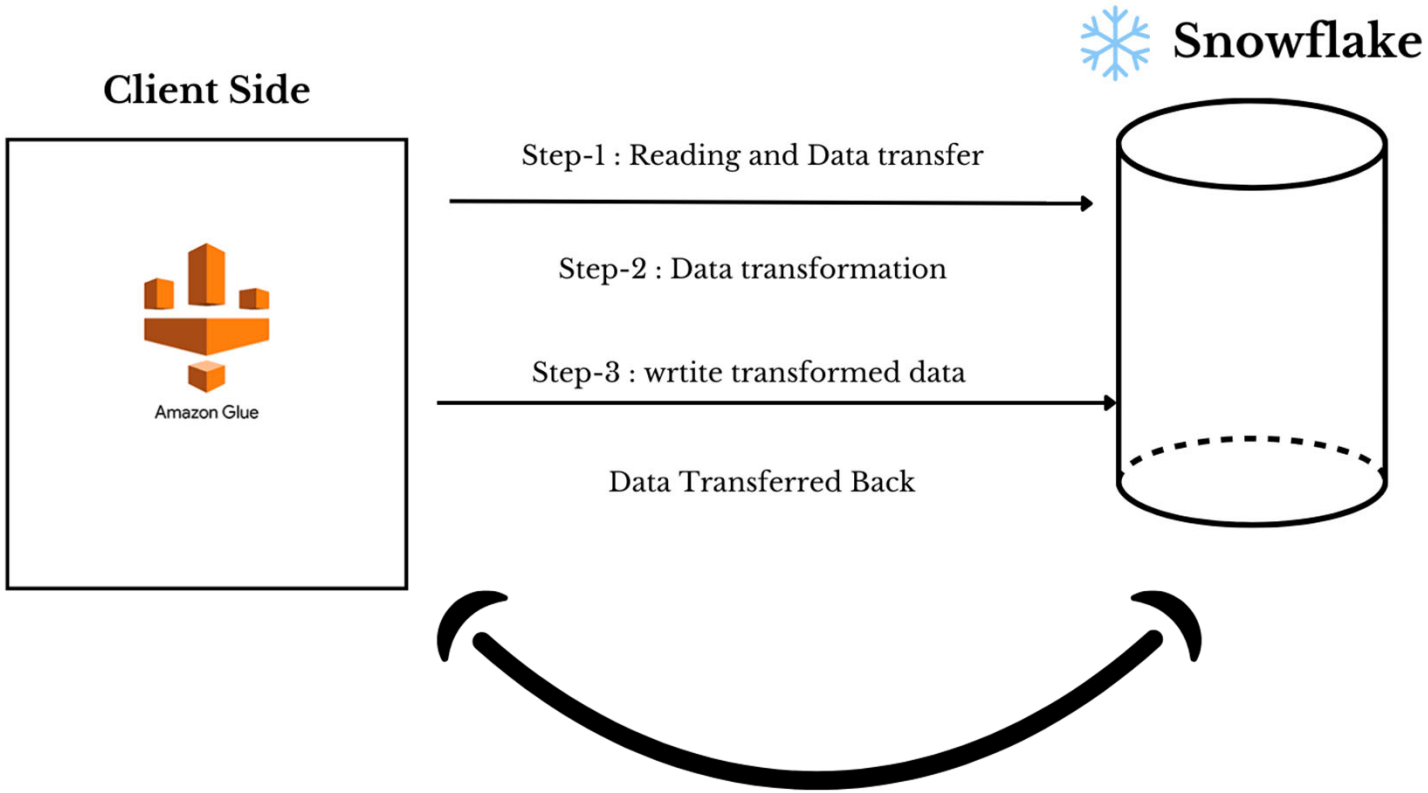
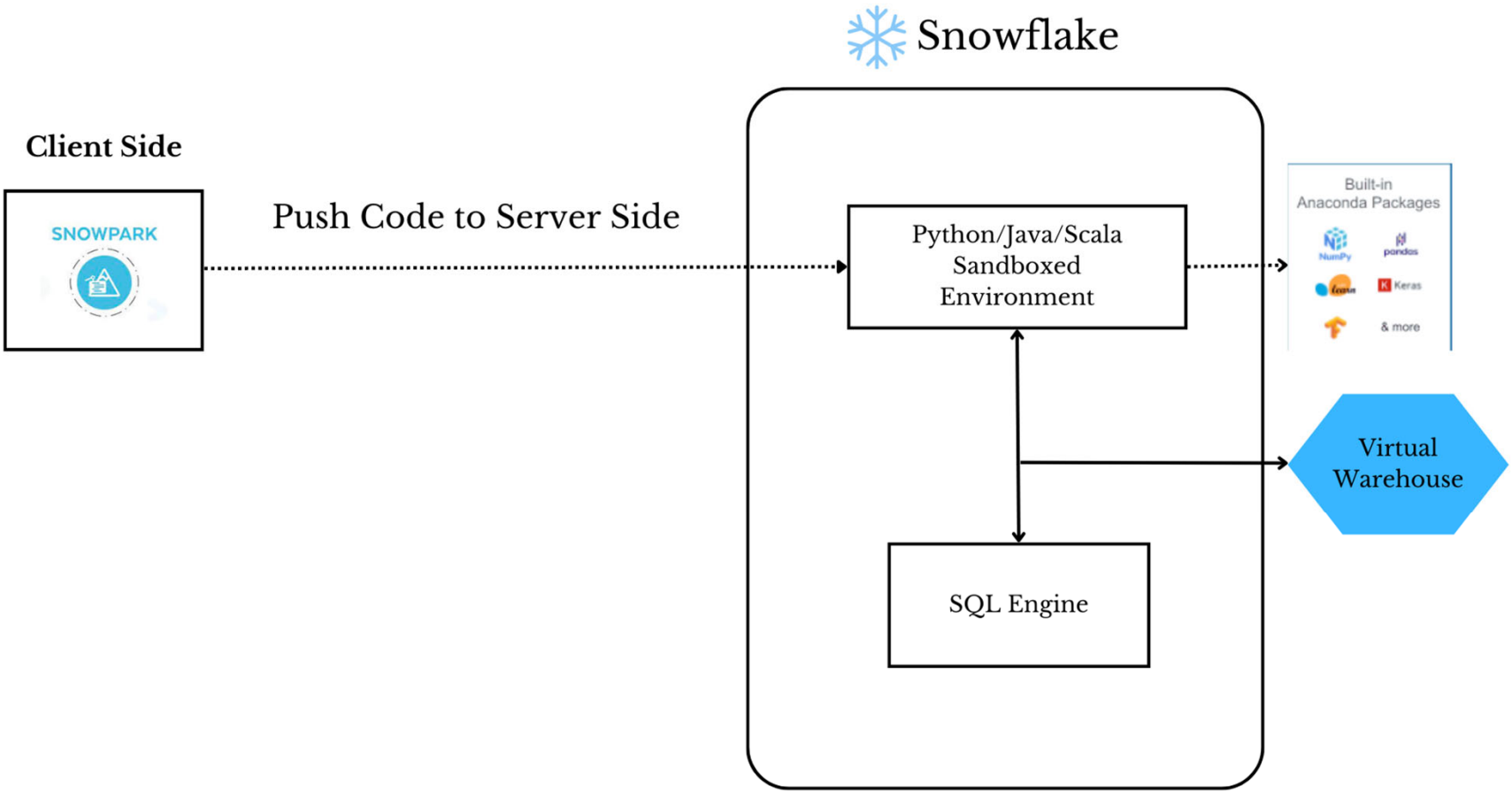
SNOWPARK API

- The Snowpark library provides an intuitive API for querying and processing data in a data pipeline
- The core abstraction in Snowpark is the DataFrame and also supports embedded SQL
- **Use-Cases :**
  - Data Transformations using Dataframes
  - Machine Learning deployments
- Snowpark operations are executed lazily on the server
- With Snowpark , you reduce or eliminate the need to move data outside of your snowflake servers
- Snowpark does not require a separate cluster outside of Snowflake for computations as all of the computations are done within Snowflake


Snowflake
Client Side
Amazon Glue
Step-1 : Reading and Data transfer
Step-2 : Data transformation
Step-3 : wrtite transformed data
Data Transferred Back
Data Moves to the location of the code and back


Snowflake
Client Side
SNOWPARK
Push Code to Server Side
Python/Java/Scala
Sandboxed
Environment
Built-in
Anaconda Packages
NumPy
pandas
Keras
& more
Virtual
Warehouse
SQL Engine

SNOWPARK DRAWBACKS

- **Architectural / Design complexity**
  - SQL only/low-code design to a more developer/engineering oriented approach

- **Vendor dependency / Non-cloud-agnostic pipelines**
  - Snowpark only works with Snowflake