SNOWFLAKE: CHANGING THE GAME WITH CLOUD DATA WAREHOUSING

KENT GRAZIANO, CHIEF TECHNICAL EVANGELIST I @KentGraziano

© 2019 Snowflake Inc. All Rights Reserved

My Bio

- Chief Technical Evangelist, Snowflake Computing
- Oracle ACE Director, Alumni (DW/BI)
- OakTable Network
- Blogger – The Data Warrior
- Certified Data Vault Master and DV 2.0 Practitioner
- Former Member: Boulder BI Brain Trust (#BBBT)
- Member: DAMA Houston & DAMA International
- Data Architecture and Data Warehouse Specialist
  - 30+ years in IT
  - 25+ years of Oracle-related work
  - 20+ years of data warehousing experience
- Author & Co-Author of a bunch of books (Amazon)
- Past-President of ODTUG and Rocky Mountain Oracle User Group

© 2019 Snowflake Inc. All Rights Reserved

AGENDA

Data challenges today

What is a Data Warehouse as a Service?

Introducing Snowflake

Top 10 (or so) cool features of Snowflake

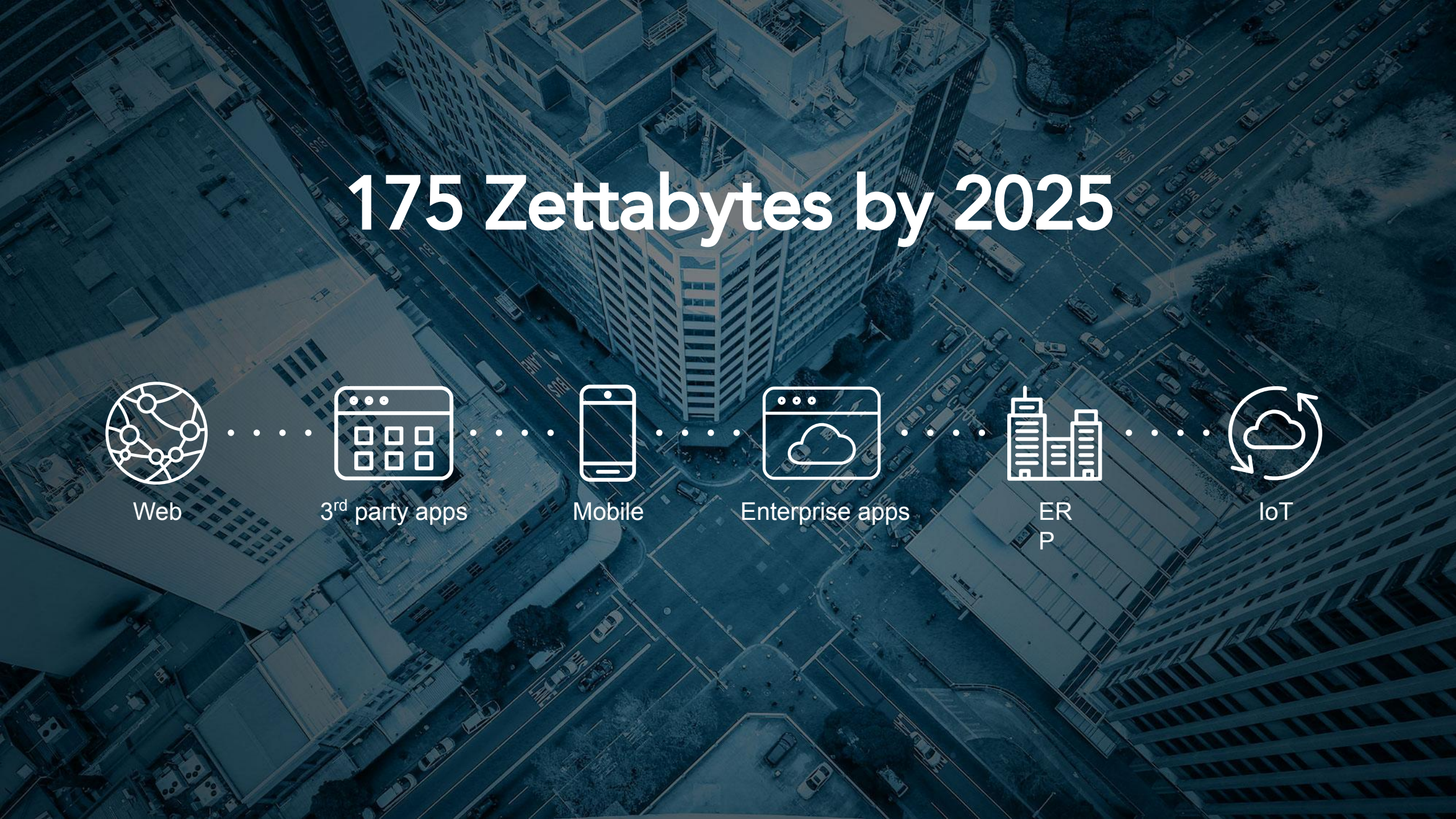
Snowflake in Action

© 2019 Snowflake Inc. All Rights Reserved

Data Challenges Today

© 2019 Snowflake Inc. All Rights Reserved

175 Zettabytes by 2025



It’s not the data itself

it’s how you take full advantage of the insight it provides

Most firms don’t consistently turn data into action

73%

of firms aspire to be ***data-driven.***

All Possible data

All Possible Action

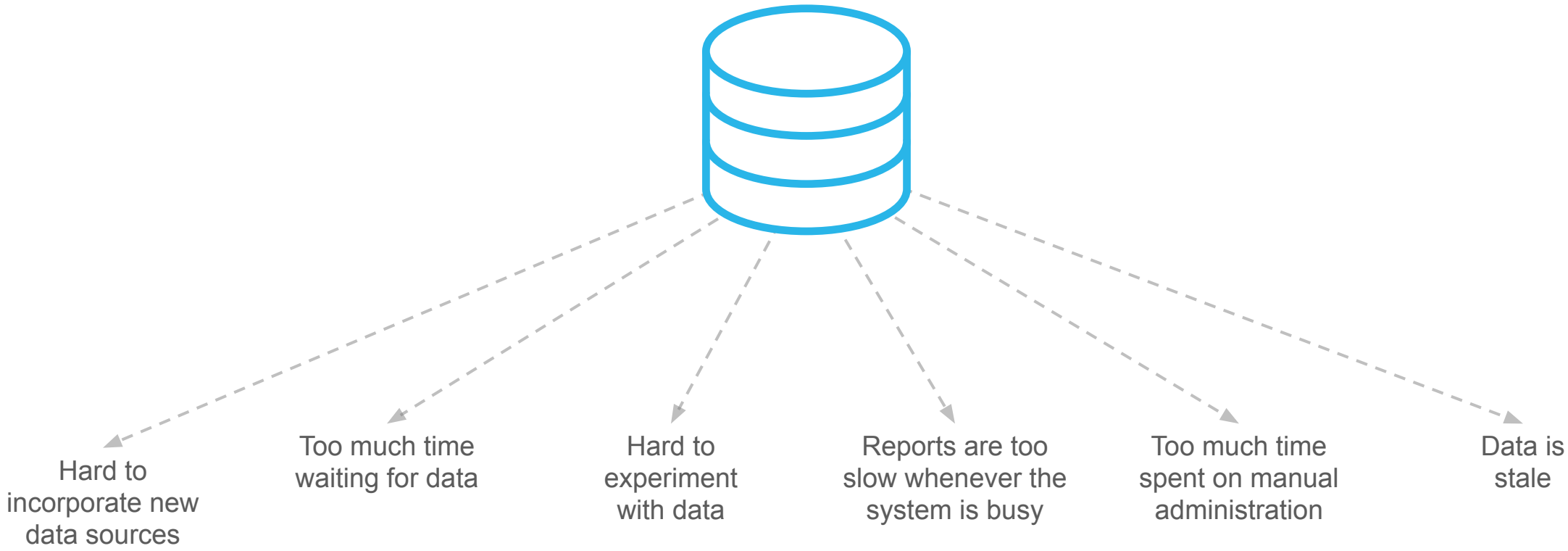
29%

of firms are good at turning data into **action.**

Source: Forrester

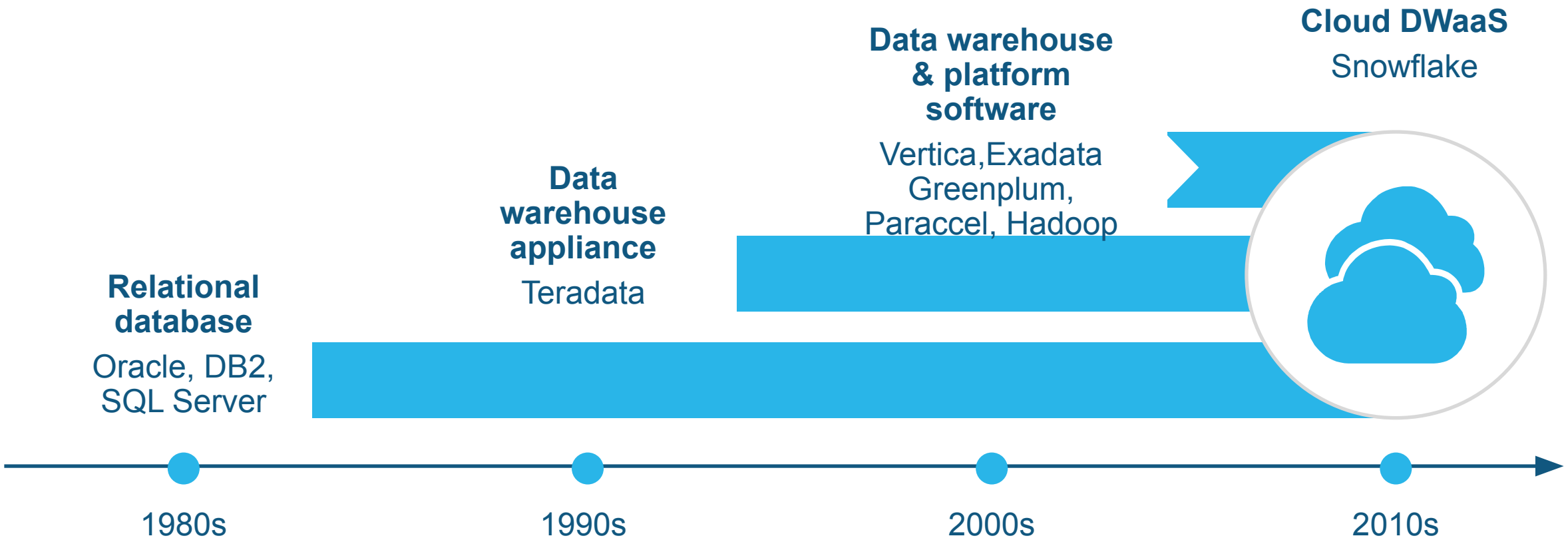
© 2019 Snowflake Inc. All Rights Reserved

THE DATA STRUGGLE



© 2019 Snowflake Inc. All Rights Reserved

The evolution of data platforms



© 2019 Snowflake Inc. All Rights Reserved

What is a Cloud DWaaS?

DW- Data Warehouse

- Relational database
- Uses standard SQL
