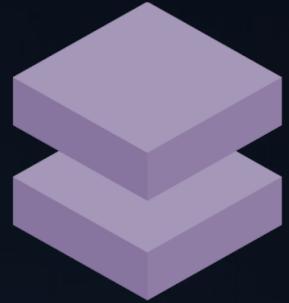




ONE WAY
SOLUTION



One Way Solution

Patterns & Common Use-Cases

Data Engineering – [Day 5]

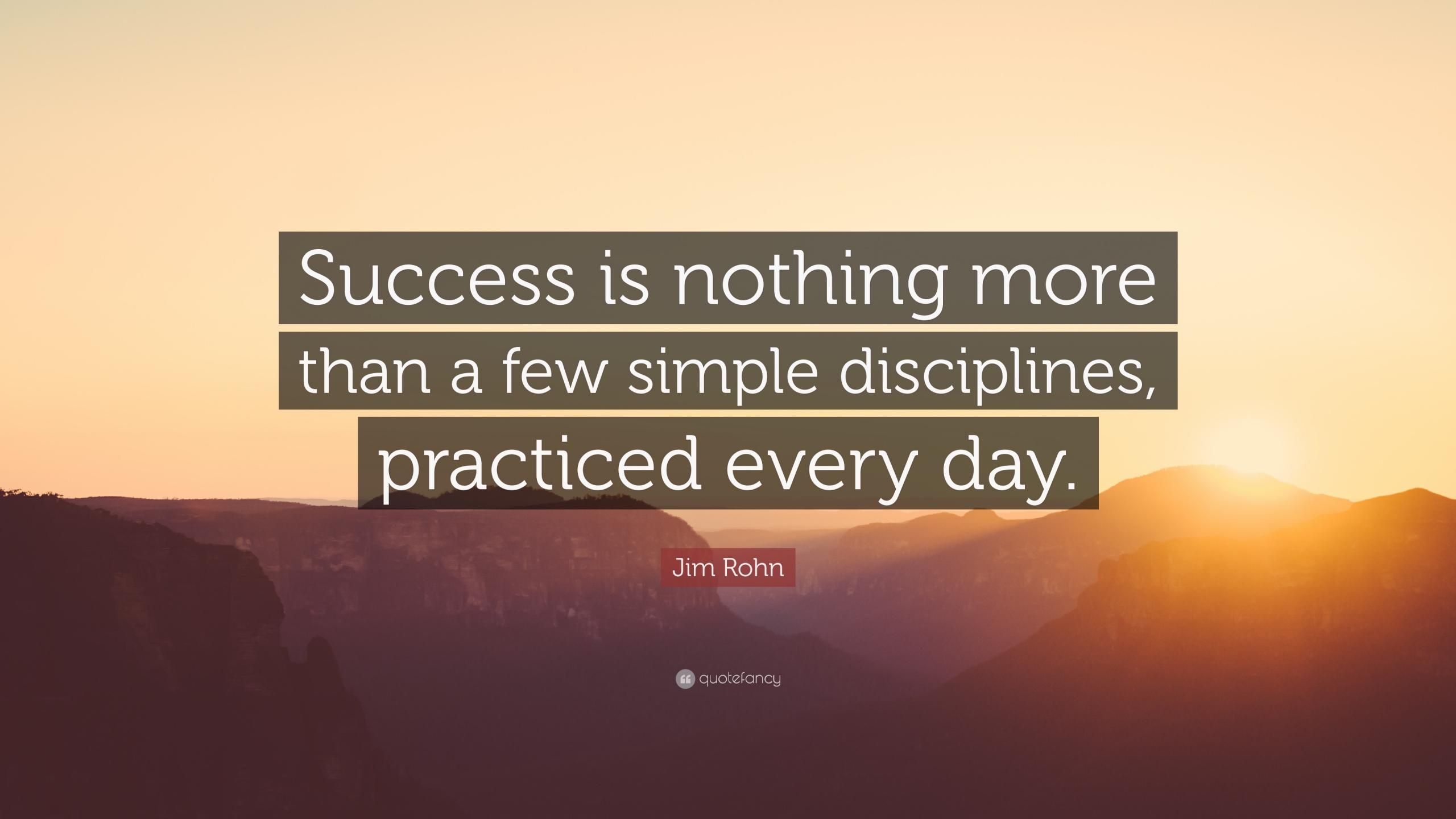


LUAN MORENO

CEO & CDO

Data Engineer & Data Platform MVP





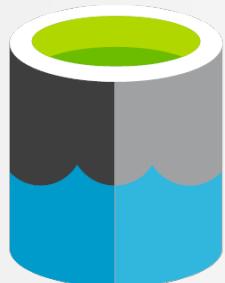
Success is nothing more
than a few simple disciplines,
practiced every day.

Jim Rohn

The Spark Lifecycle ⚡

Data Lake

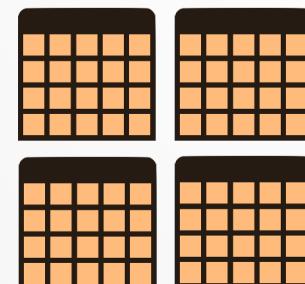
Repository of Raw Data
Without Schema Enforcement



Raw Ingestion

Apache Spark

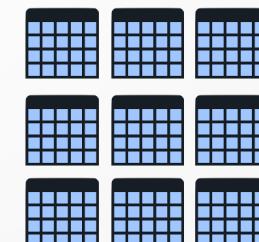
Distributed Cluster-Computing Framework
Optimized for Memory Computation



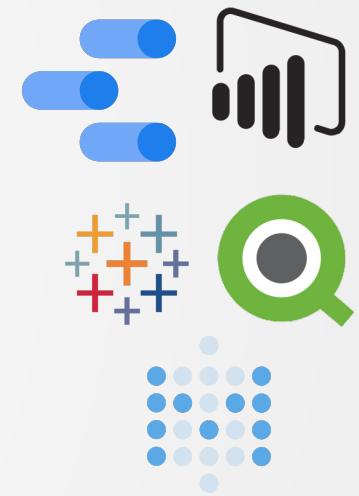
Transformations

Data Warehouse

Analytics Platform for Enterprises
Scalability – Horizontally & Vertically



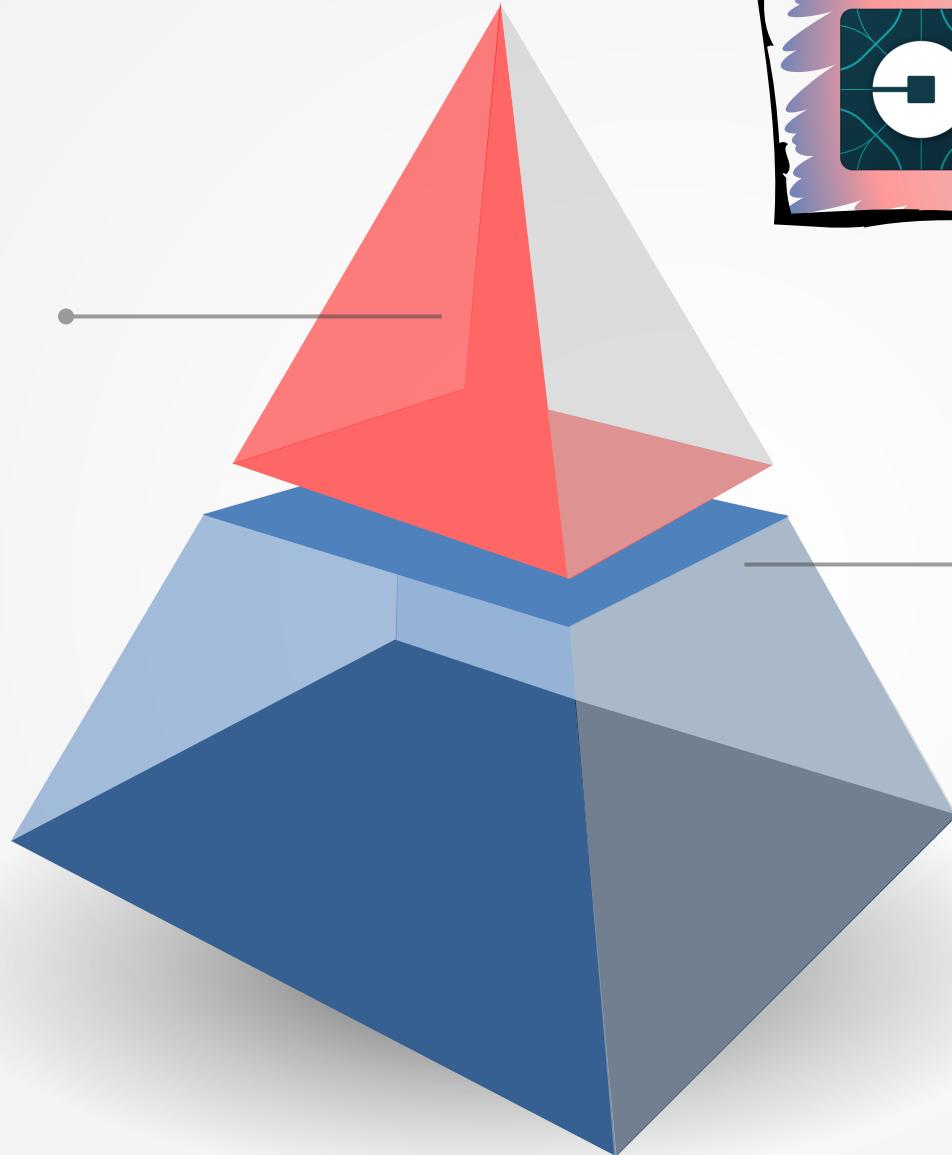
Business-Level



Data Lake vs. Data Lakehouses

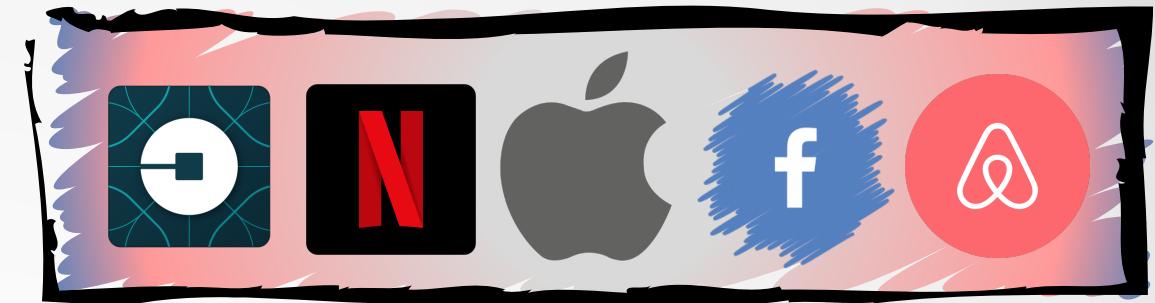
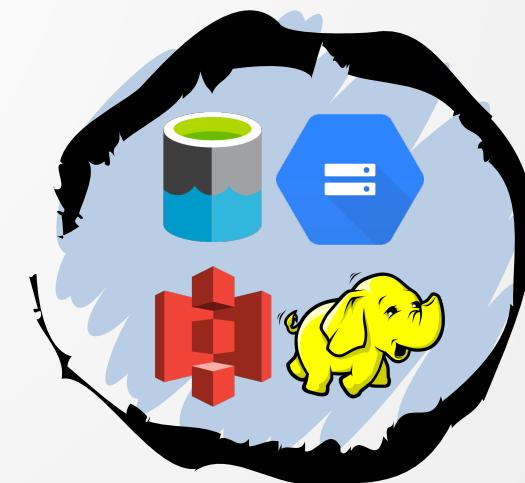
Data Lakehouses

Metadata Layers for Data Lakes
New Query Engine Design
High-Performance SQL Execution
Optimized Access of Data

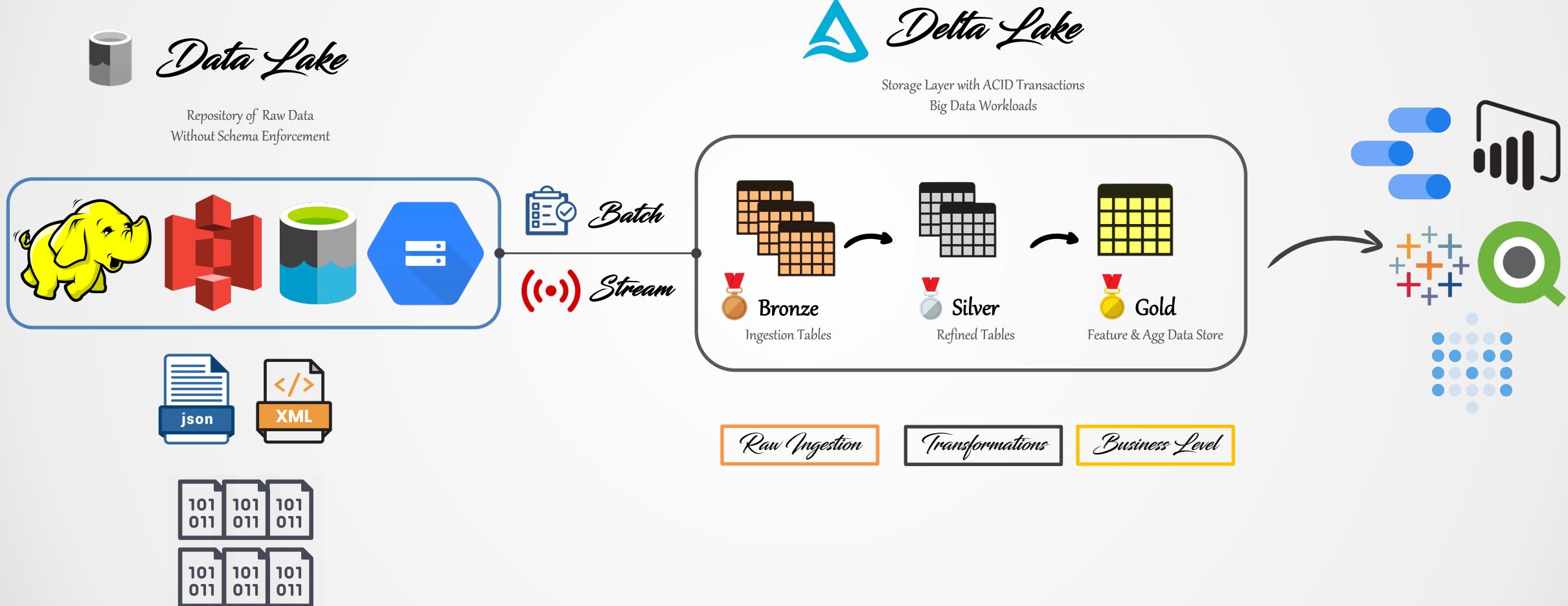


Data Lake

Repository of **Raw** Data
Unsilohed Data
Without Schema Enforcement
Data Swamp & Data Quality Issues



The Delta Architecture





What you focus on grows.

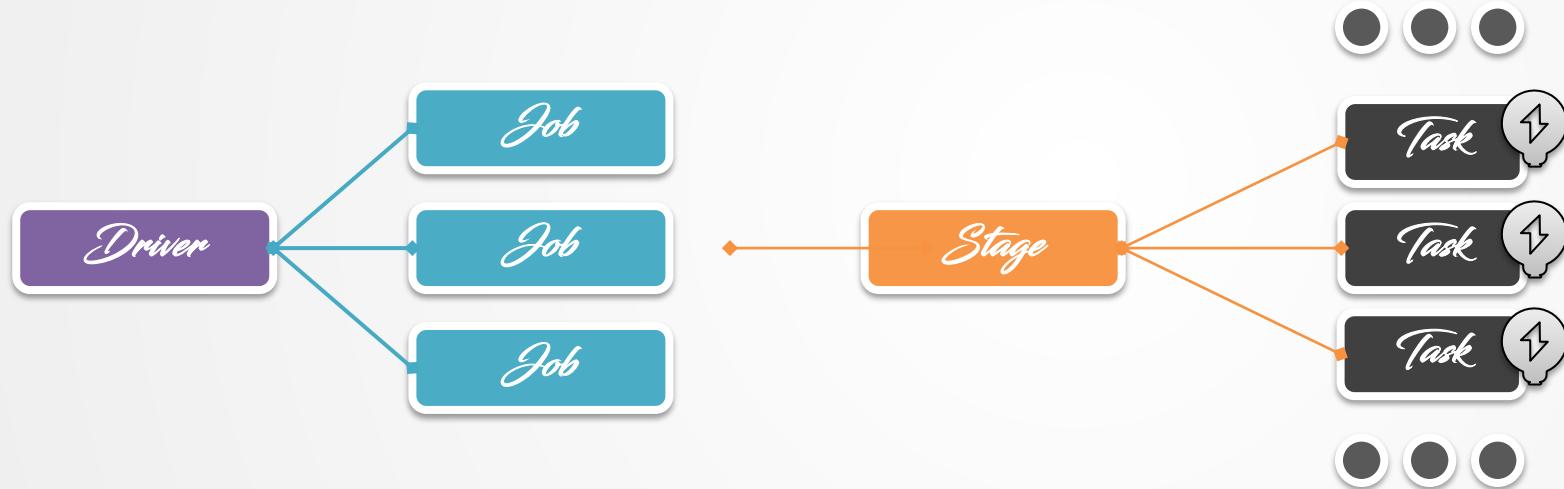
Esther Hicks

Apache Spark Query Plans Distilled ~ Query Execution Model



Query Execution Plan

Entry Point for Understanding Execution. Important for Debugging or Investigating Heavy Workloads. Understanding Query Plan is First Step Towards Optimizing Apache Spark Application Code



Job

Sequence of Stages, Triggered By an Action Such as Count(), ForeachRDD(), Collect(), Read(), Write(). Transforms Each Jobs in DAGs

Stage

Sequence of Tasks in Parallel, Based on Computation Boundaries, Based of Number of Partitions of a Dataset

Task

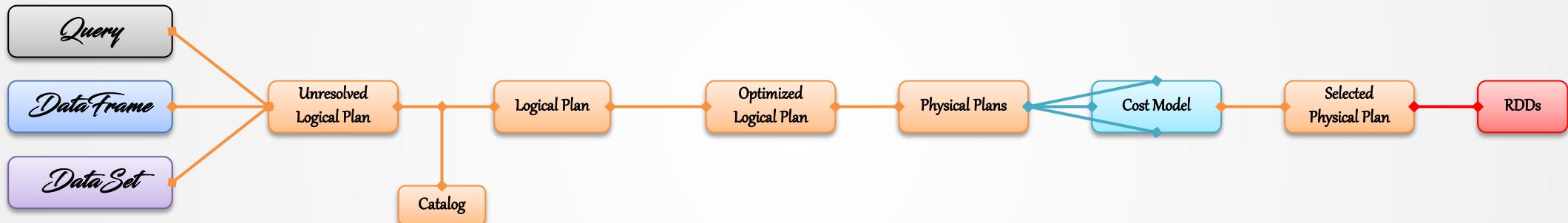
Single Operation Applied to a Single Partition. Task Executed as a Single Thread in an Executor. (Unit of Execution)

Apache Spark Query Plans Distilled ~ The Catalyst Optimizer



Logical & Physical Plans

Computational Query & Converts into an Execution Plan. Stages ~ Analysis, Logical Optimization, Physical Planning & Code Generation. Catalyst Optimizer Provides Rule-Based and Cost-Based Optimizations



Analysis

Optimizer Takes Unresolved Plan & Cross Checks with Catalog, Verify If Plan is Correct

During Resolution, Tries to Identify Data Type, Existence, and Location of Columns. Analyzer Validates Operations

If Query Resolves Successfully, Analyzed Query Plan, Additional Information Included

Logical Optimization

Once Query is Analyzed, Catalyst Optimizes Query using Rule-Based Optimization ~ Derive an Optimized Logical Plan

Physical Planning

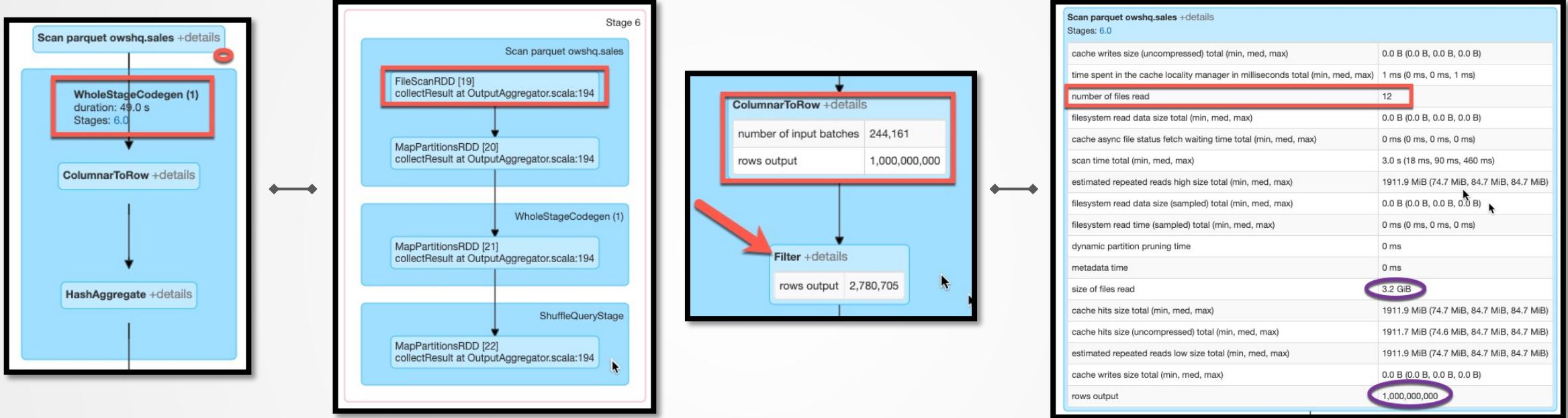
Optimizer uses Optimized Logical Plan & Generates Physical Plans. Apache Spark Decides Which Algorithm Must Be Used for Every Operator ~ SortMergeJoin & BroadcastHashJoin

Best Plan is Selected using Cost-Based Model ~ Model Costs for Engine

Code Generation

Once Best Physical Plan is Chosen, Apache Spark uses Tungsten Backend to Generate Java ByteCode to Run on Each Machine ~ Executor

Apache Spark Query Plans Distilled ~ Query Plan Operators



Collapse Code Gen Stages

Operators Grouped. During Physical Planning, Catalyst Optimizer Follows a Rule, `CollapseCodeGenStages` and Groups Operators ~ Support Code Generation Together ~ Speed Up Execution Process

Scan Parquet

Read Operations on Source Files ~ Apache Parquet & Delta. Objective ~ Pull Data from Source, Return Only Requested and Selected for **Column Pruning**, **Filter Rows using Pushed & Partition Filters**

Additional Info

Additional Information Regarding Reading from Storage System, **Number of Files Read and Size of Files. Details Used** ~ Understanding About Source Data

Apache Spark Query Plans Distilled ~ Query Plan Operators



Exchange +details	
shuffle records written	1,000,000,000
shuffle write time total (min, med, max)	30.6 s (250 ms, 956 ms, 3.3 s)
records read	1,000,000,000
local bytes read total (min, med, max)	8.1 GiB (88.6 MiB, 330.7 MiB, 330.7 MiB)
fetch wait time total (min, med, max)	561 ms (0 ms, 0 ms, 372 ms)
remote bytes read total (min, med, max)	177.3 MiB (88.6 MiB, 88.7 MiB, 88.7 MiB)
local blocks read	26
remote blocks read	8
data size total (min, med, max)	29.8 GiB (319.1 MiB, 1191.7 MiB, 1191.8 MiB)
remote bytes read to disk	0.0 B
shuffle bytes written total (min, med, max)	8.3 GiB (88.6 MiB, 330.7 MiB, 330.7 MiB)

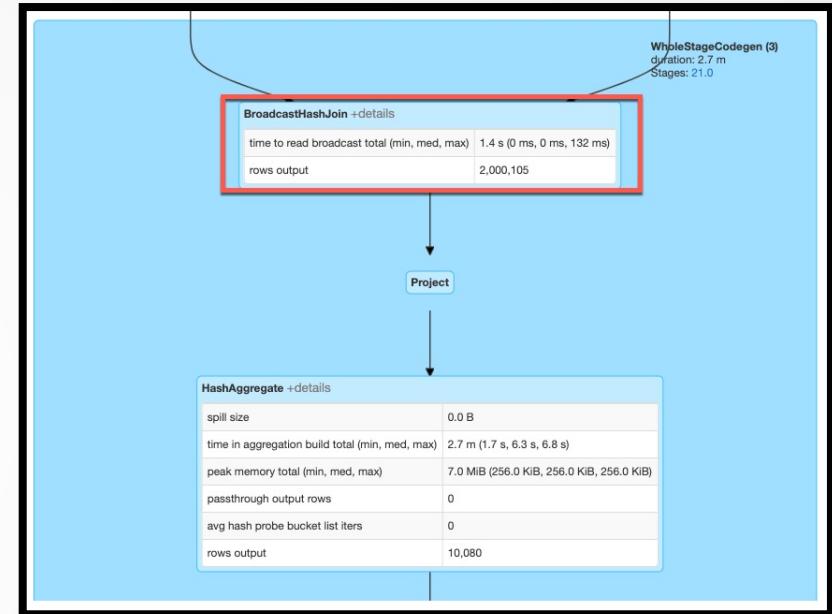


Exchange +details	
shuffle records written	300,151
shuffle write time total (min, med, max)	465 ms (26 ms, 33 ms, 60 ms)
records read	300,151
local bytes read total (min, med, max)	2.5 MiB (245.5 KiB, 253.3 KiB, 254.8 KiB)
fetch wait time total (min, med, max)	30 ms (0 ms, 0 ms, 19 ms)
remote bytes read total (min, med, max)	488.1 KiB (243.4 KiB, 244.7 KiB, 244.7 KiB)
local blocks read	10
remote blocks read	2
data size total (min, med, max)	6.9 MiB (579.3 KiB, 586.8 KiB, 591.0 KiB)
remote bytes read to disk	0.0 B
shuffle bytes written total (min, med, max)	2.9 MiB (243.4 KiB, 252.3 KiB, 254.8 KiB)

Exchange

Simply Means Shuffle, Meaning Physical Data Movement in Cluster.
One of Most Expensive Operations, Triggered

- **Joins** ~ Between DataSets, DataFrames
- **Repartition** ~ Repartition Data ~ Reduce Data Skew
- **Coalesce** ~ Move All Data ~ Single Executor ~ Output of CSV
- **Sort** ~ Output Data Sorted



Joins

Types of Joins Used By Apache Spark Engine

- **BHJ** ~ One Side is Very Small (MBs), Smaller Table is Broadcasted ~ Every Executor (Exchange) and Joined with Bigger Table using HashJoin
- **SHJ** ~ One Side is 3x Smaller, and Average of Partition Size is Small Enough for a Broadcast. During Join Partitions are Broadcasted and Joined using HashJoin
- **SortMergeJoin** ~ Most Common Join, Cannot Apply Available Options. During Join Data on Both Sides are Sorted and Joined using Merge Sort

Apache Spark 3.0 New Features



Ignore the noise,
focus on your work.

Anonymous

SaaS Data Pipeline Orchestration Options

The image features a grayscale world map with three specific regions highlighted by circles:

- Azur Data Factory (North America):** Represented by a blue factory icon.
- AWS Glue (Europe):** Represented by an orange cube icon.
- Cloud Data Fusion (Asia):** Represented by a blue hexagonal icon.

Azure Data Factory

Fully Managed, ServerLess Data Integration Solution for Ingesting, Preparing, and Transforming Data at Scale

1. Easy-to-Use = Rehost SSIS Effortlessly
2. Cost-Effective = Pay-as-You-Go
3. + 90 Built-In Connectors

AWS Glue

ServerLess Data Integration Service, Fully Managed ETL Service and Cost-Effective for Data Clean, Enrich & Move

1. Discover, Prepare, & Combine Data for Analytics, Machine Learning, & Application Development
2. Automatic Schema Discovery using Crawlers
3. Manage and Enforce Schema for Data Streams [AWS Glue Schema Registry]

Cloud Data Fusion

Fully Managed, Cloud-Native Data Integration Service at Any Scale, Open Core, Delivering Hybrid Integration using CDAP

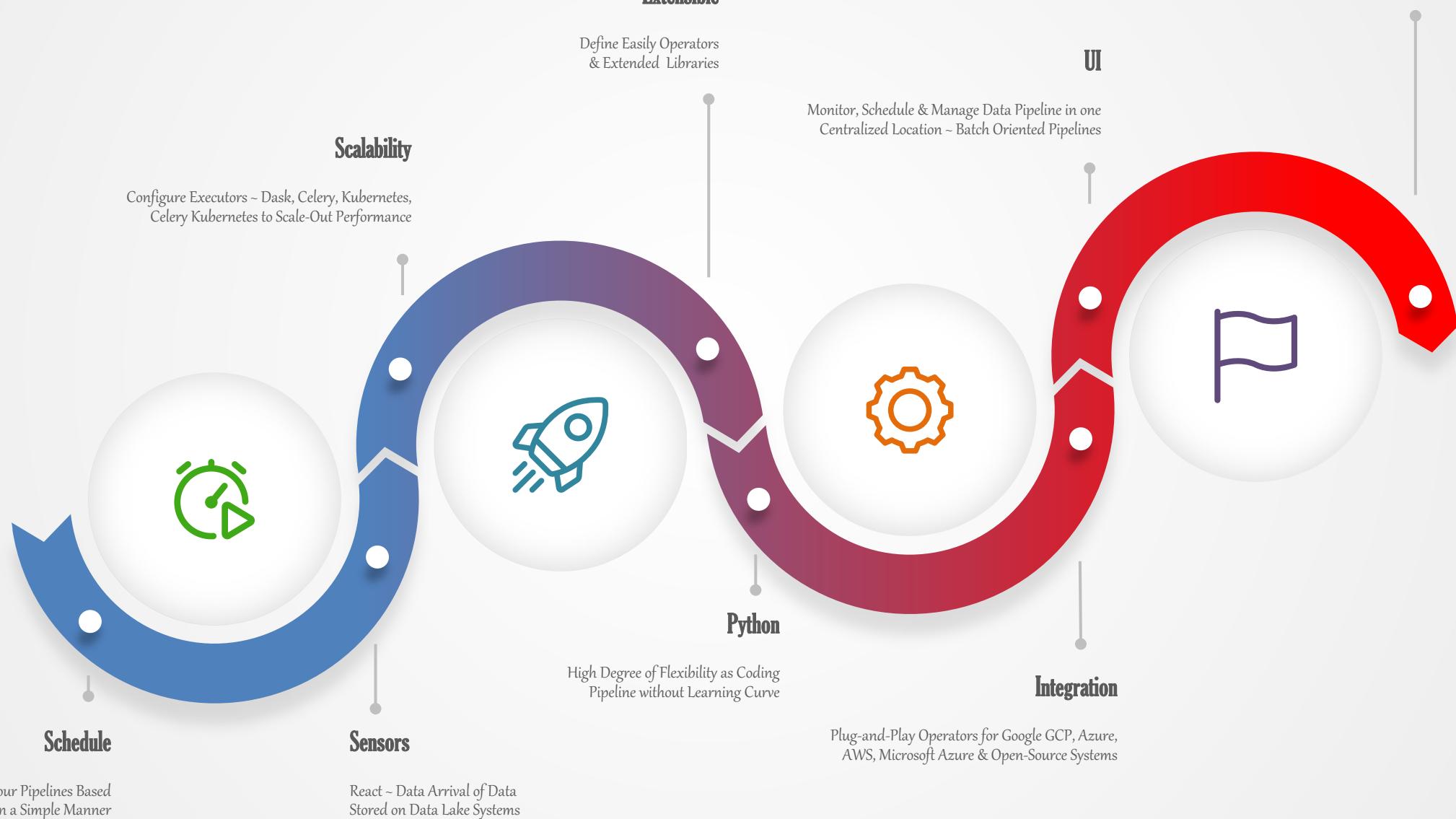
1. Visual Point-and-Click Interface Enabling Code-Free Deployment of ETL/ELT Data Pipelines
2. Broad Library of 150+ Pre-Configured Connectors & Transformations
3. Natively Integrated Best-in-Class Google Cloud Services

Apache Airflow

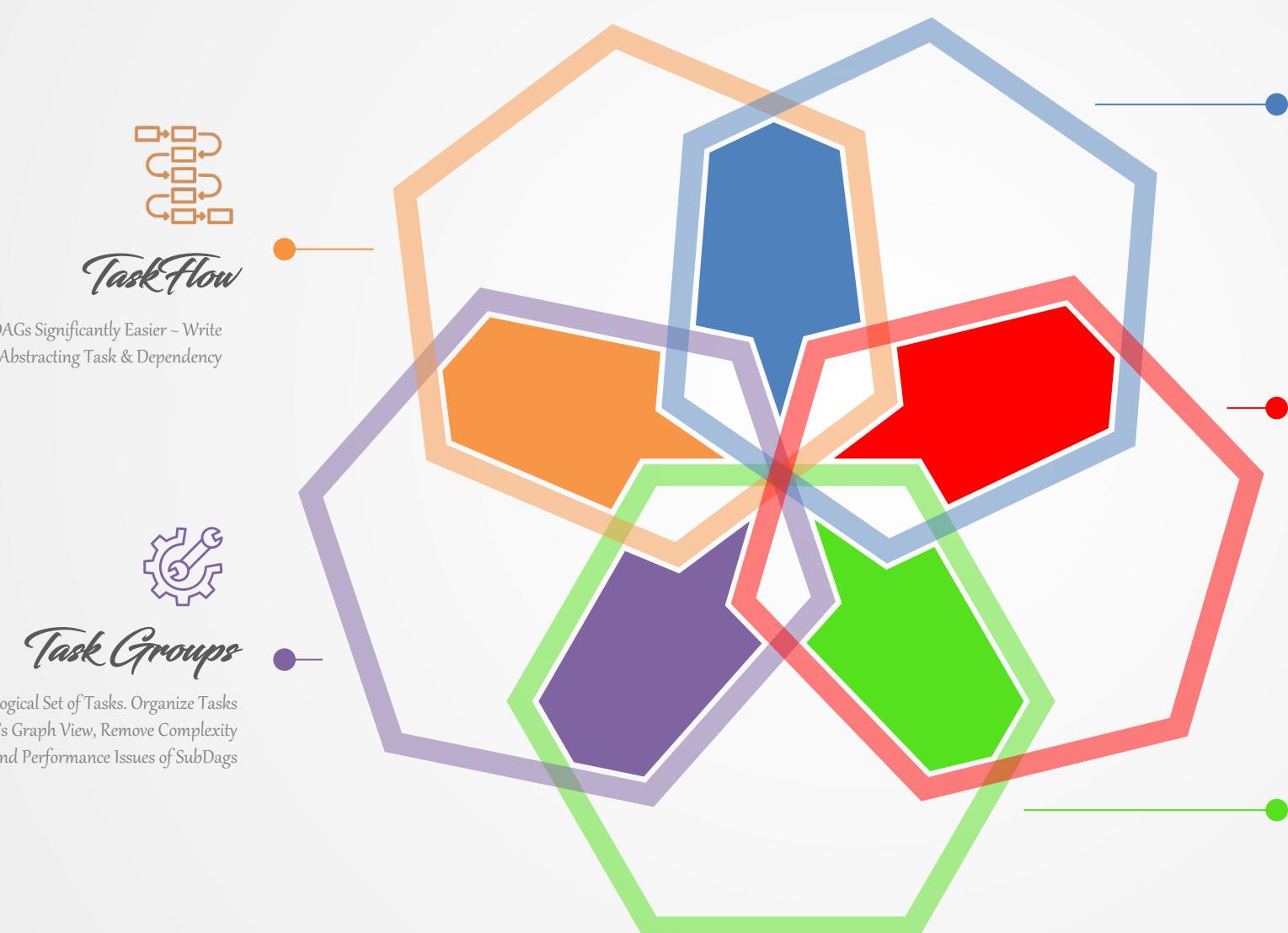


Open-Source

Apache Airflow is Open-Source and Official Kubernetes
Helm Chart Supported by Community is Available



New Features in Apache Airflow 2.0



Apache Airflow Managed Deployment Options



Amazon Managed Workflows for Apache Airflow [MWAA]

Managed Orchestration Service for Apache Airflow, Operate End-to-End Data Pipelines in Cloud at Scale with Minimum Effort & Configuration

1. Data Secured by Default Running in an Isolated and Secure Cloud Environment using VPC, Data is Automatically Encrypted using KMS
2. Connect ~ AWS or On-Premises Resources Required for Workflows Including Athena, Batch, Cloudwatch, DynamoDB, DataSync, EMR, Fargate, EKS, Firehose, Glue, Lambda, Redshift, SQS, SNS, Sagemaker & S3



Google Cloud Composer

Fully Managed Workflow Orchestration Service Built On Apache Airflow

1. Author, Schedule, and Monitor Pipelines – Span Across Hybrid and Multi-Cloud Environments
2. Frees You from Lock-In and is Easy to Use
3. Supports Hybrid and Multi-Cloud

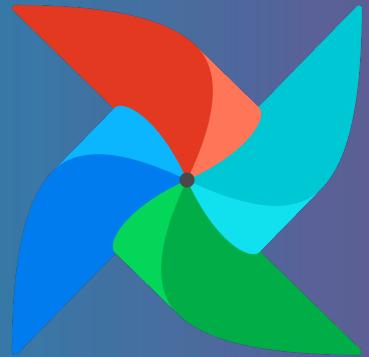


Kubernetes

Open-Source System for Automating Deployment, Scaling, and Management of Containerized Applications

1. Planet Scale
2. Runs Anywhere
3. Batch Execution
4. Self Healing
5. Designed for Extensibility
6. Multi-Cloud Approach

Apache Airflow on Kubernetes





A goal without a
plan is just a wish.

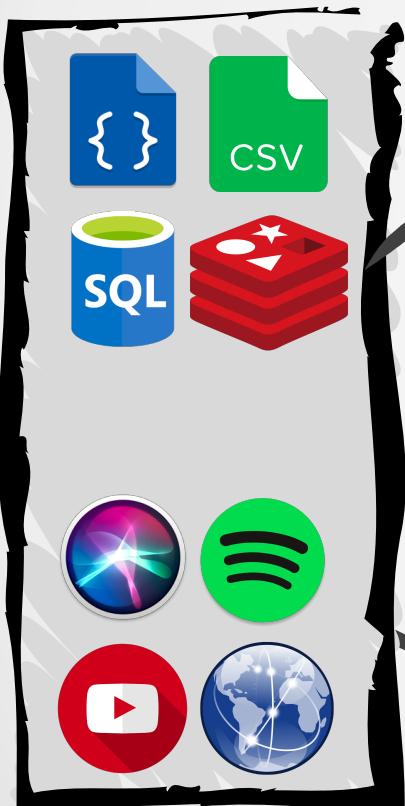
Antoine de Saint-Exupéry

Lambda Architecture – Cloud Agnostic & Simplified



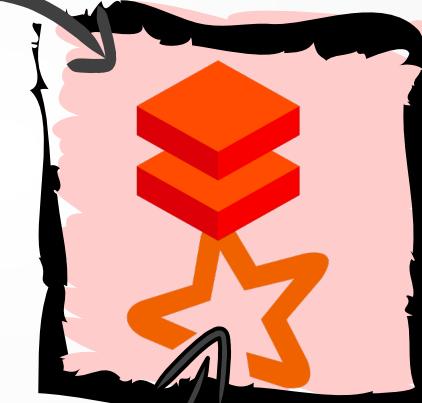
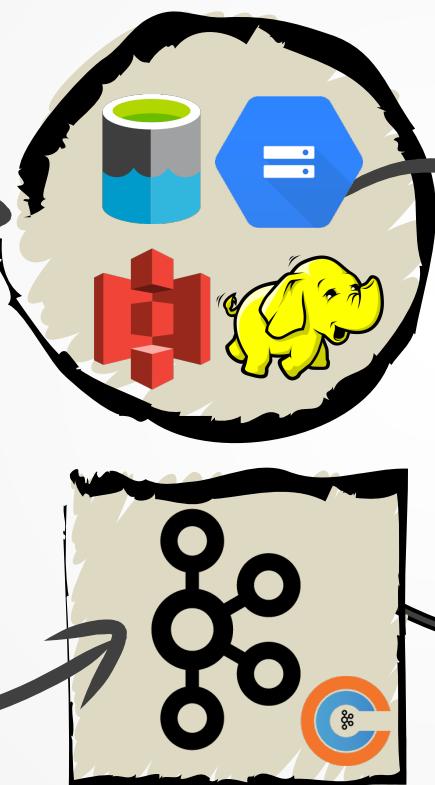
Data Source

JSON & CSV
SQL Server & Redis
Internet – Siri | Spotify | YouTube



Batch-Layer

Data Storage - Data Lake Storage Gen2 | GCS | S3 | HDFS
Batch-Processing - Apache Spark | Databricks



