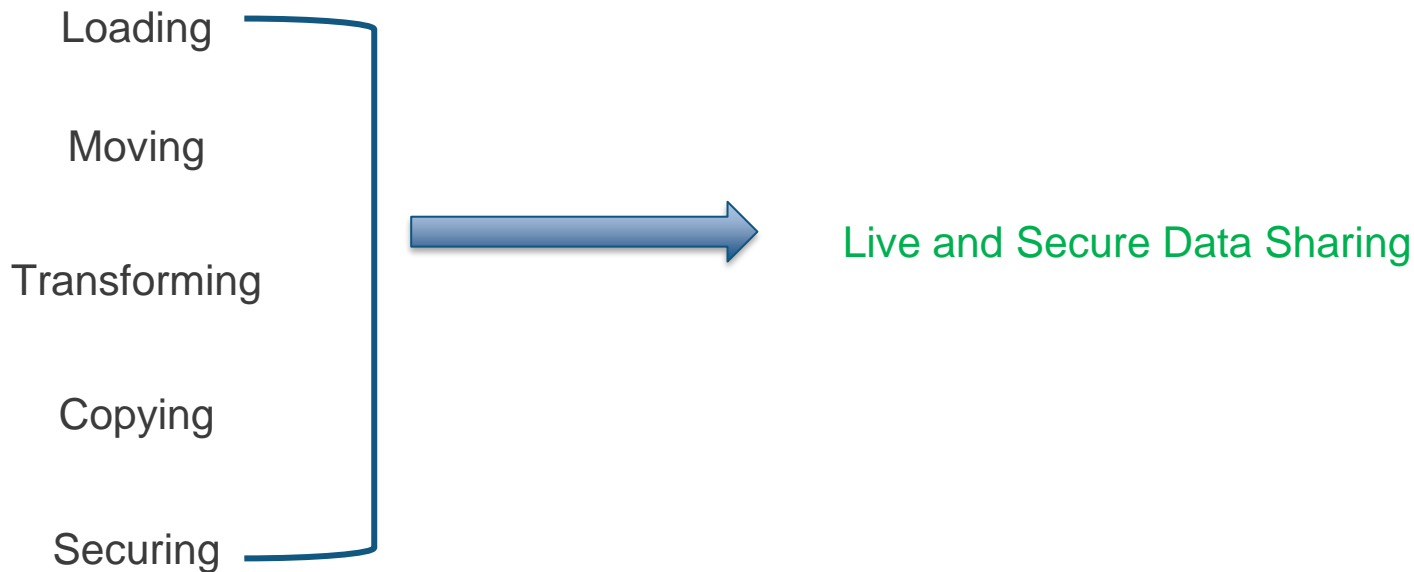
Query Tuning

Statistic Collection	→	Automatic – at DML time
Memory Management	→	Automatic – Cooperative Memory Brokering
Parallelism	→	Automatic – Adaptive
Query Plan Hinting	→	Robust adaptive execution strategy dynamic join filters, adaptive push down and distribution methods, join skew resilience
Workload Management	→	Virtual warehouse per workload Auto-scale multi-cluster warehouse

Availability

Setup High Availability → Out-of-box: Snowflake Architecture Multi-AZ
Disaster Recovery: Cross-region Replication

Handle Hardware Faults → Automatic: Snowflake Cloud Services detects and replace faulty hardware

Backups → Automatic: blob storage with 11 9's durability, Undrop, Clone as-of, time travel, Fail-Safe

CONCLUSION

SNOWFLAKE DATA CLOUD

Worldwide Web of Data

Single System

Collaborative

Simplicity