- Optimized for fast loads and analytic queries

---

aaS – As a Service

- Like SaaS (e.g. SalesForce.com)
- No infrastructure set up
- Minimal to no administration
- Managed for you by the vendor
- Pay as you go, for what you use

© 2019 Snowflake Inc. All Rights Reserved

Introducing Snowflake

© 2019 Snowflake Inc. All Rights Reserved

3 years in stealth + 4 years GA

Founded 2012 by industry veterans with over 120 database patents

Over $920M in venture funding from leading investors

First customers 2014, general availability 2015

1300+ employees
Over 2000 customers today

Fun facts:

Queries processed in Snowflake per day:	Largest single table:	Largest number of tables single DB:	Single customer most data:	Single customer most users:
260 million	**39 trillion rows**	**1.1 million!**	**> 15PB**	**> 11,000**

© 2019 Snowflake Inc. All Rights Reserved

Snowflake: a team of data experts

Benoit Dageville

CTO

Lead architect of Oracle parallel execution and a key manageability architect

Marcin Zukowski

Founder & VP of Engineering

Inventor of vectorized query execution in

Thierry Cruanes

Founder Architect

Leading expert in query optimization and parallel execution

Team

ORACLE · salesforce.com · aster data · Actian · vectorwise · Microsoft · Google · cloudera · LinkedIn · TimesTen

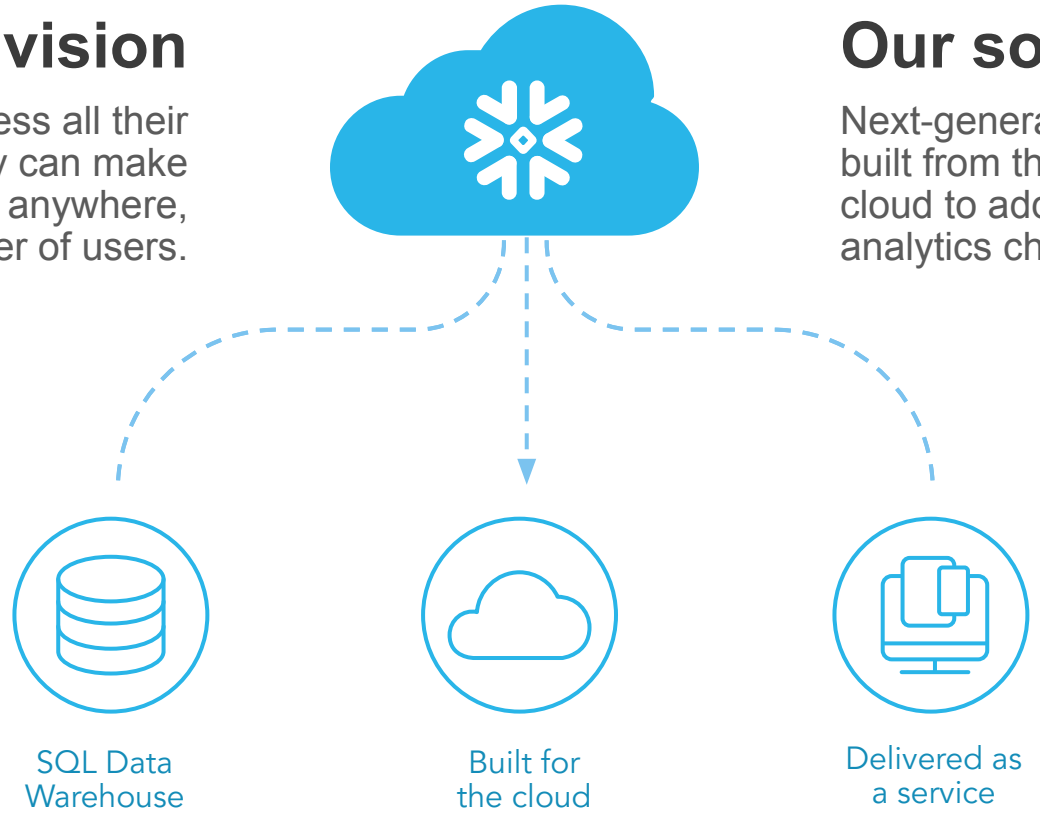
Investors

Redpoint · SUTTER HILL VENTURES · WiNG

SNOWFLAKE: A FULL DATA WAREHOUSE, BUILT FOR THE CLOUD

Our vision

Allow our customers to access all their data in one place so they can make actionable decisions anytime, anywhere, with any number of users.



Our solution

Next-generation data warehouse built from the ground up for the cloud to address today's data and analytics challenges.

