

Visualization Survival Guide

BI on Snowflake during the zombie apocalypse

bit.ly/okc-snowflake-5



LEFT 4 DEAD 2



Coach
BI Developer



Rochelle
DBA



Nick
Data Engineer



Ellis
PM



Coach
BI Developer



Rochelle
DBA



Nick
Data Engineer



Ellis
PM

Coach | BI Developer

Know your tools.

Decouple your dashboards.

Iterate often.

Coach | BI Developer

PowerBI and Excel (PowerQuery):

- No SSO
- Affordable
- PowerBI publishing requires gateways

Tableau:

- Expensive
- Weird bugs during querying (timezones)
- SSO, but can only use default roles
- Best dashboard development experience

Coach | BI Developer

Sigma Computing:

- No installations
- Best integration with Snowflake
- Painful dashboard development experience

Metabase:

- Free!
- Open source!
- Deployable on premise or in the cloud.
- Finicky backend

Coach | BI Developer

Decouple your dashboard development from underlying table structure.

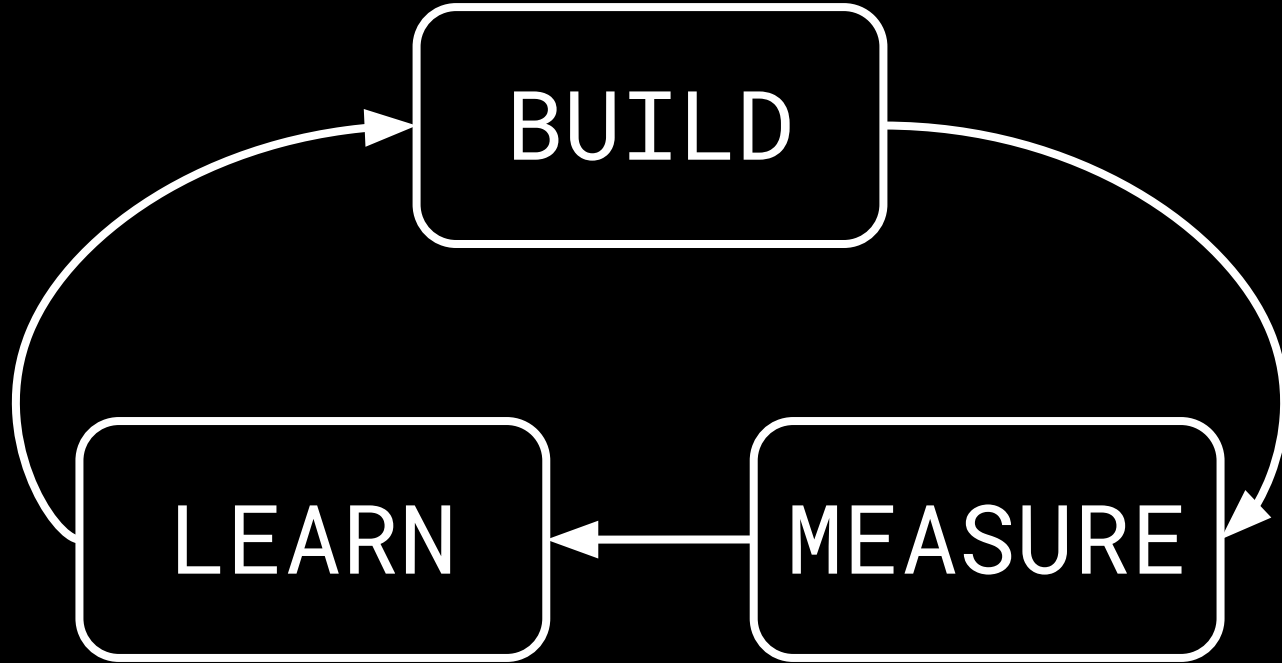
Custom SQL: `SELECT * FROM DB.SCHEMA.TABLE`

Coach | BI Developer

Decouple your dashboard development from underlying table structure.

Custom SQL: `SELECT * FROM DB.SCHEMA.TABLE`

Coach | BI Developer





Coach
BI Developer



Rochelle
DBA



Nick
Data Engineer



Ellis
PM

Rochele | DB Administrator

Use dedicated warehouses.

Size warehouses appropriately.

Use service accounts.

