THE DREAM DATA WAREHOUSE
(CIRCA 2012)

Unlimited and Instant Scaling
- No data silos
- 10x faster for the same price, no over provisioning

Store all your data
- Structured and semi-structured
- Petabyte scale at very low cost

Extreme simplicity
- No management tasks, offered as a service
- Fast out-of-box with no tuning knobs

No compromises full fledge Data Warehouse
- Full support for ACID transactions with read consistency
- ANSI SQL, RBAC
WHY THEN?
OUR VIEW OF THE CLOUD...

- Storage became dirt cheap
- Flat network offered uniform bandwidth
- Single core performance stalled
- Data warehouse and analytic workload are mostly CPU bound

Design for abundance and not scarcity of resources
THREE PILLARS

Multi-cluster shared data Architecture
- Leverage cloud elasticity and pay only what you use
- Instant scale
- Performance isolation
- Real-time Data sharing

Immutable Scalable Storage
- Extremely fast response time at scale
- Fine grain vertical and horizontal pruning on any column
- Automatically applied to any data (structured and semi-structured)

Multi-Tenant Service
- Self-tuning, self-healing
- Transparent upgrade
- Service architecture designed for availability, durability and security
ARCHITECTURE
AN ARCHITECTURE BUILT FOR THE CLOUD

Traditional Architectures

- **Shared-disk**
  - Shared storage
  - Single cluster

- **Shared-nothing**
  - Decentralized, local storage
  - Single cluster

- **Multi-cluster, shared data**
  - Centralized, scale-out storage
  - Multiple, independent compute clusters
No data silos
Storage decoupled from compute

Any data
Native for structured & semi-structured

Unlimited scalability
Along many dimensions

Low cost
Compute on demand

Instantly cloning
Isolate prod from dev & qa

Highly available
11 9’s durability, 4 9’s availability
VIRTUAL WAREHOUSE

How to allow concurrent workloads run without impacting each other?

- One or more MPP compute cluster
- Unit of fault and performance isolation
- Use multiple warehouses to segregate workload
- Resizable on the fly
- Able to access data in any database
- Transparently caches data accessed
- Transaction manager synchronizes data access
- Automatic suspend when idle and resume when needed
MULTI-CLUSTER WAREHOUSE
LEVERAGE ABUNDANCE OF COMPUTE RESOURCES

- Automatically scales compute resources based on concurrent usage
- Single virtual warehouse of multiple compute clusters
- Queries are load balanced across the clusters in a virtual warehouse
- Split across availability zones for high availability

Query scheduler

Virtual Warehouse Group

Cluster 1

Cluster 2

Cluster 3
IN THE REAL-WORLD

**Continuous Loading (4TB/day)**
- S3 <5min SLA

**Virtual Warehouse**
- Medium
- Large
- 2X-Large
- Auto Scale – X-Large x 5

**Reporting (Segmented)**
- Interactive Dashboard
  - 50% < 1s
  - 85% < 2s
  - 95% < 5s

**ETL & Maintenance**
- Virtual Warehouse
- Prod DB

4 trillion rows
3+ petabyte raw
8x compression ratio
25M+ micro-partitions
SCALABLE IMMUTABLE STORAGE
STORAGE IMMUTABILITY

- **Accumulates immutable data over time**
  Well supported by all cloud vendor object stores

- **Allow separation of storage and compute resources**
  Enable workload scalability

- **Heavily optimized for read mostly workload**
  Natural fit for analytic systems

- **Transaction management becomes a metadata problem**
  Multi-version concurrency control and Snapshot isolation semantic

- **Transaction coordination separated from storage and compute**
  Allow for consistent access across compute resources
Data is automatically partitioned at load time
Storage decoupled from compute

Columnar organization in each micro-partition
Enable both horizontal vertical pruning

Micro partition – only few 10MBs
Fine grain pruning, no skew

Metadata structure tracks data distribution
Very fast pruning at optimization time

Applied to both structured and semi-structured data
Very fast response time for both
AUTOMATICALLY APPLIED TO SEMI-STRUCTURED DATA

Semi-structured data
(JSON, Avro, XML, Parquet, ORC)

Structured data
(e.g., CSV, TSV, …)

Native support
Loaded in raw form (e.g. JSON, Avro, XML)

Optimized storage
Optimized data type, no fixed schema or transformation required

Optimized SQL querying
Full benefit of database optimizations (pruning, filtering, …)
EXAMPLE

Client Application
- ODBC Driver
- JDBC Driver
- Web UI

Compute
- Custom Reports
- Campaign Analysts
- Loading WH

Cloud Services
- Optimization
- Query Mgmt
- Warehouse Mgmt
- Security

HTTPS (JDBC/ODBC/Python)

Storage

DDL

Campaign Analysis

Loading WH
**ENABLE DATA SHARING**

**Providers**
- Secure and integrated Snowflake’s access control model
- Only pay normal storage costs for shared data
- No limit to the number of consumer accounts with which a dataset may be shared

**Consumers**
- Get access to the data without any need to move or transform it.
- Query and combine shared data with existing data or join together data from multiple publishers
ENABLE GLOBAL REPLICATION
MULTI-TENANT SERVICE
## DATA WAREHOUSE AS A SERVICE

<table>
<thead>
<tr>
<th>Multi-Tenant Service</th>
<th>Availability</th>
<th>Durability</th>
</tr>
</thead>
<tbody>
<tr>
<td>No administration, self-tuning and healing,</td>
<td>All tier distributed over multiple datacenters with active-active data replication</td>
<td>Synchronous replication of data over multiple data centers</td>
</tr>
<tr>
<td>Transparent upgrade</td>
<td>No maintenance downtime, fully transparent software &amp; hardware upgrade</td>
<td>Automatic data retention and fail safe technology to guard against any data removal</td>
</tr>
<tr>
<td>Service architecture designed for high availability and durability</td>
<td>Automatic repair of any failed servers with transparent re-execution of any failed queries</td>
<td></td>
</tr>
<tr>
<td>Security is at the core</td>
<td>Persistent session for load-balancing and transparent fail-over</td>
<td></td>
</tr>
</tbody>
</table>
SNOWFLAKE SERVICE

Three independent layers

Authentication & Access Control
- Infrastructure manager
- Optimizer
- Transaction manager
- Security

Metadata

Cache

Cloud services
- Compilation and Management

Data processing
- Virtual warehouses

Storage
- Databases
MANAGED SERVICE
BUILT-IN DISASTER RECOVERY AND HIGH AVAILABILITY

Scale-out of all tiers
metadata, compute, storage

Resiliency across multiple availability zones
geographic separation separate power grids built for synchronous replication

Fully online updates & patches
zero downtime

Back pressure and throttling
all the way back to the client
ADAPTIVE ALL THE WAY TO THE CORE
SELF TUNING & SELF HEALING INTERNALS

- Adaptive
- Self-tuning
- Do no harm!
- Automatic
- Default

Automatic Memory Management
Automatic Distribution Method
Automatic Degree of Parallelism
Automatic Fault Handling
Automatic Workload Management
No Vacuuming
No Statistics
EXAMPLE: AUTOMATIC SKEW AVOIDANCE

1. Detect popular values on the build side of the join
2. Use broadcast for those and directed join for the others

- Adaptive \[\rightarrow\] popular values detected at runtime
- Self-tuning \[\rightarrow\] number of values
- Do no harm! \[\rightarrow\] no performance degradation
- Automatic \[\rightarrow\] kicks in when needed
- Default \[\rightarrow\] enabled by default for all joins
WHAT’S NEXT?
SERVERLESS DATA SERVICES

- Target predictable well-identified database workloads
- Horizontal scaling is automatic
- Fine grain unit of work allow for degree of parallelism to be arbitrarily small or large
- Secure since handled by the service
- Transparent retry on failures
- Service state entirely managed by the service
- Monitoring and observability of the service
CLOUD NATIVE ARCHITECTURE
A GIFT THAT KEEPS ON GIVING