© 2019 Snowflake Inc. All Rights Reserved

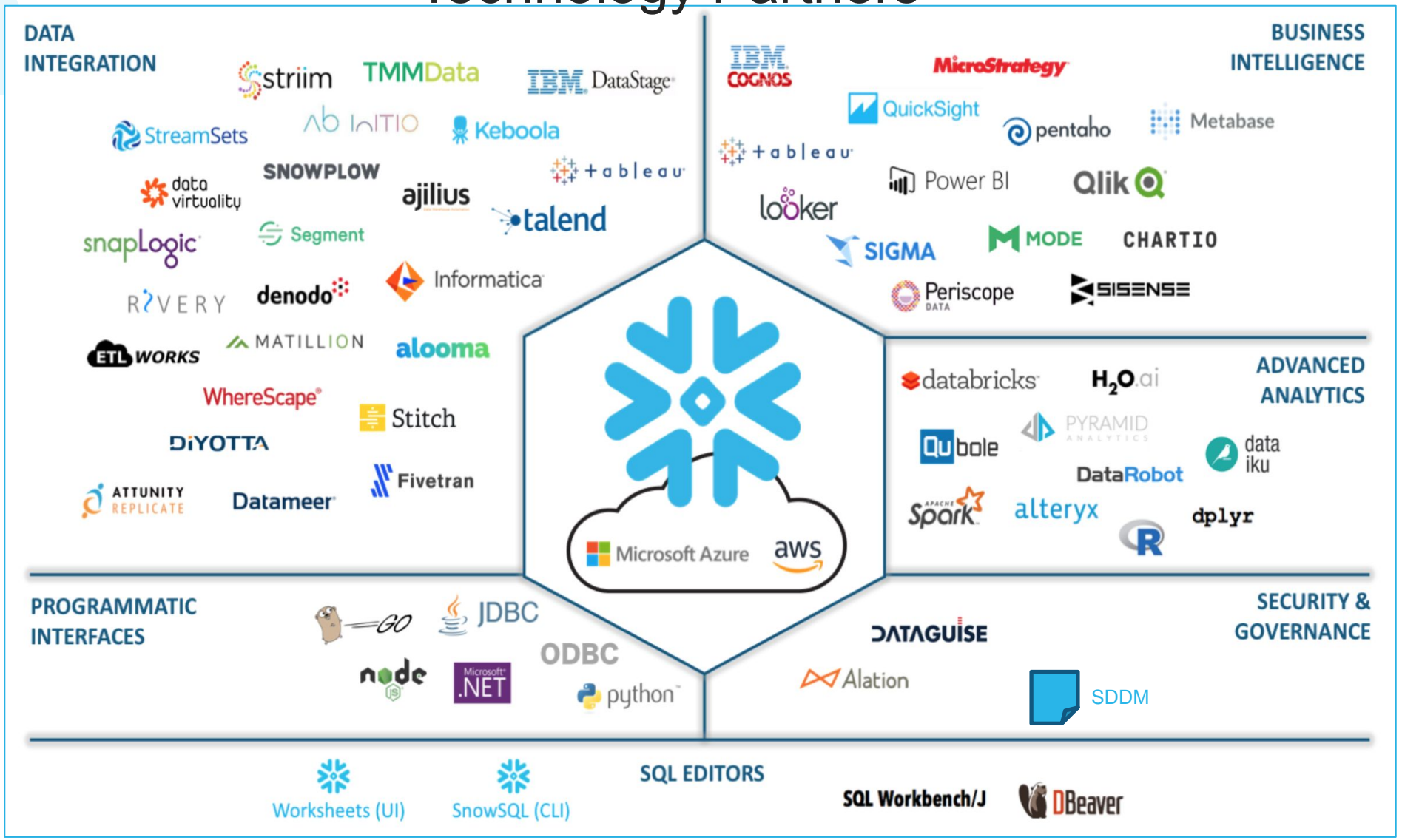
The Data Warrior's Top 10+ Cool Things About Snowflake

(A Data Geeks Guide to DWaaS)

© 2019 Snowflake Inc. All Rights Reserved

#10 Works with the tools and skills you already have.

Technology Partners

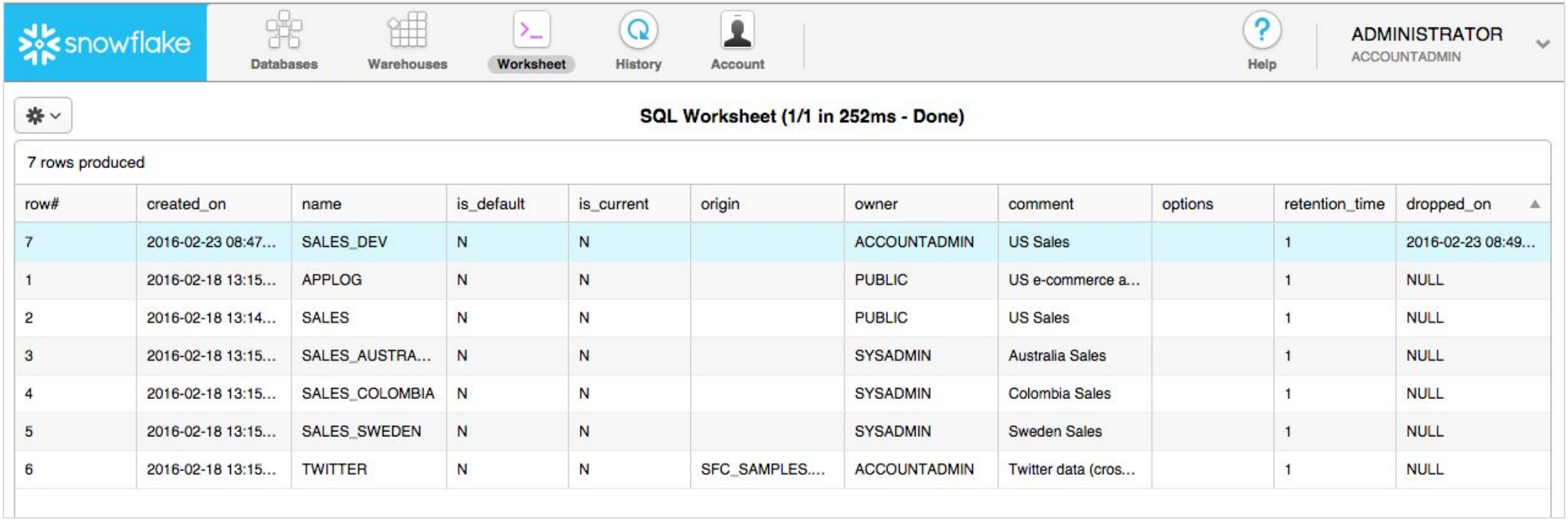


© 2019 Snowflake Inc. All Rights Reserved

#9 - UNDROP

UNDROP TABLE <table name>
UNDROP SCHEMA <schema name>
UNDROP DATABASE <db name>

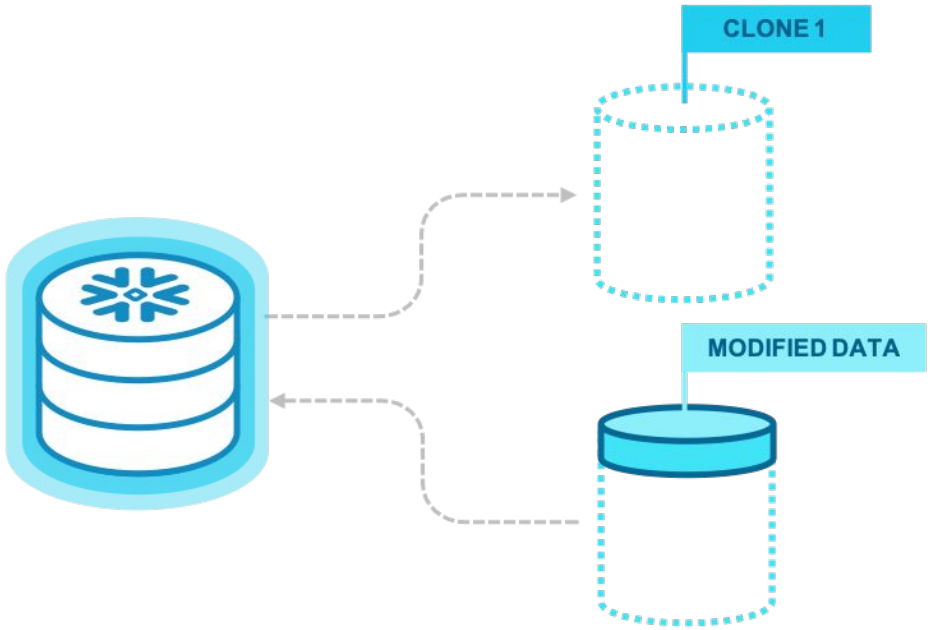
Part of Time Travel feature: **AWESOME!**



row#	created_on	name	is_default	is_current	origin	owner	comment	options	retention_time	dropped_on
7	2016-02-23 08:47...	SALES_DEV	N	N		ACCOUNTADMIN	US Sales		1	2016-02-23 08:49...
1	2016-02-18 13:15...	APPLOG	N	N		PUBLIC	US e-commerce a...		1	NULL
2	2016-02-18 13:14...	SALES	N	N		PUBLIC	US Sales		1	NULL
3	2016-02-18 13:15...	SALES_AUSTRA...	N	N		SYSADMIN	Australia Sales		1	NULL
4	2016-02-18 13:15...	SALES_COLOMBIA	N	N		SYSADMIN	Colombia Sales		1	NULL
5	2016-02-18 13:15...	SALES_SWEDEN	N	N		SYSADMIN	Sweden Sales		1	NULL
6	2016-02-18 13:15...	TWITTER	N	N	SFC_SAMPLES....	ACCOUNTADMIN	Twitter data (cros...		1	NULL

© 2019 Snowflake Inc. All Rights Reserved

#8 - ZERO-COPY DATA CLONING



Instant data cloning operations

Databases, schema, tables, etc

Metadata-only operation

Modified data stored as new blocks

Unmodified data stored only once

No data copying required, no cost!