Rochele | DB Administrator

Do not do this:

- TABLEAU_WH
- DATABRICKS_WH

Do this:

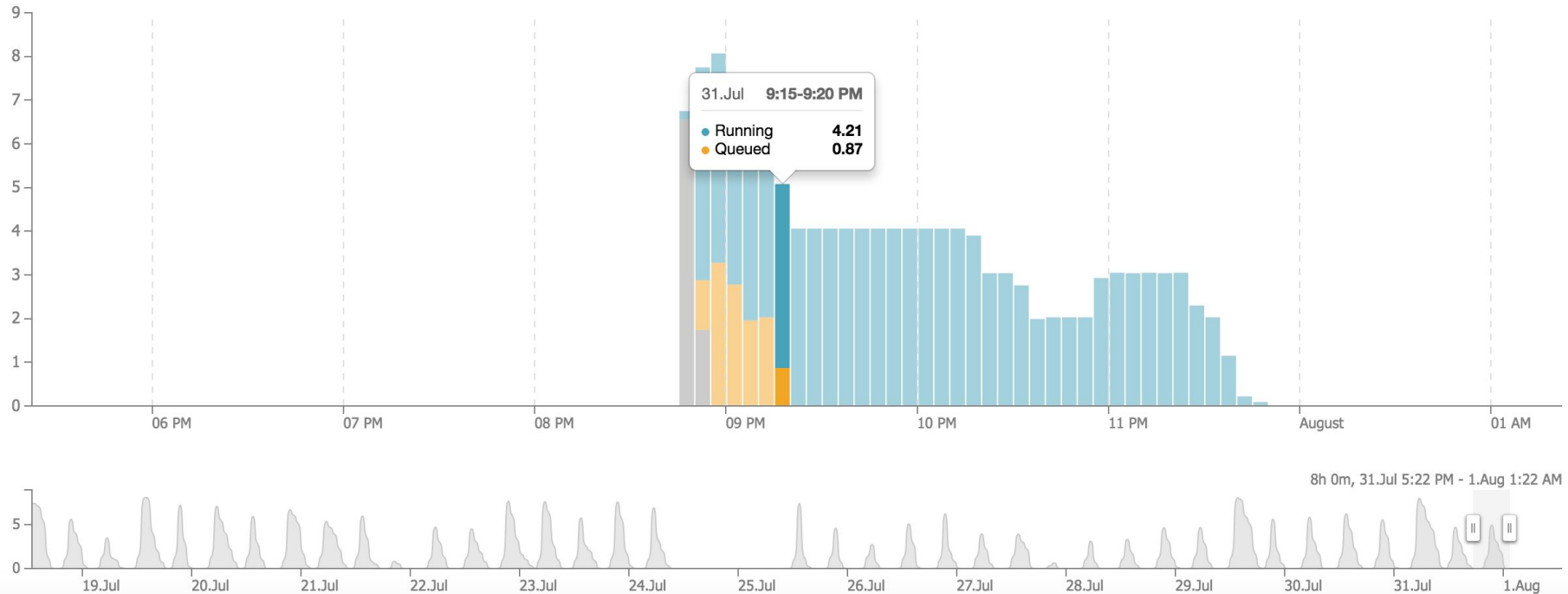
- TABLEAU_SECURITY_WH
- TABLEAU_OPS_EFFICIENCY_WH
- TABLEAU_SALES_KPI_WH
- DATABRICKS_REVENUE_PRED_WH
- <WHO>_<WHAT>_WH

Roche11e | DB Administrator

Increase warehouse size to speed up a single query.

Multic1uster to increase the number of queries a warehouse can handle at once.

Rochele | DB Administrator



Rochele | DB Administrator

Snowflake SS0 has a serious flaw:

- Only a default role can be used.
- This leads to over-permissioning and habitual use of highest privilege access.

Use BI service accounts to:

- grant minimum access to the account.
- tightly couple the warehouse to the service account (same rules for service accounts as warehouse names).



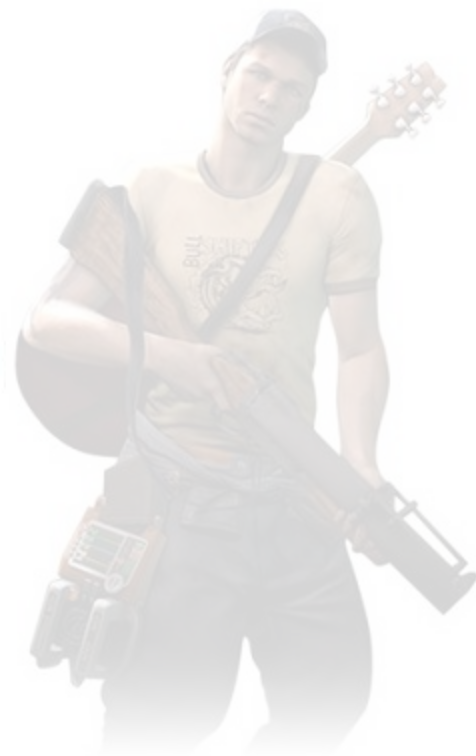
Coach
BI Developer



Rochelle
DBA



Nick
Data Engineer



Ellis
PM

Nick | Data Engineer

Know how to build analytics tables.

