Modern Cloud-based Data Stack Architecture

The Modern Data Stack Architecture

The modern data stack is a set of tools that together compromise a new approach to building data warehouse, from ingestion to transformation and serving of the data. They comprise of a set of independent tools and technologies that work together to enable businesses to collect, process, store and analyze data at scale.

These tools and technologies include the following:

* Tooling for ingesting data (e.g python, Stitch, fivetran, airbyte)
* A cloud data warehouse (e.g Snowflake, Redshift, BigQuery, Databricks, Azure Synapse)
* Transformation tools ( e.g dbt, python)
* Orchestration tools like airflow, prefect, dagster
* Data visualization tools like Looker, Mode, Power BI, Shiny
* Data Quality and Governance like great expectations
* Monitoring and alert: datadog, pageduty, airflow, acceldata

My preference: