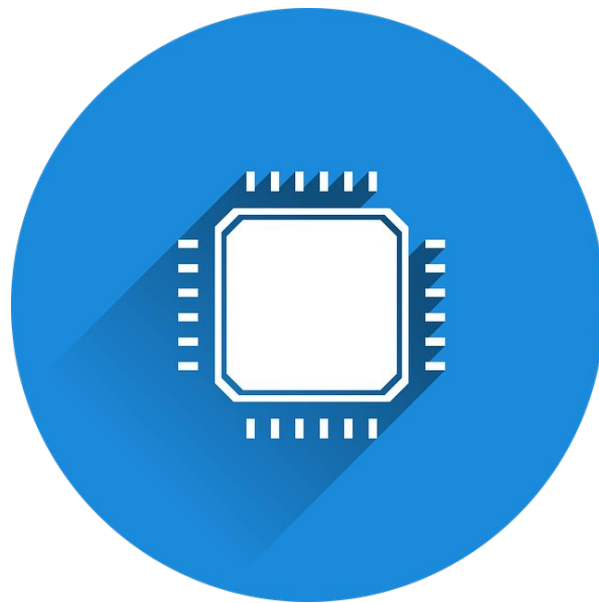


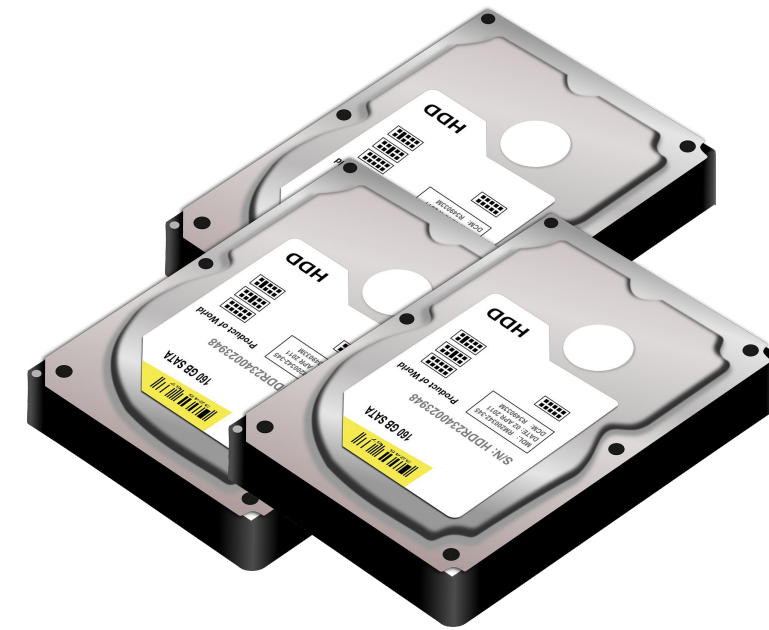


Summary

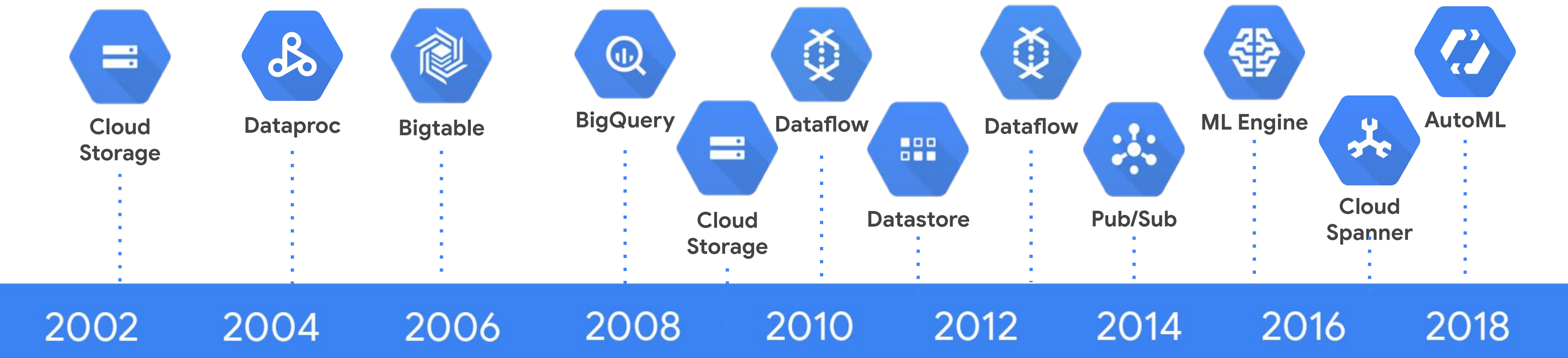
Google's data center network speed enables the separation of compute and storage