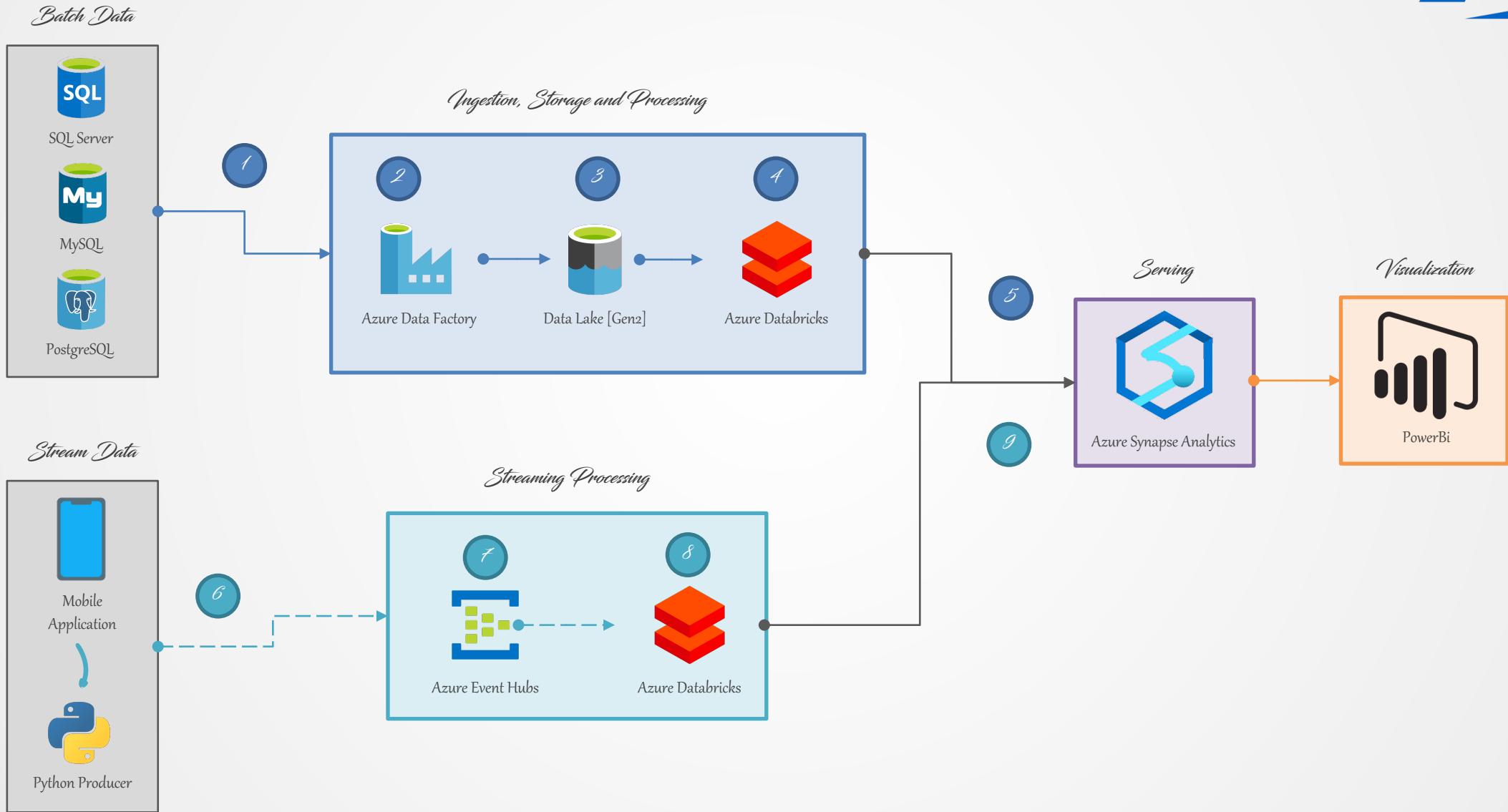
Speed-Layer

Real-Time Ingestion - Apache Kafka [Confluent]
Stream Processing - Apache Kafka [Confluent] | Apache Spark [Databricks]

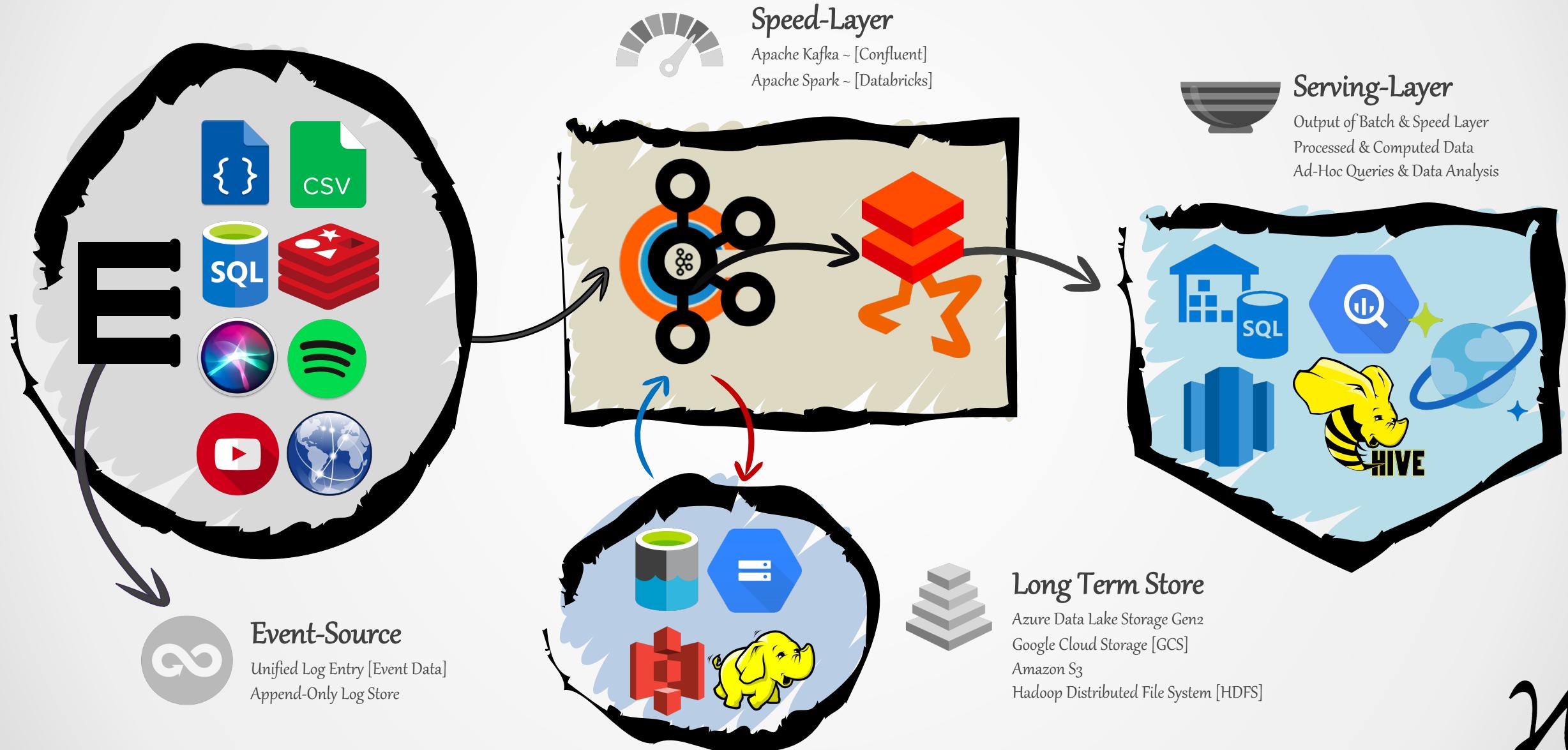
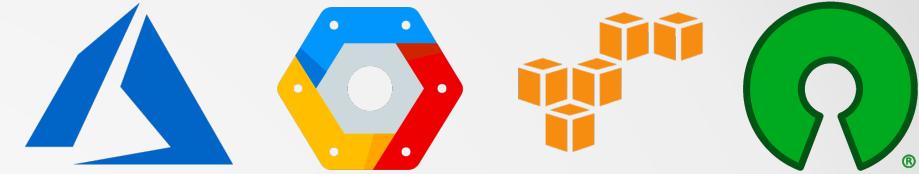




Lambda Architecture



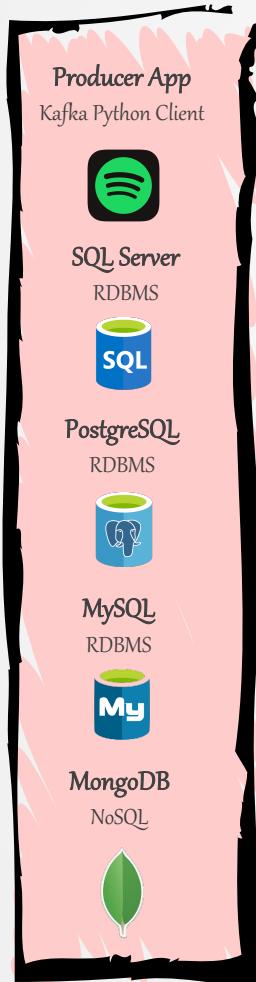
Kappa Architecture – Cloud Agnostic & Simplified





Producers

sending data in



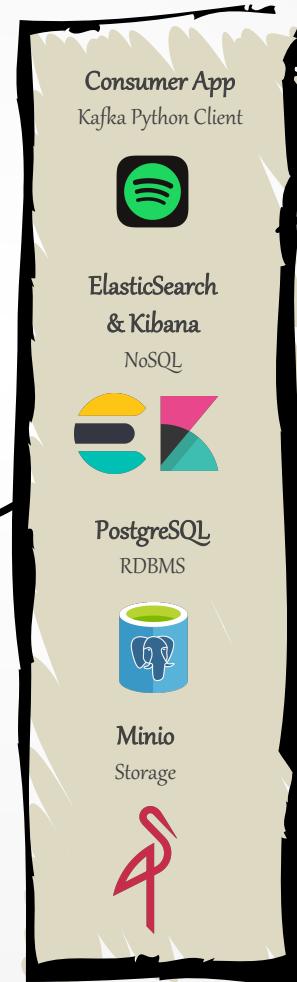
Streams

process data inside of



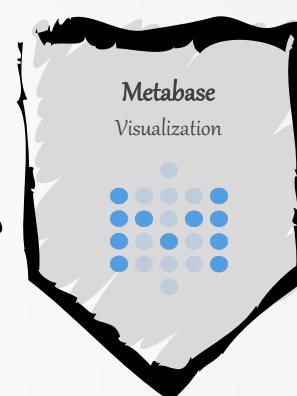
Consumers

getting data out



Serving

visualize data



[Orion] - Big Data as a Service

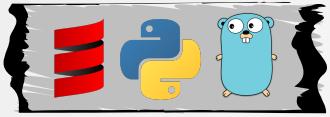
data-driven architecture for microservices & data pipelines. this process allows tighten and seamless integration between applications and data stores for analytics at scale.

