Start Ending Your Data Struggle
A 30 Day Plan

Alessando Pireno, Sales Engineer

Snowflake Computing

Common data struggles

Data Loading

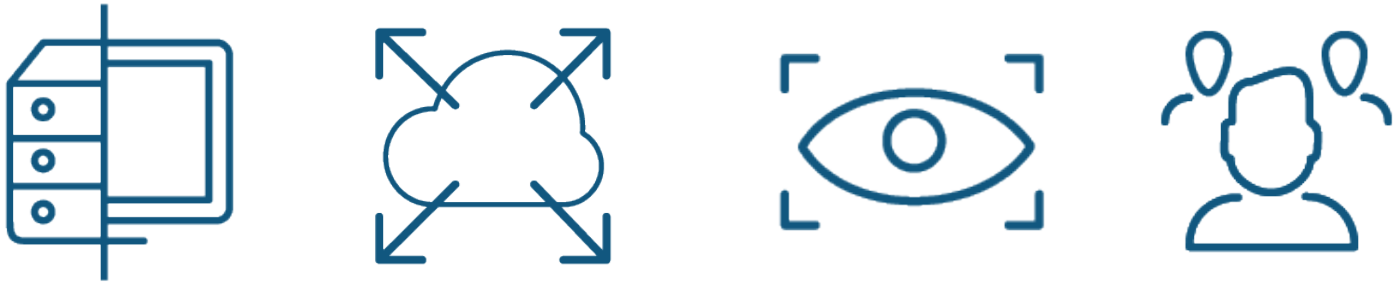
Data Integration

Analytics

Collaboration

Agenda

1. Common data struggles and how Snowflake addresses them



2. A practical 30 day plan to start ending your struggles with Snowflake

Data Loading

Struggle to Load Data

Preparing disparate data to load

– Have to flatten to store semi-structured (or use noSQL)

Capacity Planning

– Storage and compute are limited

Resource Contention

– Architecture forces linear compute capacity

"Where can I connect to that new JSON web log data?"
-BI Team

Tackle loading challenges with Snowflake

Disparate data ✓

- Variant column type supports semi-structured
- No more flattening (unless you want to)

Capacity ✓

- Built on the cloud (S3, EC2)
- Scale data and compute to load any data

Contention ✓

- Unlimited virtual warehouses allow independent compute
- Isolate loading and other tasks

Data Integration

Struggle to Integrate Data

Making sense of data in silos

– Hard to transform different datasets while in different silos/formats

Editing and transforming data

– noSQL tools complex, not all data stores ACID complaint
– Contention an issue while transforming

Support evolving business logic and disparate use cases

– No way to easily experiment with and add business logic
– Different people have different use cases

"Are the updated KPI's in the sensor data tables?"
- Data scientist

Improve data integration with Snowflake

Silos ✓

- Native storage for semi-structured, ANSI standard SQL and dot notation to use it
- Combine all of your data fluidly

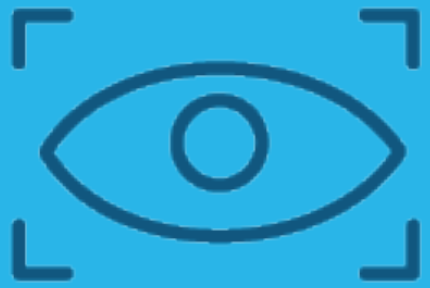
Editing and transforming ✓

- ACID compliant with virtual data warehouses
- Edit, transform, insert, delete, however or whenever you want

Business logic ✓

- Zero-copy cloning
- Rapidly iterate, test and promote business logic for multiple people

Data analytics



Struggle to Analyze Data

Queues

– Analysts are always the end of the resource priority queue

Delays

– Even with unlimited access, database is non-performant

"How come the dashboard isn't working?"
- Sales director

Analyzing Efficiently with Snowflake

Queues ✓

- Independent virtual warehouses
- Scale up, down or out to serve analytics use cases

Delays ✓

- Autoscaling and multi-cluster warehouses
- Automatically match compute to even massive demand

Collaboration

Struggle to Collaborate

Incessant fixing

– Fixing loading, integration and analytics struggles burns time
– Conflicts from those struggles reduce morale

Siloed teams

– Technical and business teams often not working together (physically or otherwise)

"I'm so buried under this queue I can't make the BI standup"
- IT team member

"I could ask IT for an updated table, but I'm not sure who was working on it."
- BI team member

Start Collaborating with Snowflake

Fixing ✓

- Address the other struggles as referenced
- Free more time for collaboration and discussion

Siloed teams ✓

- With new time, start new discussions around data
- Build updates and additions into a scheduled meet-up

A 30-day Plan to Start Ending Your Struggle with Snowflake

Start Ending Your Data Struggle – Week 1

Start from the beginning – what's the analytics goal?

1. Define the team
2. Discuss blocking issues and a place to start
3. Define the scope
4. Define success criteria
5. Try Snowflake On-Demand
6. Plan status updates going forward

Start Ending Your Data Struggle – Week 2

1. Find data to load

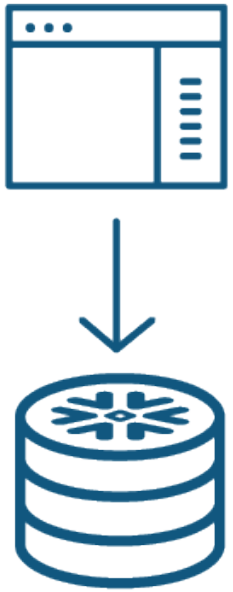
- Work within defined scope, agree as a team
- Use data that's new, challenging, or semi-structured

2. Create a Warehouse

- Will need this to load data

3. Load data

- Create a database and a table
- Stage your data
- Load from stage to database



Start Ending Your Data Struggle – Week 3

1. Test and deploy business logic

– Discuss metrics, KPIs, transformations to add
– Use zero-copy cloning to test and then promote

2. Optional: Create Integration WH

– Isolate integration and transformation

3. Optional: Plan ongoing loading and transform

– Use zero-copy cloning to test iterations safely and promote

Start Ending Your Data Struggle – Week 4

1. Create Warehouses for BI

- Avoid queues with isolated compute resources
- Optionally, set up auto-scaling

2. Create analytics users

- Spread the value of the data
- Use this as an opportunity to share and discuss

3. Connect your BI to Snowflake

- Use Tableau, Looker, etc. to query your data live
- Consider publishing dashboards with live connect

After 30 days you should see improvements

1. Your team should be talking and collaborating more
2. You should be able to easily load and combine data
3. You should have accurate business logic in your data
4. You should be finding more insight

Thank You to Our Partners

Platinum

amazon web services

informatica + UNICO SOLUTION

looker

talend

Gold

alooma

alteryx

Cervello

Fivetran

Silver

ATTUNITY

Datasticians

Decisive Data

DUNN

INQUIDIA CONSULTING

snapLogic

SONATA SOFTWARE

TEKNION data solutions

USEREADY

snowflake

CLOUD ANALYTICS CITY TOUR