Instant test/dev environments

Test code on your entire production dataset

Swap tables into production when ready

© 2019 Snowflake Inc. All Rights Reserved

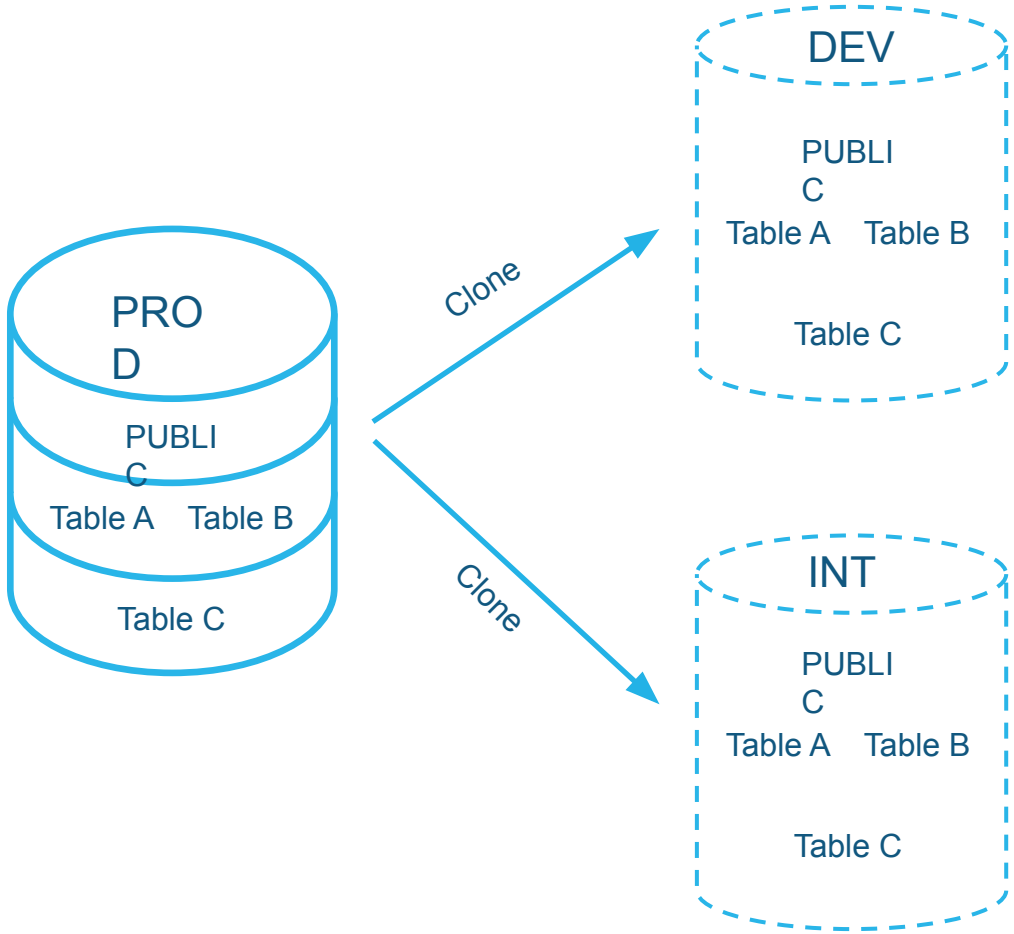
Simply SQL to Clone – Fast!

Instant copy of table, schema, or database:

```
CREATE OR REPLACE TABLE MyTable_V2
        CLONE MyTable;
```

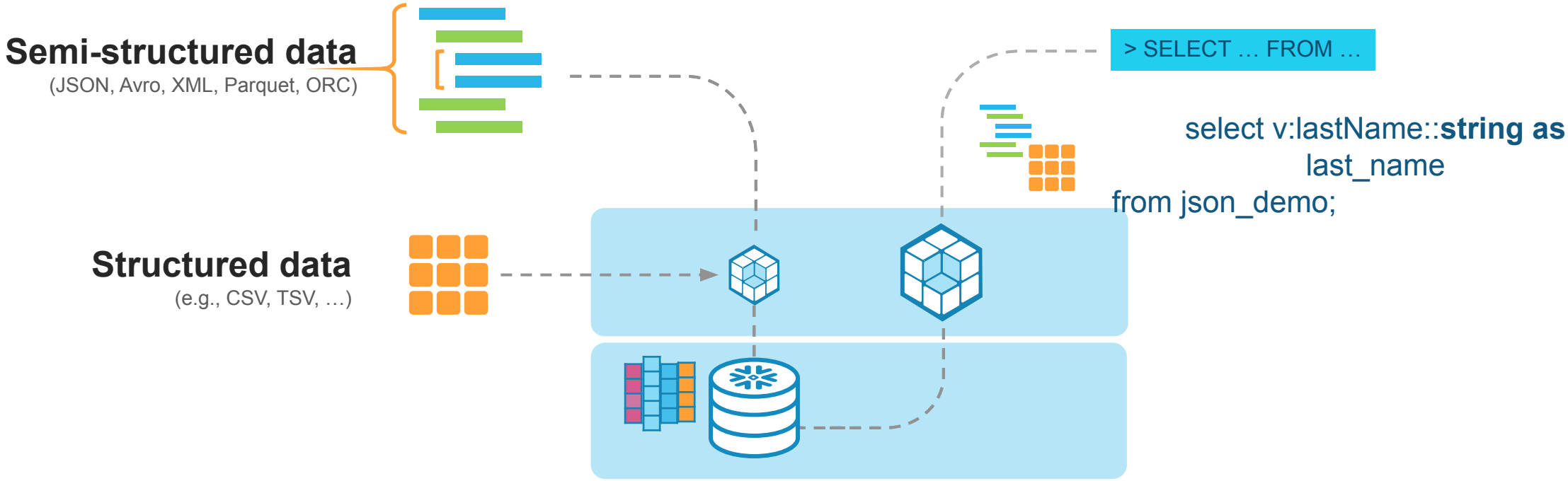
With Time Travel:

```
CREATE SCHEMA mytestschema_clone_restore
CLONE testschema
    BEFORE (TIMESTAMP =>
        TO_TIMESTAMP(40*365*86400));
```



© 2019 Snowflake Inc. All Rights Reserved

#7 – JSON Support with SQL!



Storage optimization

Transparent discovery and storage optimization of repeated elements

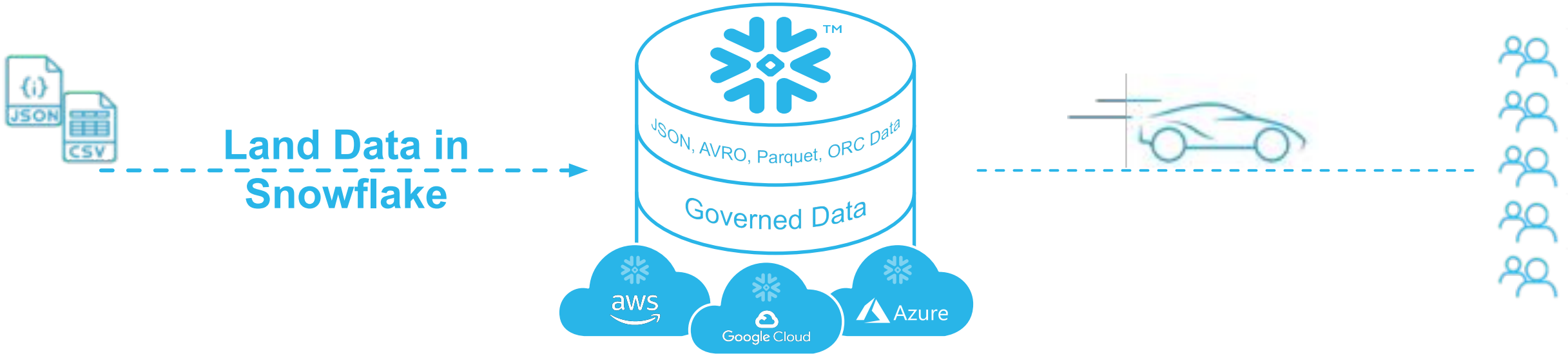
+

Query optimization

Full database optimization for queries on semi-structured data

© 2019 Snowflake Inc. All Rights Reserved

#6 - Snowflake as a Data Lake!