1 Petabit/sec of total bisection
bandwidth



Google Cloud opens up that innovation and infrastructure to you



There are many roles in a data-driven organization

Data Analyst

Analyst

Data Engineer

Data Engineer

Applied ML Engineer

Data Scientist

Ethicist

Statistician

Social Scientist

Applied ML
Engineer

Researcher

Tech Lead

Analytics Manager

Decision Maker

Big Data Challenges

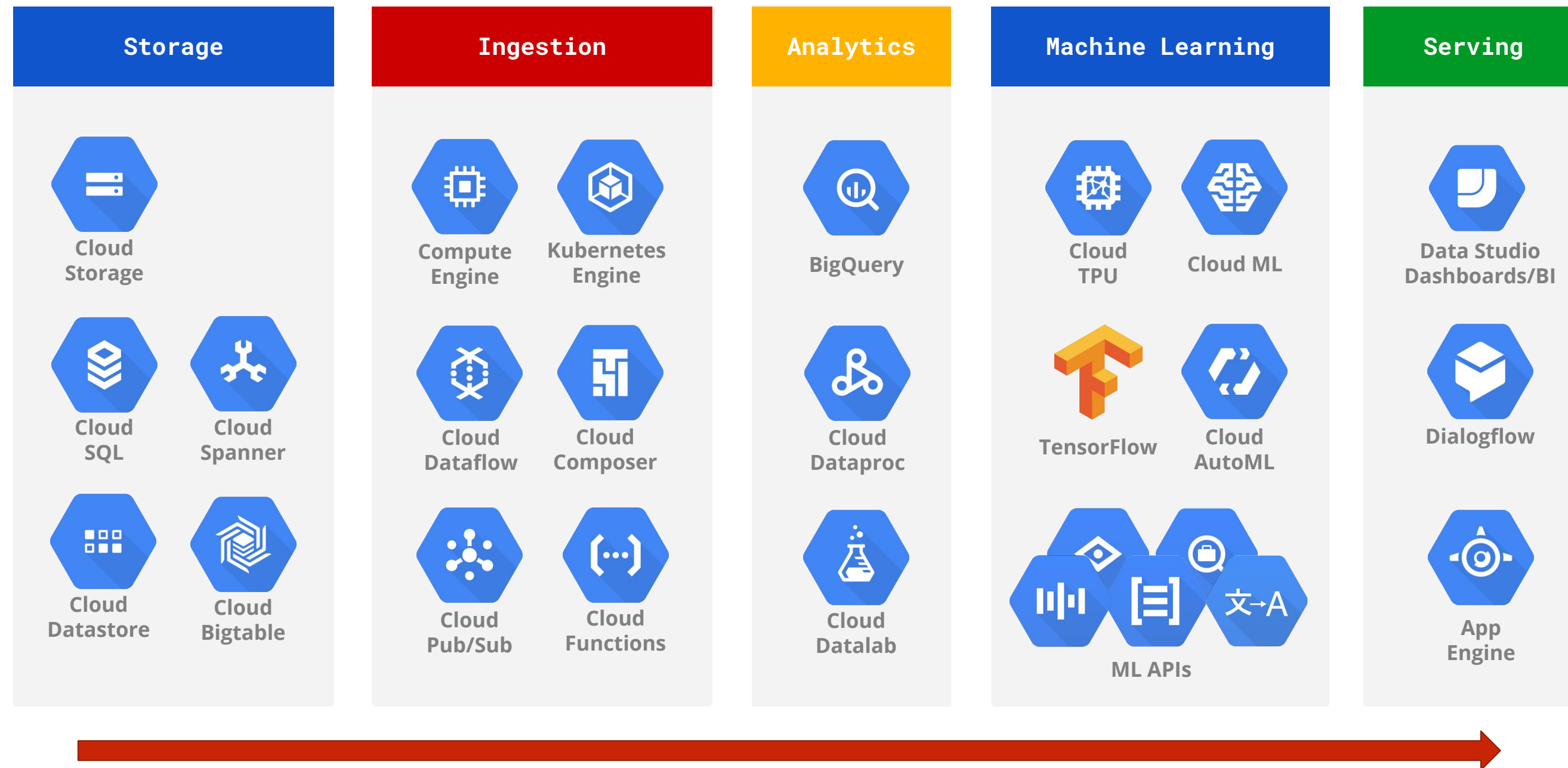
**Migrating existing
data workloads
(ex: Hadoop, Spark jobs)**