Kappa Architecture



Application Ingestion Layer

Kafka Producer API



Data Store Ingestion Layer

Kafka Connect Source API



Kubernetes

Open-Source Container-Orchestration System

Enterprise Data Hub [EDH]

Apache Kafka [Strimzi]



Deep Storage

MinIO



Processing Layer

KSQLDB, Apache Spark & Faust



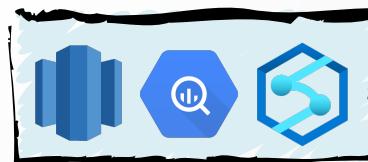
OSS Data Serving

Apache Druid & YugabyteDB



Cloud Data Serving

Redshift, BigQuery, Synapse Analytics



Integration & Exploration Layer

Trino & Zeppelin



Data Visualization

Metabase & Superset



Continuous Delivery

ArgoCD - GitOps



Orchestration Layer

Airflow



Logging

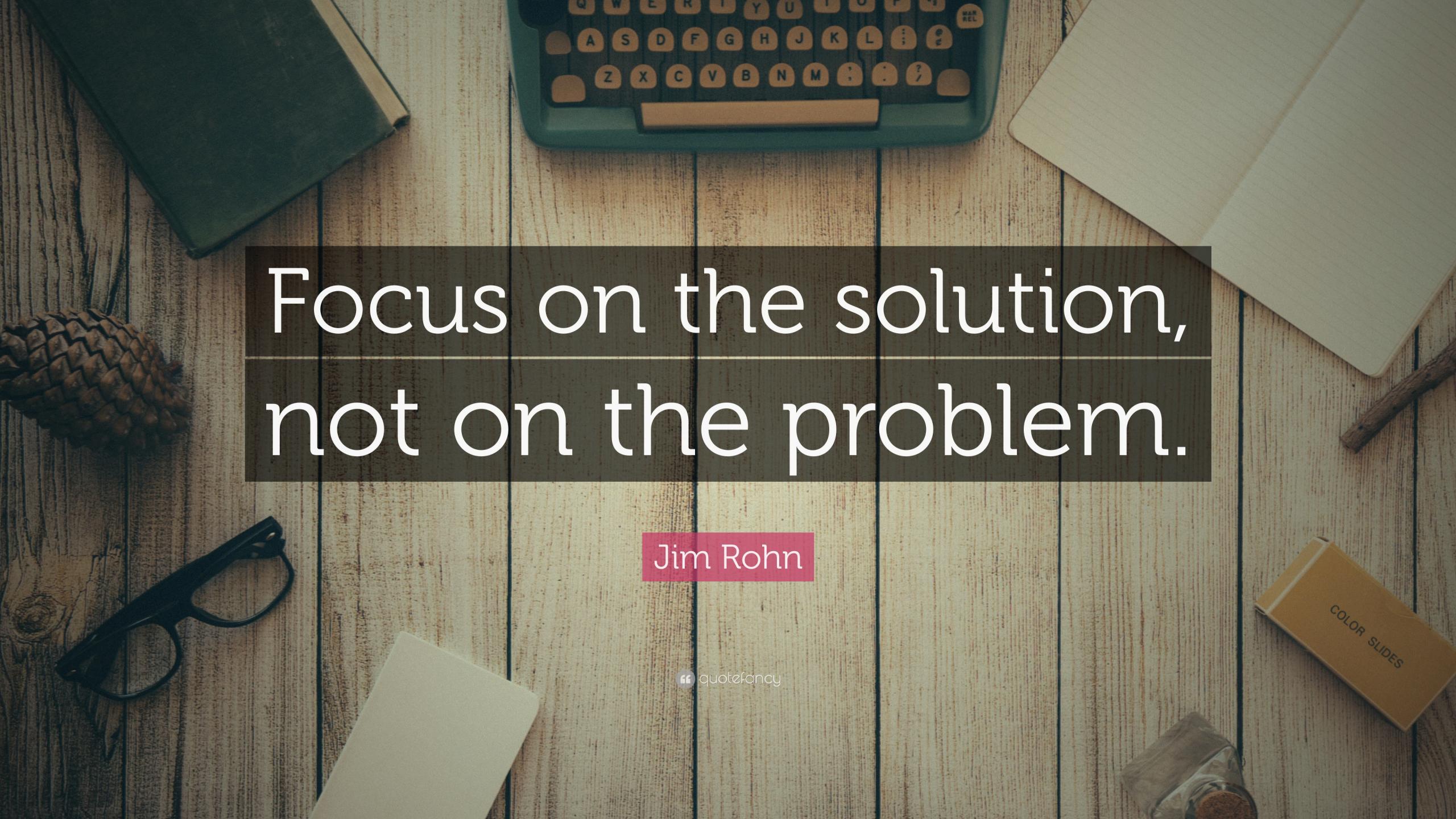
FileBeat, ElasticSearch & Kibana



Monitoring

Prometheus, Alert Manager & Grafana





Focus on the solution,
not on the problem.

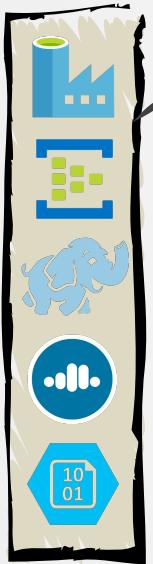
Jim Rohn

Microsoft Azure Big Data Landscape for Data Pipelines



Data Ingestion

Azure Data Factory
Azure Event Hubs
HDInsight - [Apache Kafka]
Confluent Cloud
Azure Blob Storage



Shared Resources
Shared Among Pipeline



Data Lake

Azure Data Lake Gen 2



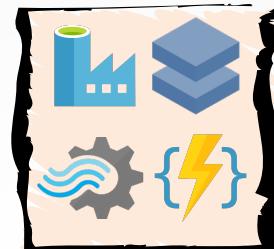
Data Discovery

Azure Purview



Data Processing

ADF ~ Mapping Data Flows
Azure Databricks
Azure Stream Analytics
Azure Functions



Data Serving

HDInsight ~ [Apache Hive]
HDInsight ~ [Interactive Query]
Azure CosmosDB
Azure Synapse Analytics
Snowflake