Fast Analytics

Built-in performance optimizations and data pruning

Simple Service

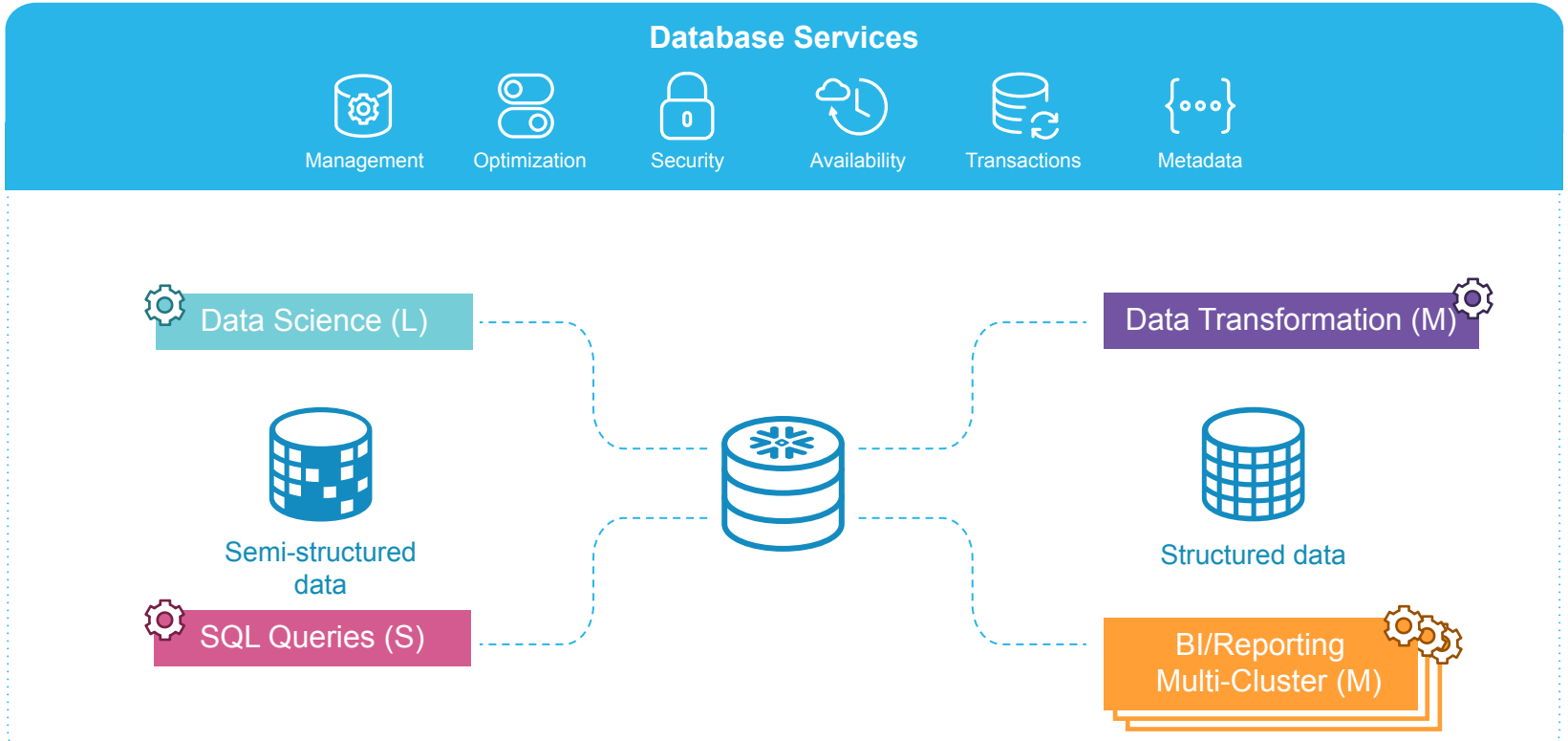
Single, integrated service, across clouds
Responsive customer support

Governed by Design

Role based access
Resource isolation and control
Automatic metadata and history

© 2019 Snowflake Inc. All Rights Reserved

#5 – Separation of Storage & Compute
New multi-cluster, shared data architecture



Flexible Cloud Storage For All Kinds of Data

Unlimited Compute Clusters to Serve Every Use

Easy-to-use Service with No Management

© 2019 Snowflake Inc. All Rights Reserved

#4 – Support Multiple Workloads

Deliver faster analytics
at any scale

Accelerate the data pipeline

Run loading & analytics at any time, concurrently, to get data to users faster

Scale compute to support any workload

Scale processing horsepower up and down on-the-fly, with zero downtime or disruption

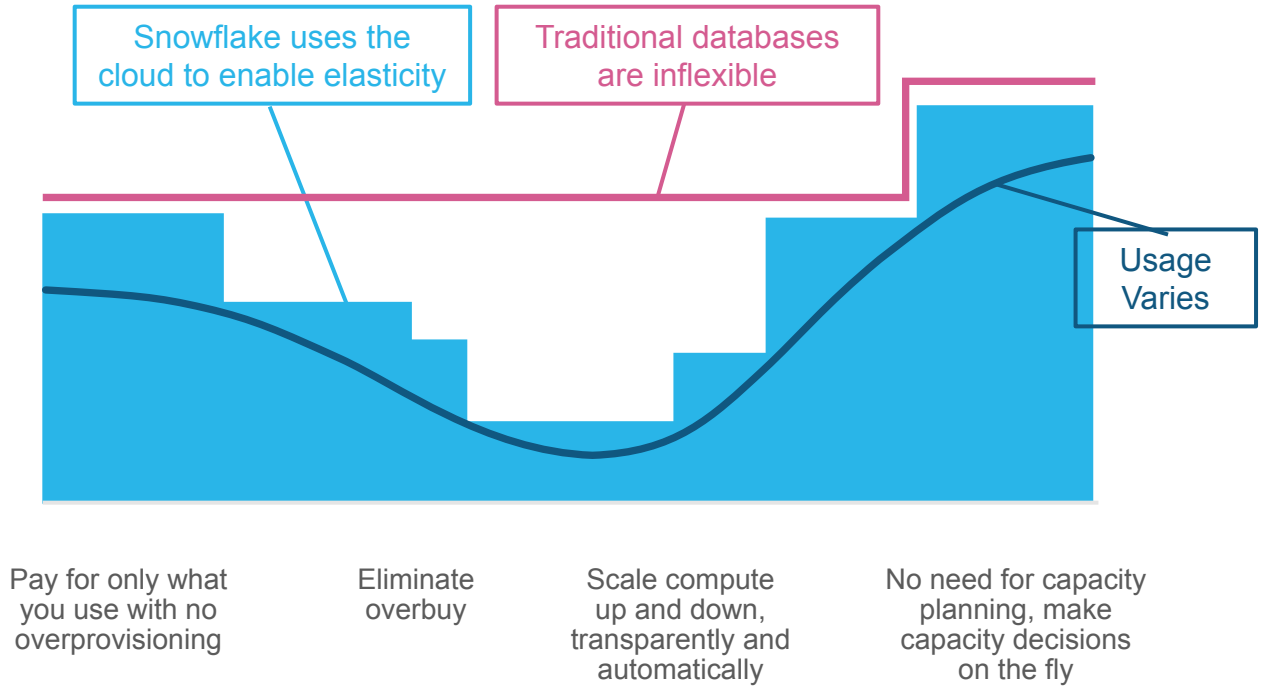
Scale concurrency without performance impact

Multi-cluster “virtual warehouse” architecture scales concurrent users & workloads without contention

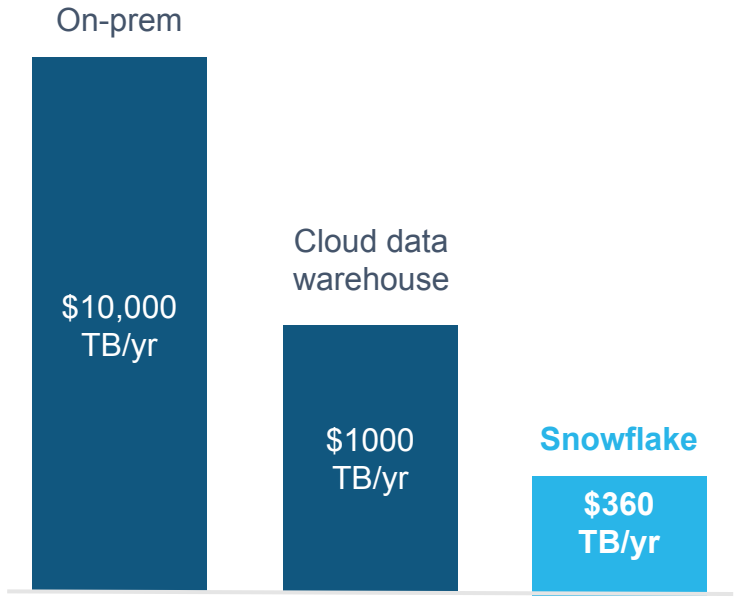
© 2019 Snowflake Inc. All Rights Reserved