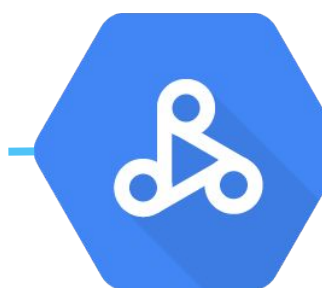
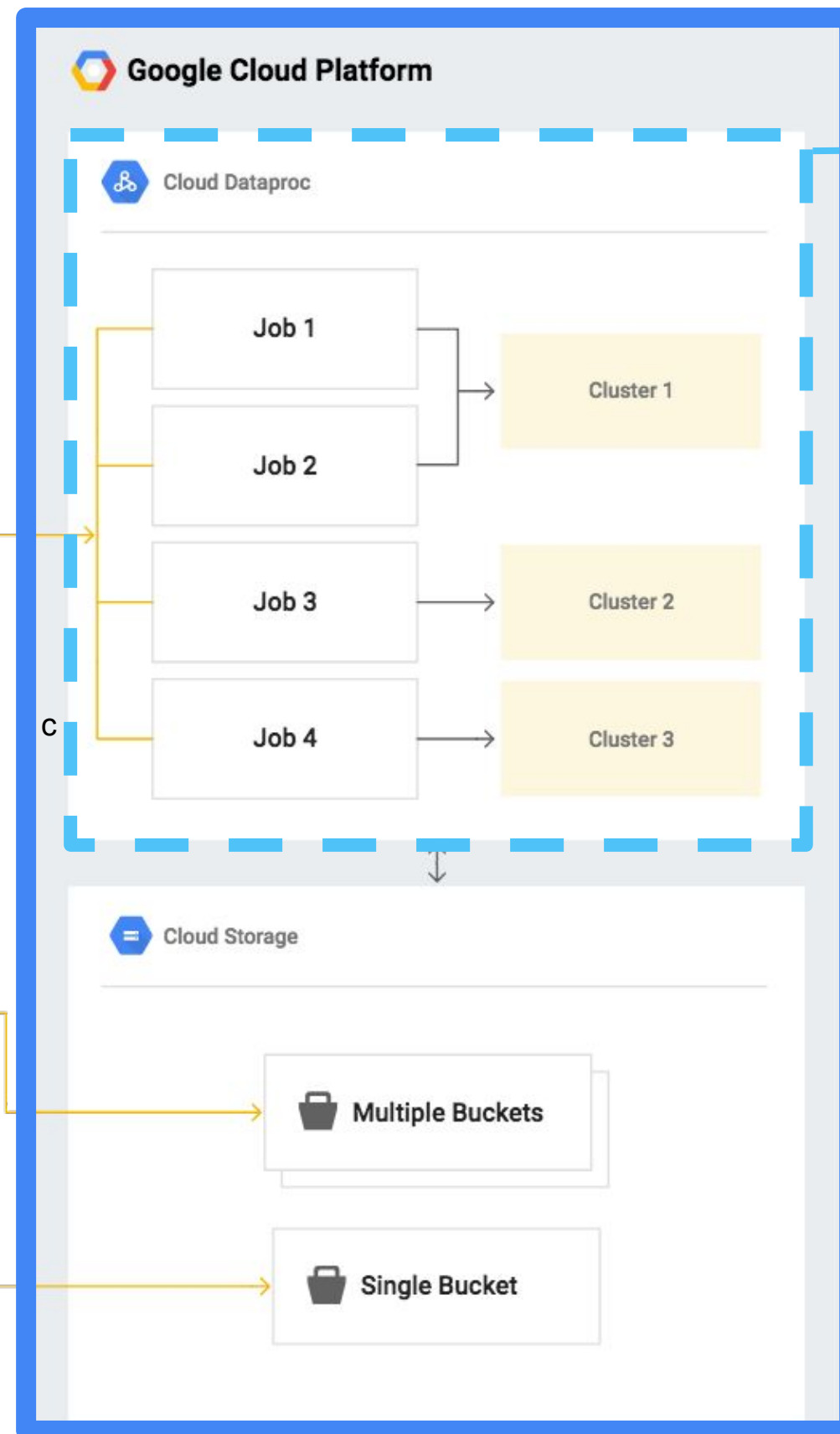