Know your table types.

Use the query profiler.

Nick | Data Engineer

Analytics Tables / Data Marts / Top Tables:

- DENORMALIZE data for better BI.
- Avoid VARIANT columns.
- Use (partial) row hashing.
- Always store more data to save warehouse credits

Nick | Data Engineer

Views:

- Cheap to build, cheap to destroy
- Worse performance (expensive to read)

Materialized Views:

- Extremely limited today (1 source table)
- Cheap to build, but expensive to maintain.
- Better read performance than views (cheap to read).

Nick | Data Engineer

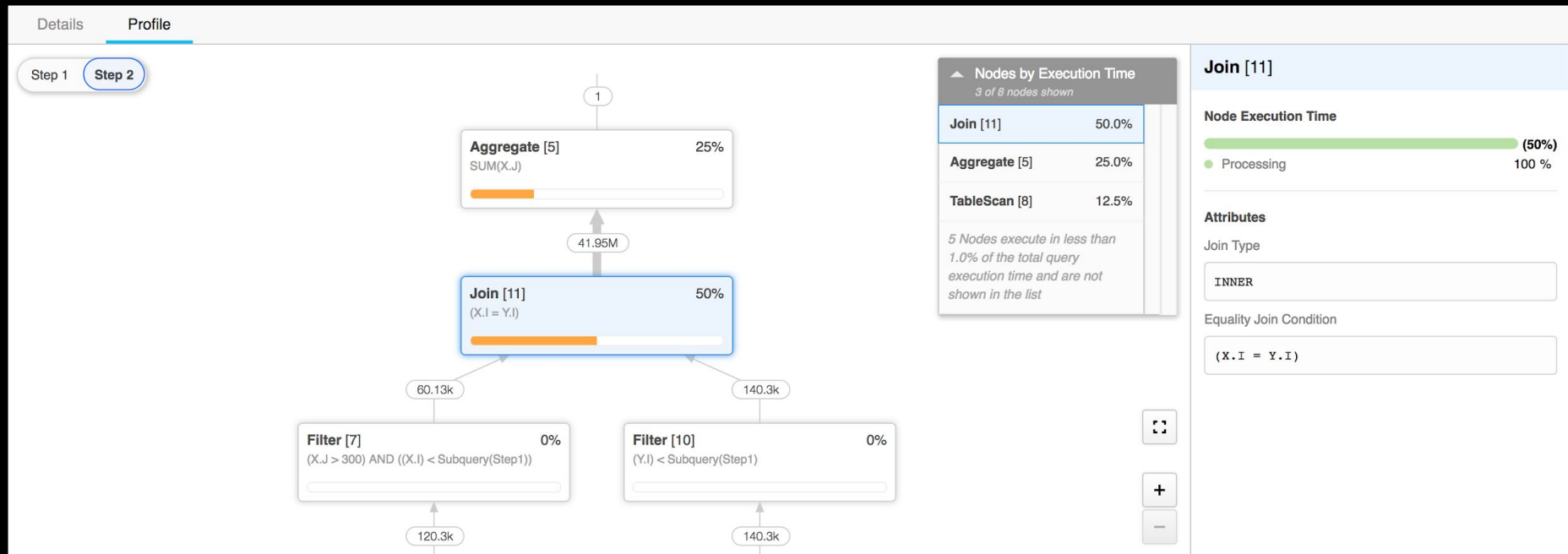
Table:

- Typically requires more complex transformation pipelines.
- Best read performance.
- Time travel

Clone:

- Ideal to get the benefits of a high-churn table with reduced cache impacts.

Nick | Data Engineer





Coach
BI Developer



Rochelle
DBA



Nick
Data Engineer



Ellis
PM

Ellis | Project Manager

Uncover problems for the team.

Remove problems for the team.

questions?

case study

bit.ly/okc-snowflake-5