TRUE ELASTICITY

Pay for only what you use, at cloud economies of scale



Pay only for what you use



Cloud economies of scale

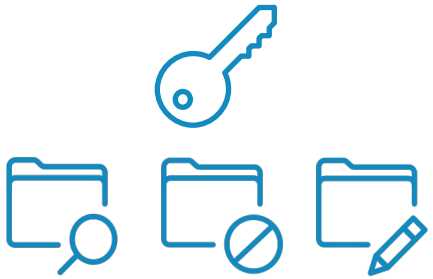
© 2019 Snowflake Inc. All Rights Reserved

#3 - SECURE BY DESIGN

Authentication

- Embedded multi-factor authentication
- Federated authentication available

Access control



- Role-based access control model
- Granular privileges on all objects & actions

Data encryption

- All data encrypted, always, end-to-end
- Encryption keys managed automatically

External validation

- Certified against enterprise-class requirements

© 2019 Snowflake Inc. All Rights Reserved

#2 – Automatic Query Optimization

Load data and run queries,
we do all the rest

Zero infrastructure and admin costs

Secure and highly available

Fully managed with no knobs
or tuning required

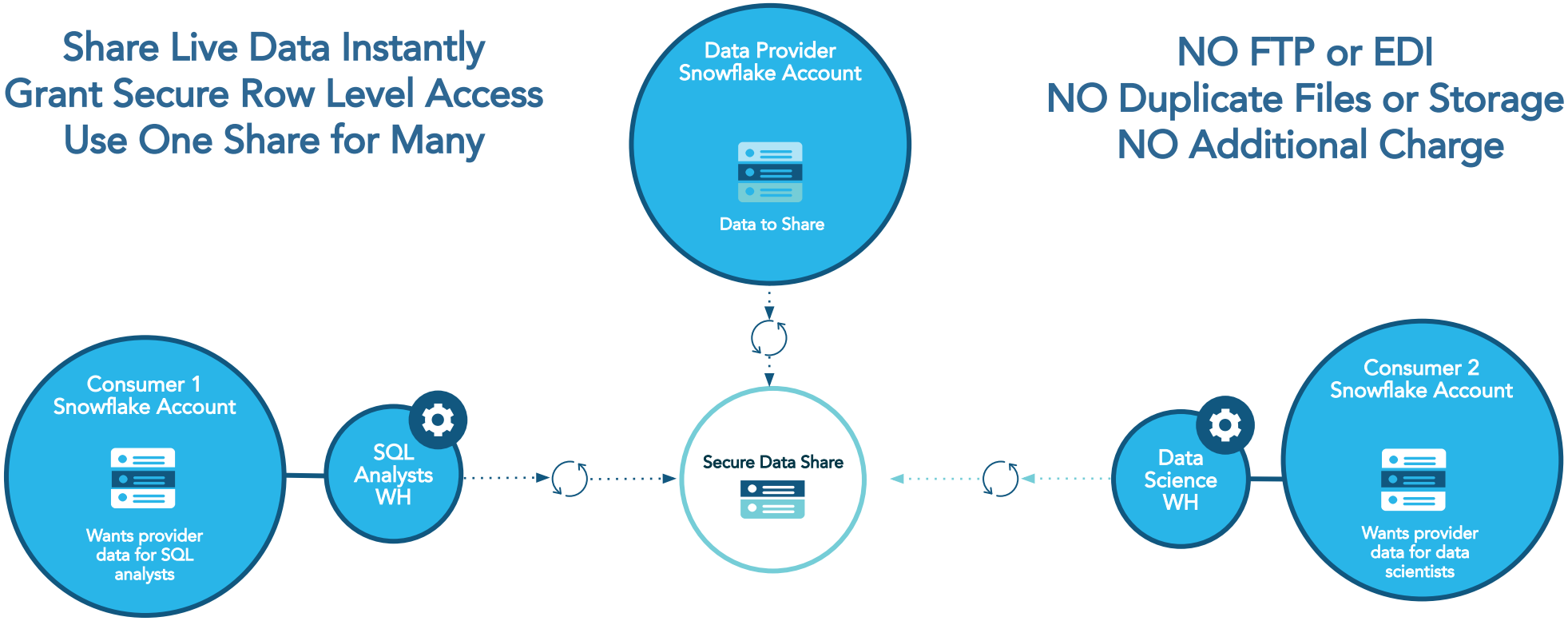
No indexes, distribution keys,
partitioning, or vacuuming

© 2019 Snowflake Inc. All Rights Reserved

#1 Live Data Sharing

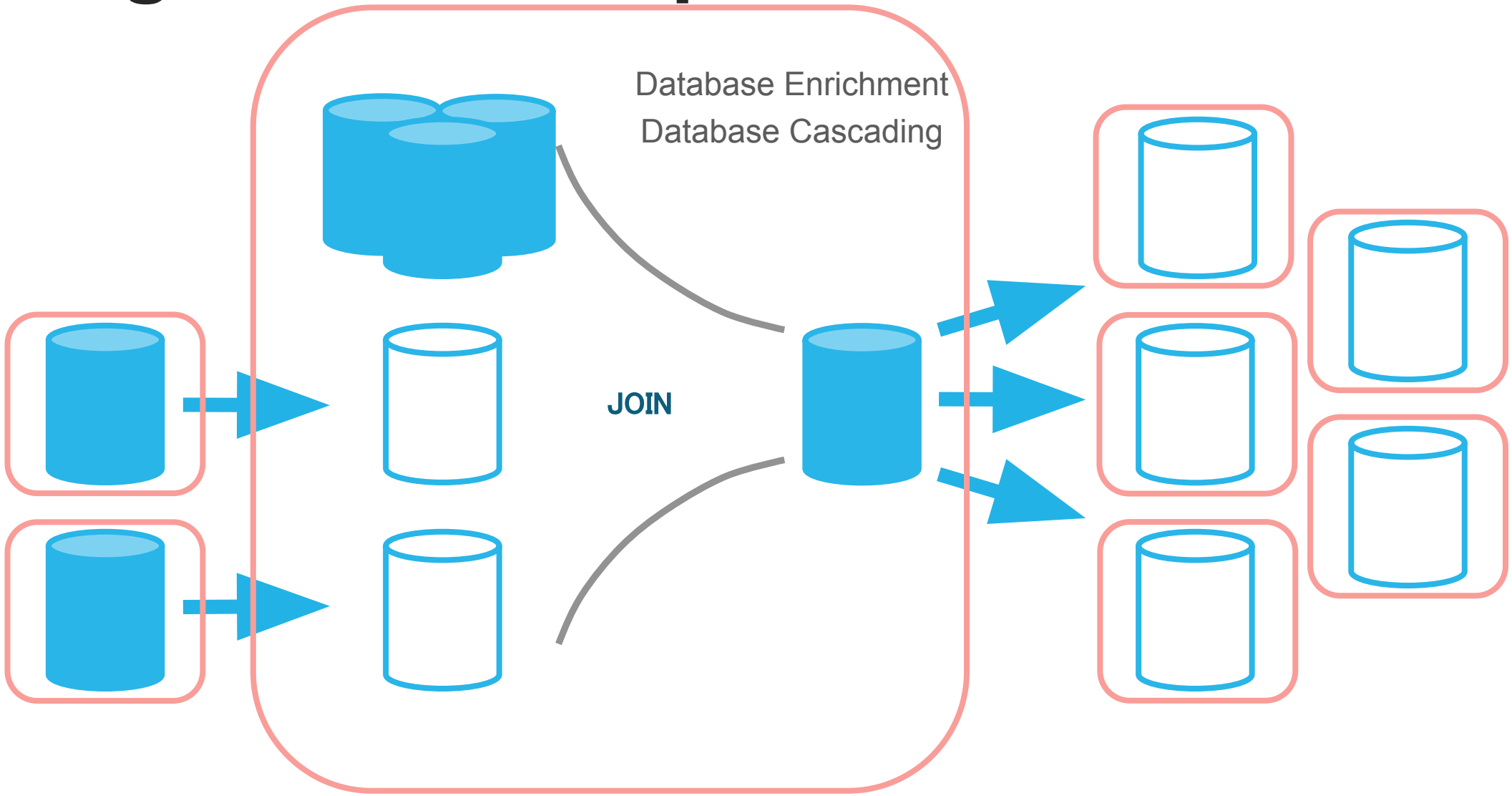
Share Live Data Instantly
Grant Secure Row Level Access
Use One Share for Many