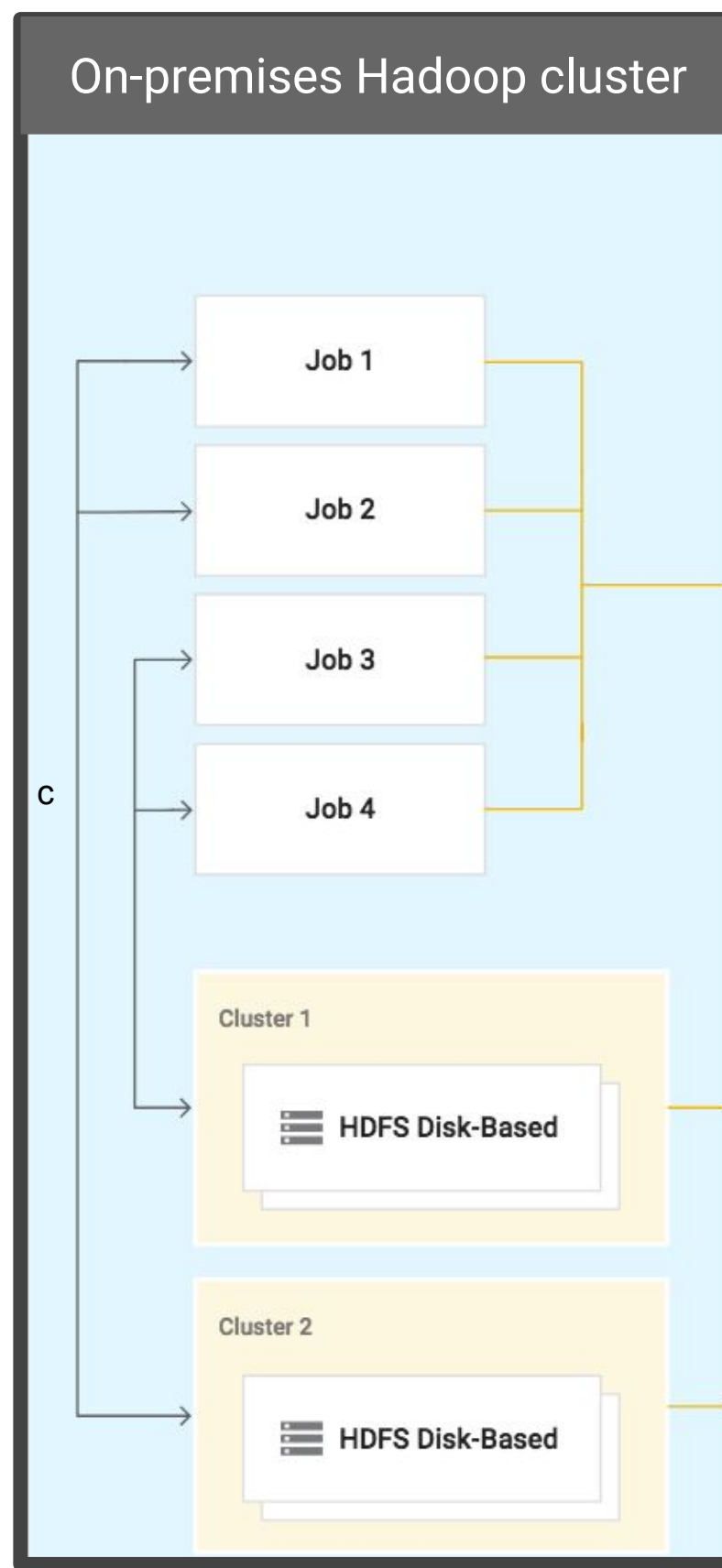
**Analyzing large
datasets at scale**

**Building streaming
data pipelines**

**Applying machine
learning to your data**

The suite of big data products on Google Cloud Platform





Cloud
Dataproc

Quickly launch,
resize, and manage
clusters to meet
workload needs

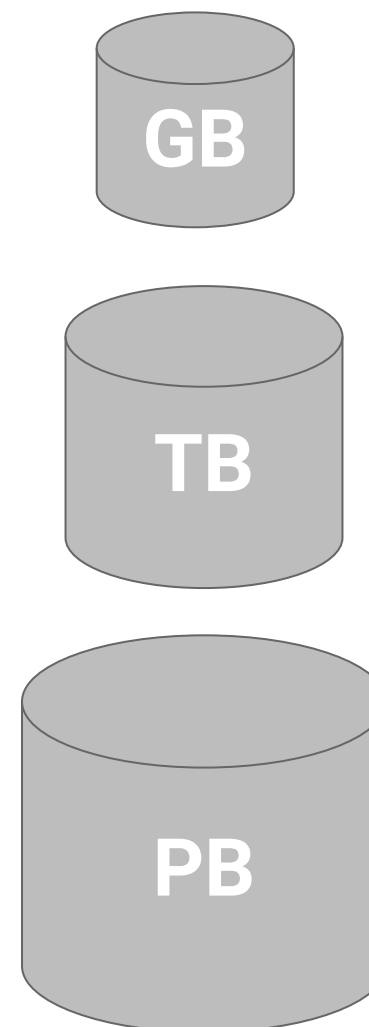


Modern big data pipelines face many challenges

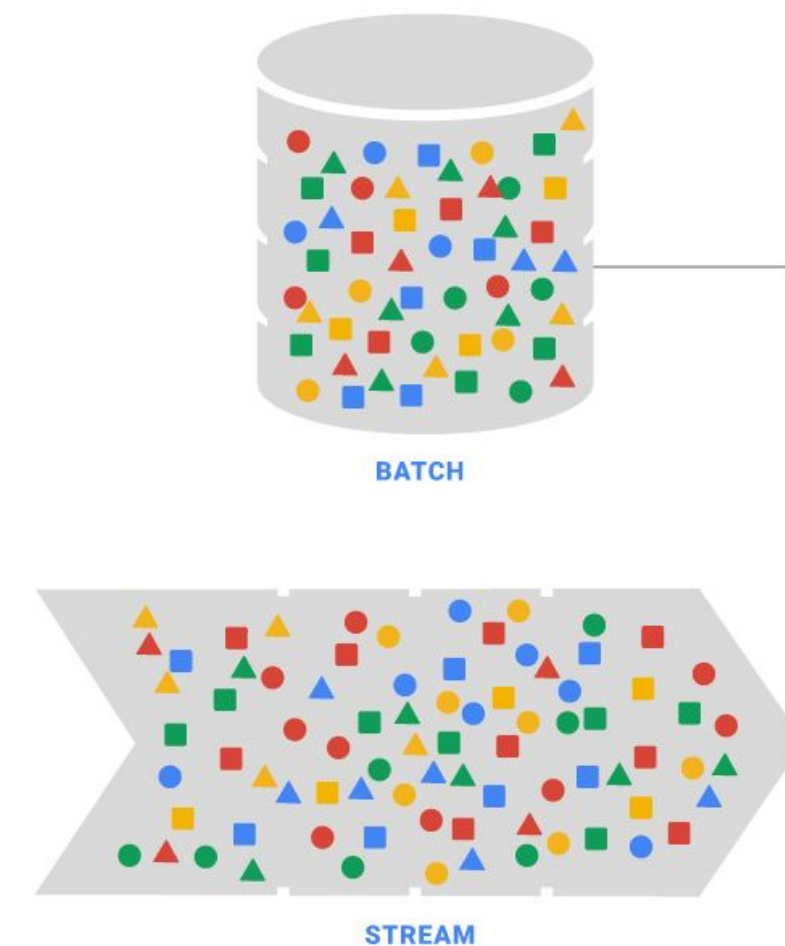
Variety



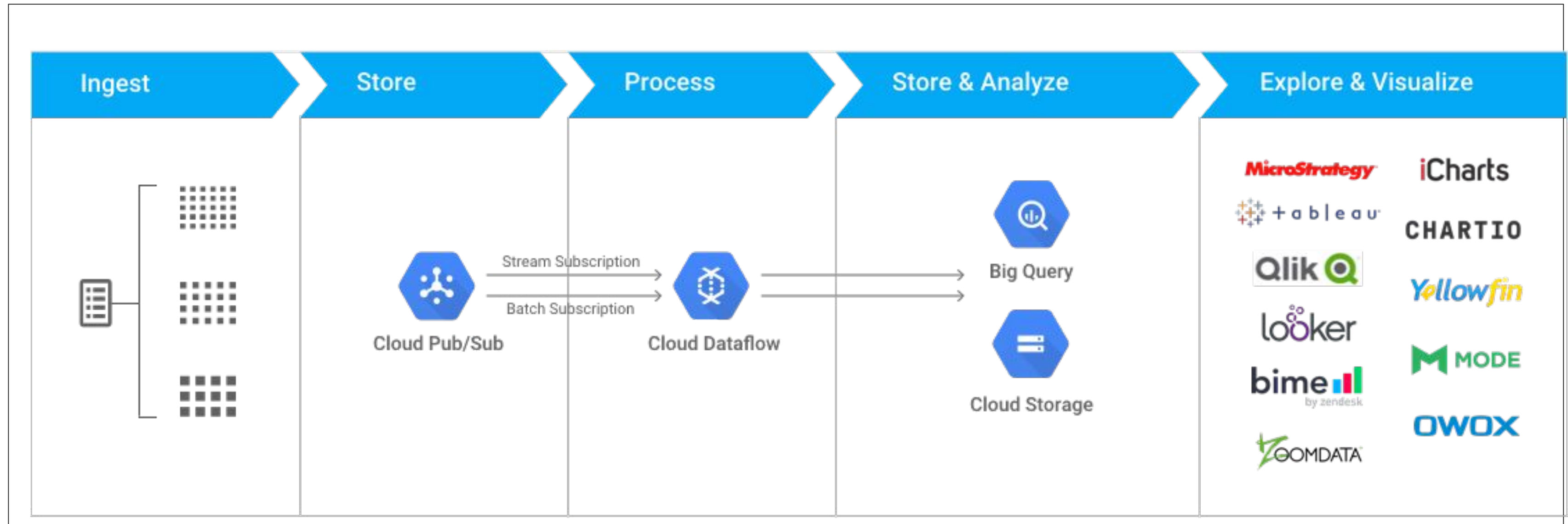
Volume



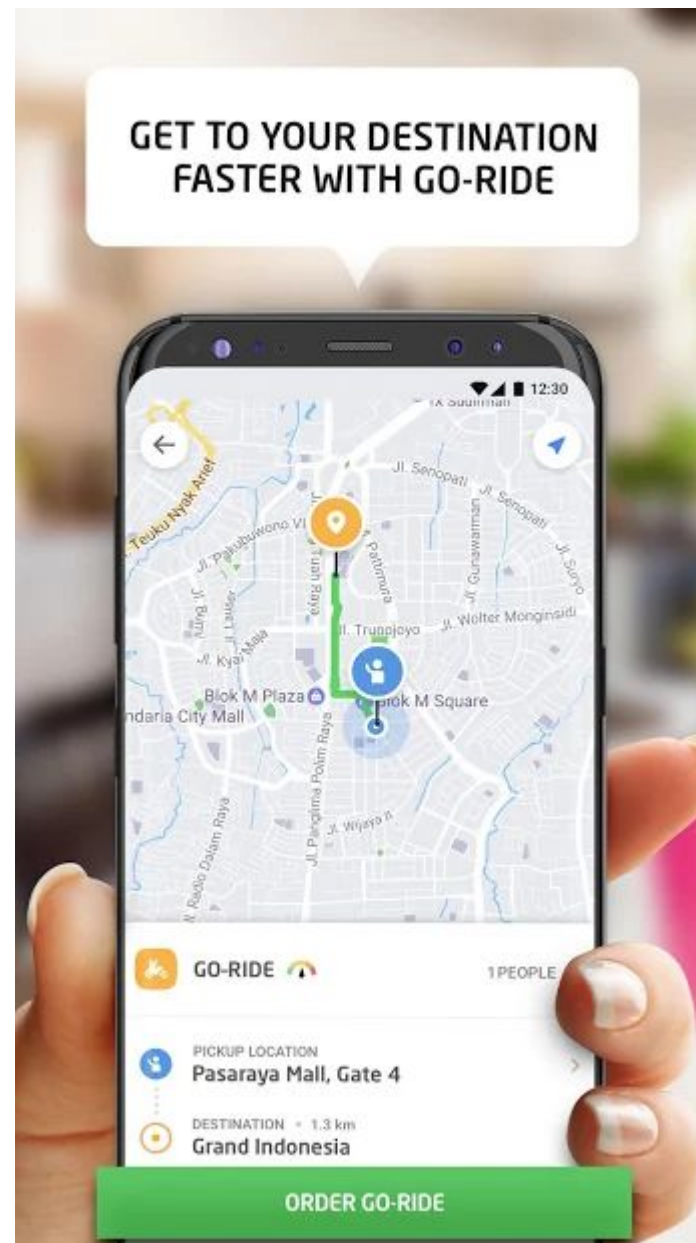
Velocity



Google Cloud Serverless Big Data Pipeline



GO-JEK brings goods and services to over
2 million families in 50 cities in Indonesia



Recap: BigQuery ML



1

Dataset

2

Create/train

3

Evaluate

4

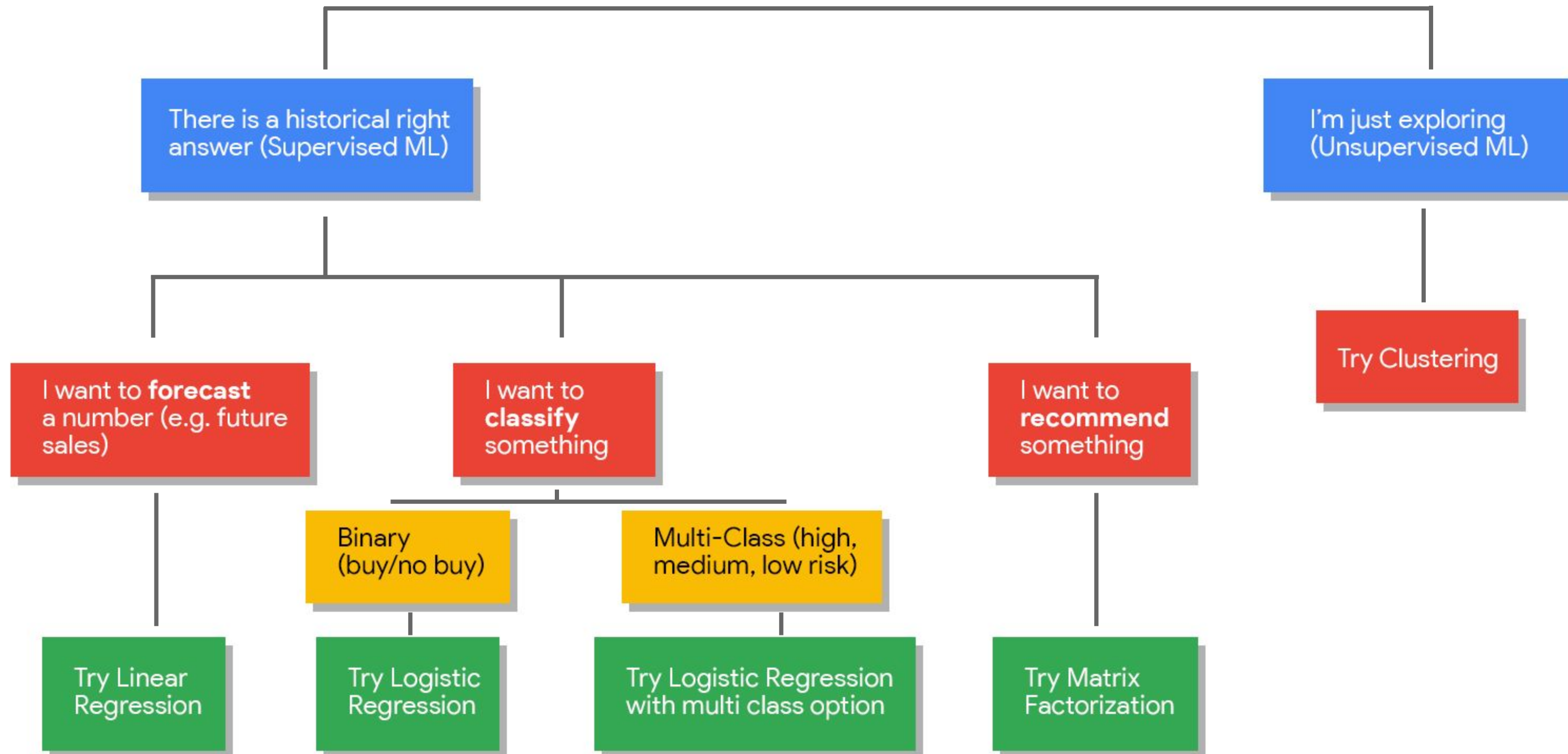
Predict/classify

```
CREATE MODEL `bqml_tutorial.sample_model`  
OPTIONS (model_type='logistic_reg') AS  
SELECT
```

```
FROM  
  ML.EVALUATE (MODEL  
    `bqml_tutorial.sample_model`,  
    TABLE eval_table)
```

```
FROM  
  ML.PREDICT (MODEL  
    `bqml_tutorial.sample_model`,  
    table game_to_predict) )  
AS predict
```

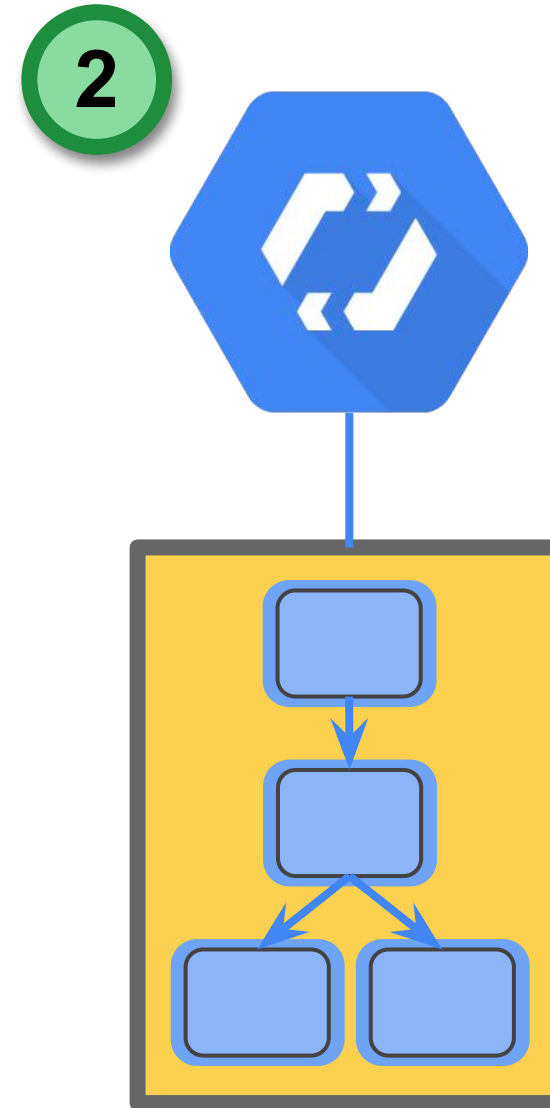

Choose the right model type for your structured data use case



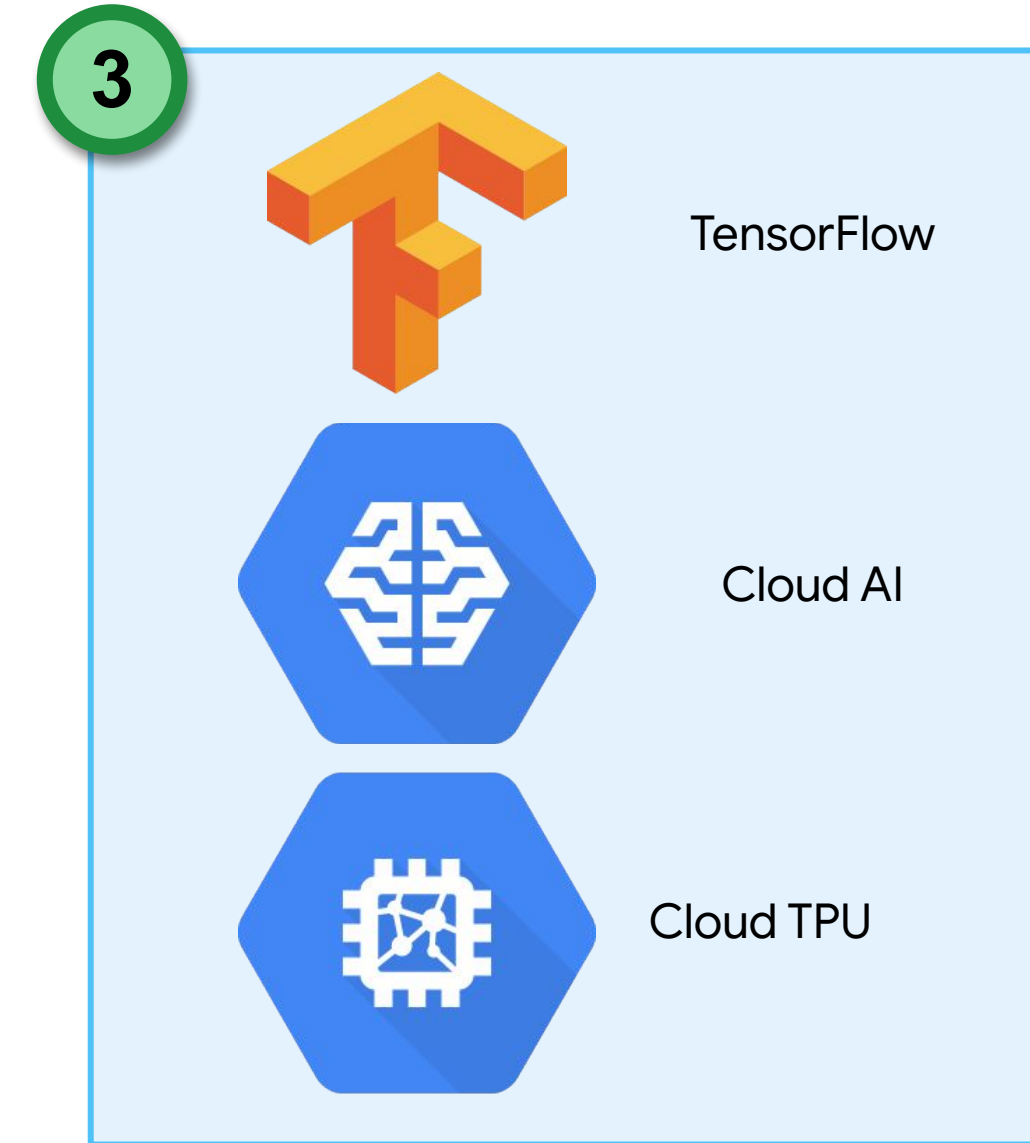
Artificial Intelligence application strategy



Use pre-built AI



Add Custom Models



Create new Models

This course is the start to your big data journey

**Big Data & ML
Fundamentals**

From Data to Insights

Data Engineering

ML on GCP

Adv ML on GCP

400 + Self-Paced Qwiklabs:
[qwiklabs.com/catalog](https://www.qwiklabs.com/catalog)

Public Github repo:
github.com/GoogleCloudPlatform/training-data-analyst/

Thank you!