NO FTP or EDI
NO Duplicate Files or Storage
NO Additional Charge



© 2019 Snowflake Inc. All Rights Reserved

Enabling Database Pipelines



© 2019 Snowflake Inc. All Rights Reserved

Snowflake in Action Today

© 2019 Snowflake Inc. All Rights Reserved

Common customer scenarios

noSQL replacement

Replace use of noSQL system (e.g. Hadoop) for transformation and SQL analytics of multi-structured data

Data warehouse for SaaS offerings

Use Cloud DW as back-end data warehouse supporting data-driven SaaS products

Data warehouse modernization

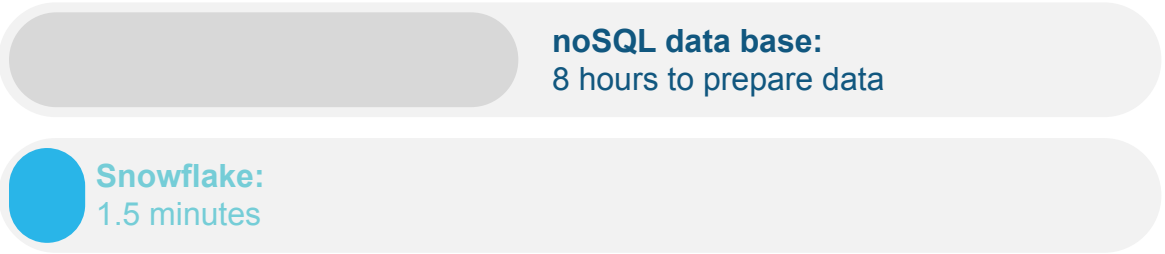
Consolidate legacy datamarts and support new projects

© 2019 Snowflake Inc. All Rights Reserved

Delivering compelling results

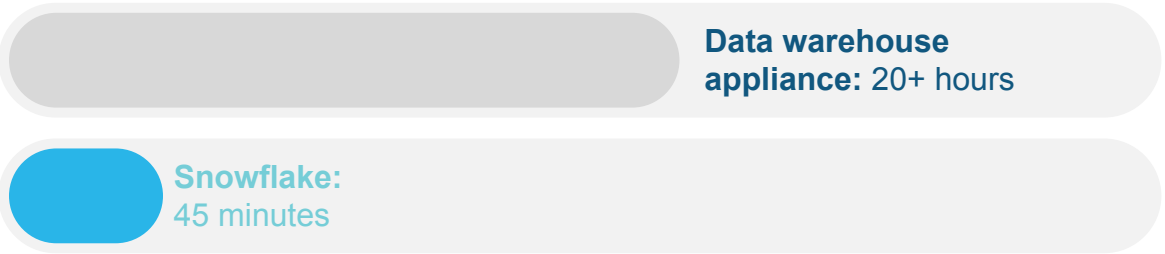
Simpler data pipeline

Replace noSQL database with Snowflake for storing & transforming JSON event data



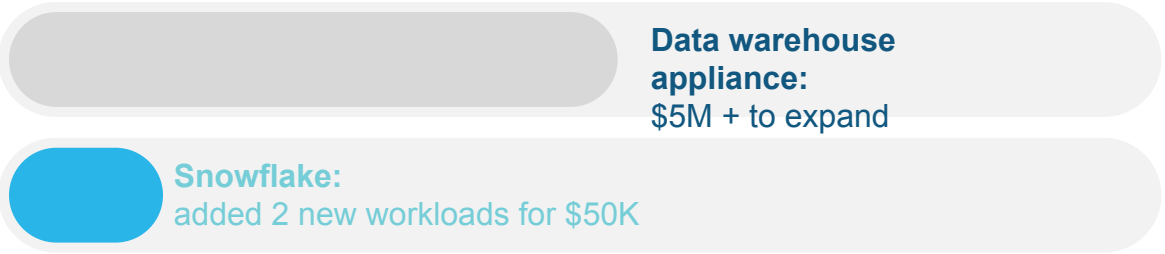
Faster analytics

Replace on-premises data warehouse with Snowflake for analytics workload



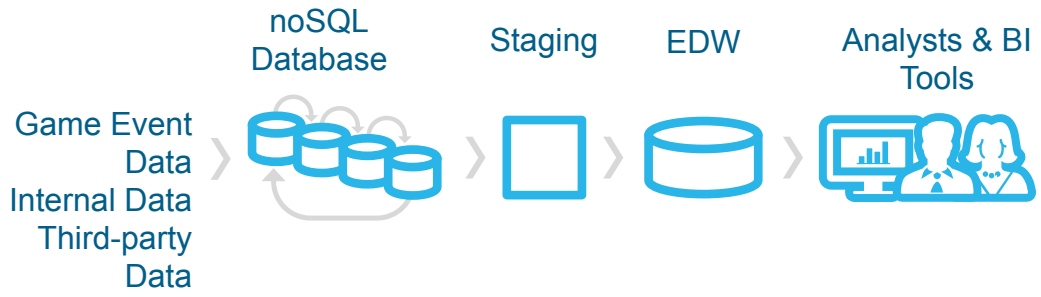
Significantly lower cost

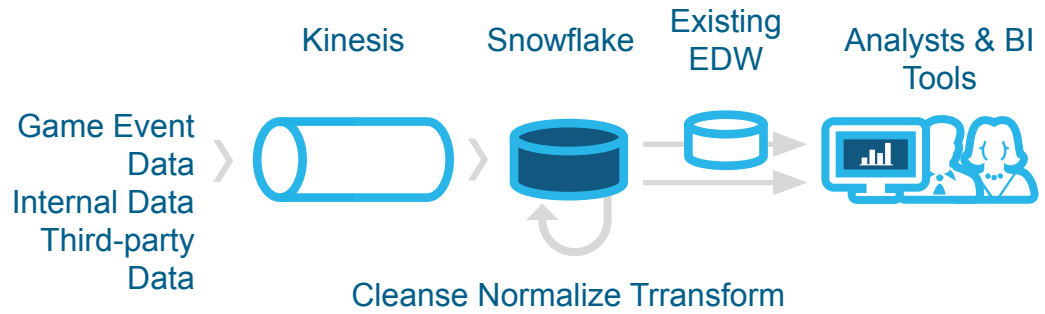
Improved performance while adding new workloads--at a fraction of the cost



© 2019 Snowflake Inc. All Rights Reserved

Simplifying the Data pipeline





11-24 hours

15 minutes

Scenario

Complex pipeline slowing down analytics

Pain Points

- Fragile data pipeline
- Delays in getting updated data
- High cost and complexity
- Limited data granularity

Solution

Send data from Kinesis to S3 to Snowflake with schemaless ingestion and easy querying

Snowflake Value

- **>50x** faster data updates
- **80%** lower costs
- Nearly eliminated pipeline failures
- Able to retain full data granularity

© 2019 Snowflake Inc. All Rights Reserved

DATA ANALYTICS AT EXTREME SCALE

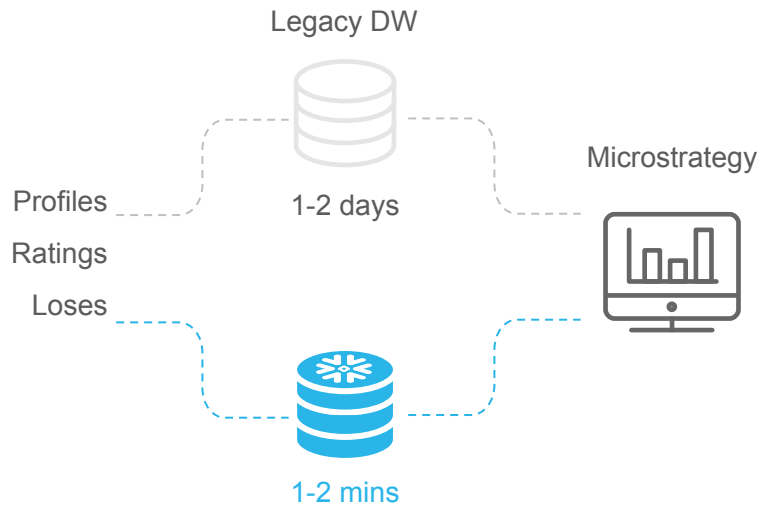
Scenario

- Financial institution with a huge focus on security
- Overburdened staff
- Business needs to run monthly reports that span 10 years of historical data
- No way to analyze semi-structured data

Pain Points

- Quoted $500,000 to replace their existing hardware appliance
- 20+ hours to run reports
- Could not continue to scale
- Users unable to query while performing ETL

CapSpecialty

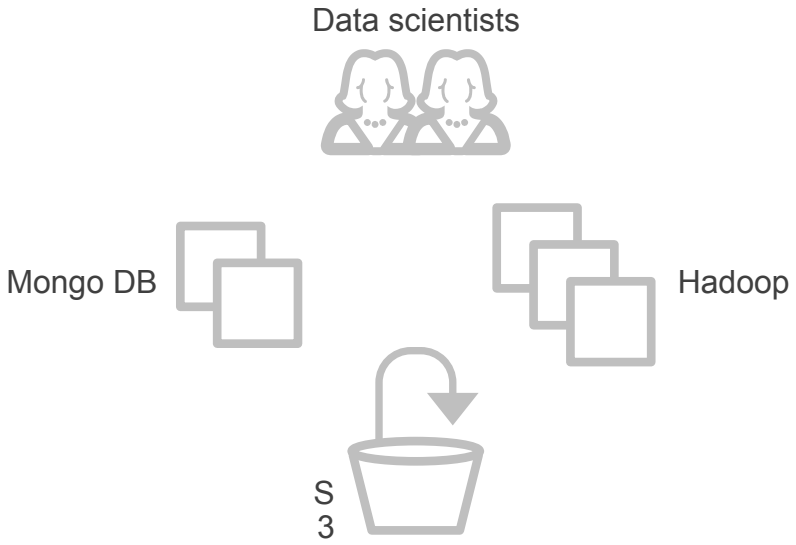


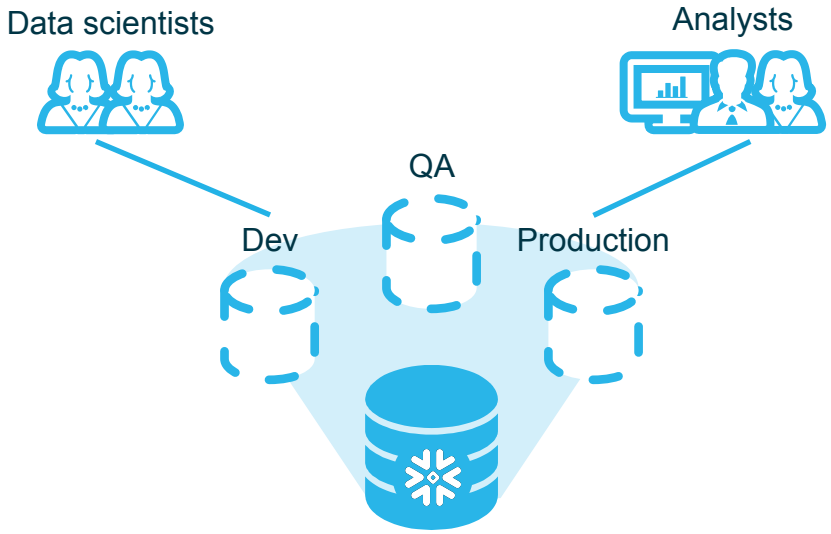
Snowflake Value

120x faster – from 20 hours to 45 minutes → Ad-Hoc analytics available to all users in minutes → Deployed in a week during the busiest time of year

© 2019 Snowflake Inc. All Rights Reserved

Data science and exploration





Scenario

- Security organization forced to use delayed reporting

Pain Points

- Many data requests unaddressed
- 24 hour turnaround time on requests
- 24 hours to push new models live

Solution

- Deploy Snowflake to accommodate analytics workloads

Snowflake Value

- 2 hours to push new models live
- Generated new research report
- Analysts can use data directly

© 2019 Snowflake Inc. All Rights Reserved

MORE THAN 2000 ENTERPRISE SUCCESS STORIES AND COUNTING

Discover the performance, concurrency, and simplicity of Snowflake

As easy as 1-2-3!

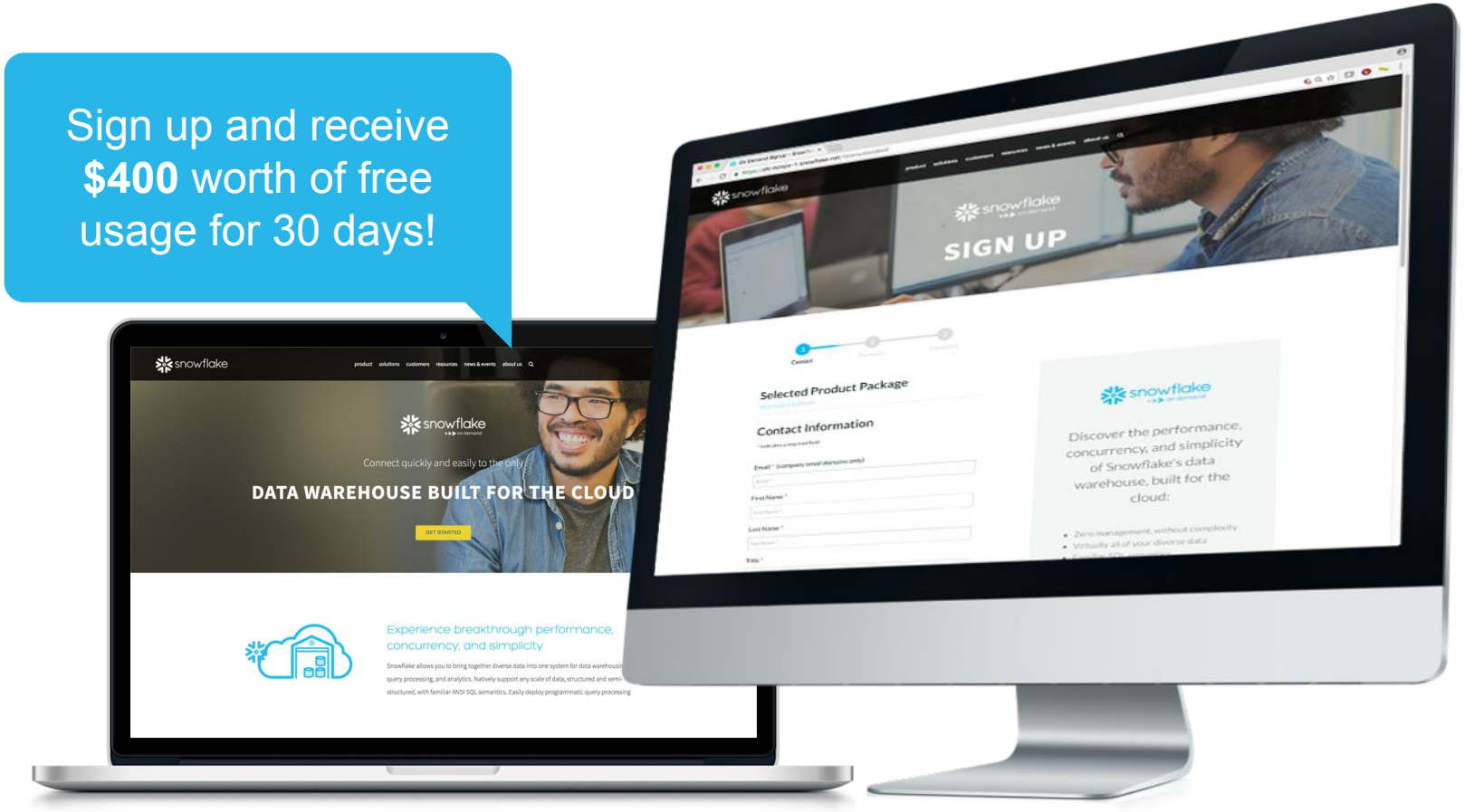
01 Visit Snowflake.net

02 Click “Try for Free”

03 Sign up & register

Snowflake is the only data warehouse built for the cloud. You can automatically scale compute up, out, or down—independent of storage. Plus, you have the power of a complete SQL database, with zero management, that can grow with you to support all of your data and all of your users. With Snowflake On Demand™, pay only for what you use.

Sign up and receive **$400** worth of free usage for 30 days!



© 2019 Snowflake Inc. All Rights Reserved

Big Data does not have to equal Big Effort

Contact Info

Kent Graziano
Snowflake Computing
Kent.graziano@snowflake.com
On Twitter @KentGraziano

More info at
http://snowflake.com

Visit my blog at
http://kentgraziano.com

The Data Warrior
Changing the world, one data model at a time. How can I help you?

© 2019 Snowflake Inc. All Rights Reserved

THANK YOU

© 2019 Snowflake Inc. All Rights Reserved