Data Viz

Power Bi



RDBMS

Azure SQL DB
Azure DB for MySQL
Azure DB for PostgreSQL



NoSQL

Azure CosmosDB
Azure Cache for Redis



Search

Azure Cognitive Search



Orchestration

Azure Data Factory

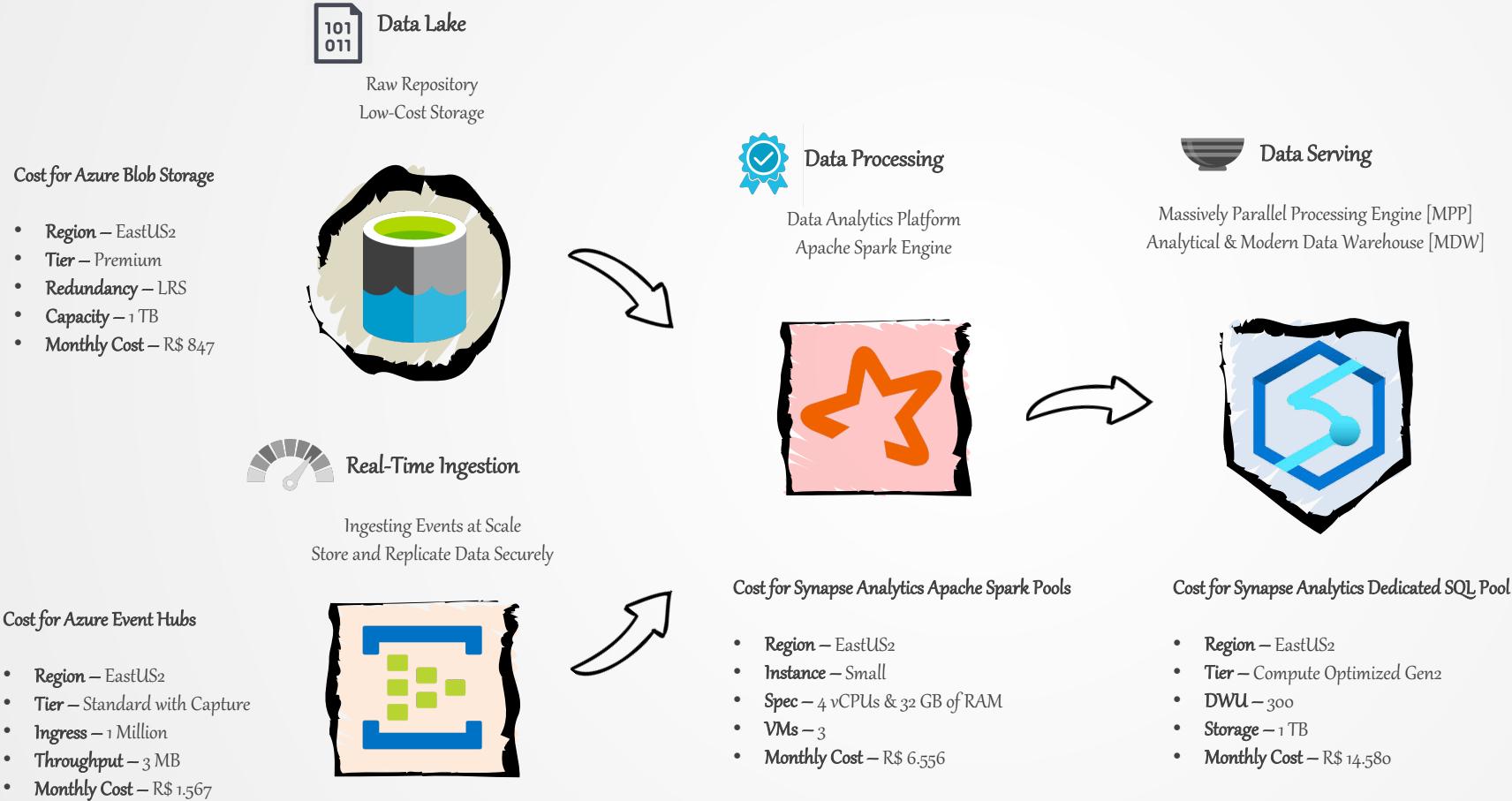


Monitoring

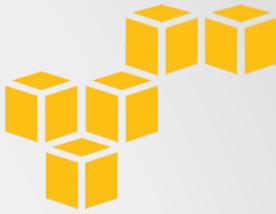
Azure Monitor



Cost of a Data Pipeline on Microsoft Azure



Amazon AWS Big Data Landscape for Data Pipelines



101 011 Data Ingestion

AWS Data Pipeline
AWS Glue
Kinesis Firehose
Kinesis Data Streams
Amazon MSK
Confluent Cloud
Amazon S3



Data Storage

Amazon S3 ~ AWS Lake Formation



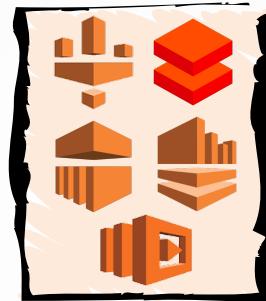
Data Exploration

Amazon Athena



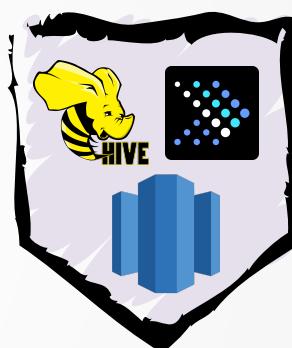
Data Processing

AWS Glue ~ DataBrew
Databricks
Amazon EMR
Kinesis Analytics
AWS Lambda



Data Serving

Amazon EMR ~ [Apache Hive]
Amazon EMR ~ [Presto]
Amazon Redshift



Data Viz

Data Studio
PowerBi
Tableau
Qlik
Metabase



Shared Resources Shared Among Pipeline



Data Discovery

AWS Glue



RDBMS

Amazon RDS



NoSQL

Amazon DynamoDB
Amazon Neptune
Amazon ElastiCache



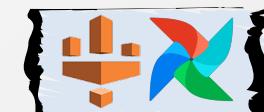
Search

Amazon CloudSearch



Data Orchestration

AWS Glue & MWAA

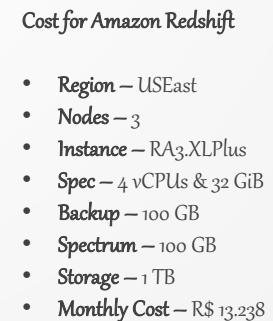
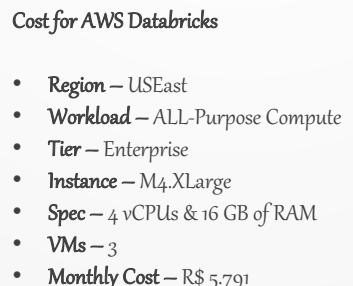
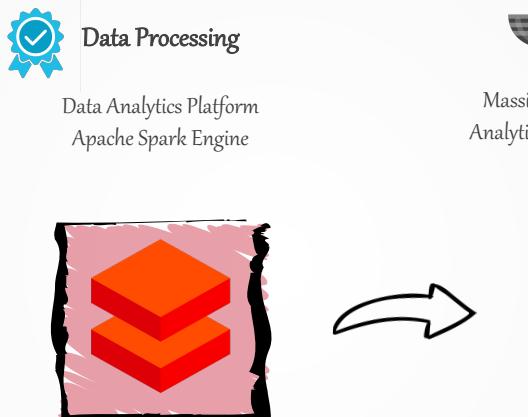
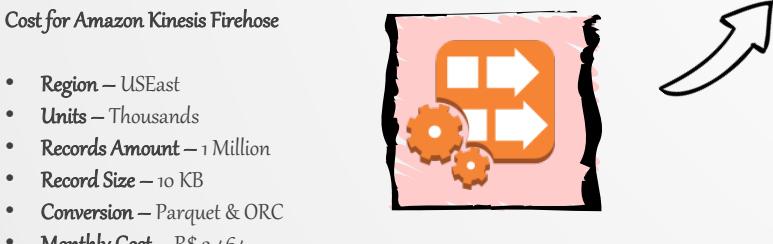
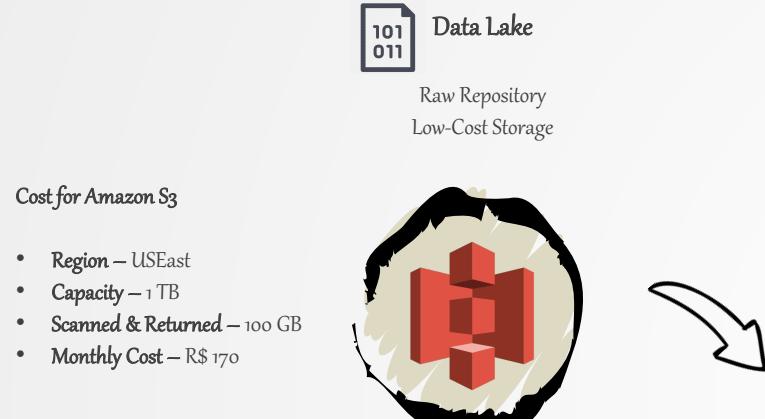


Monitoring

Amazon CloudWatch



Cost of a Data Pipeline on Amazon AWS



Total Cost for Data Pipelines on Amazon AWS

- Storage Layer = R\$ 2.634
- Data Processing Layer = R\$ 5.791
- Data Serving Layer = R\$ 13.238
- Total Monthly Cost – R\$ 21.663



Google GCP Big Data Landscape for Data Pipelines



Data Ingestion

Google Cloud Pub/Sub
Confluent Cloud
Google Cloud Storage [GCS]
Google Cloud Data Fusion



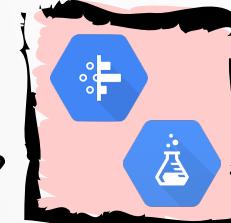
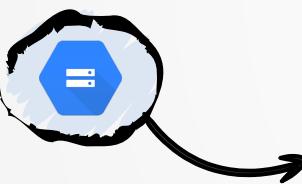
Data Storage

Google Cloud Storage [GCS]



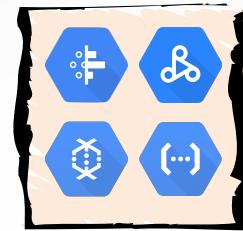
Data Exploration

Google Cloud DataPrep
Google Cloud DataLab



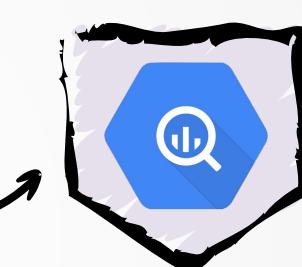
Data Processing

Google Cloud DataPrep
Google Cloud DataProc
Google Cloud DataFlow
Google Cloud Functions



Data Serving

Google BigQuery



Data Viz

Data Studio
PowerBi
Tableau
Qlik
Metabase



Shared Resources

Shared Among Pipeline



Data Discovery

Google Cloud Data Catalog



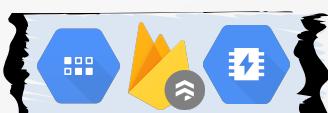
RDBMS

Google Cloud SQL
Google Cloud Spanner



NoSQL

Google Cloud BigTable
Google Cloud Firestore
Google Cloud MemoryStore



Data Orchestration

Google Cloud Composer

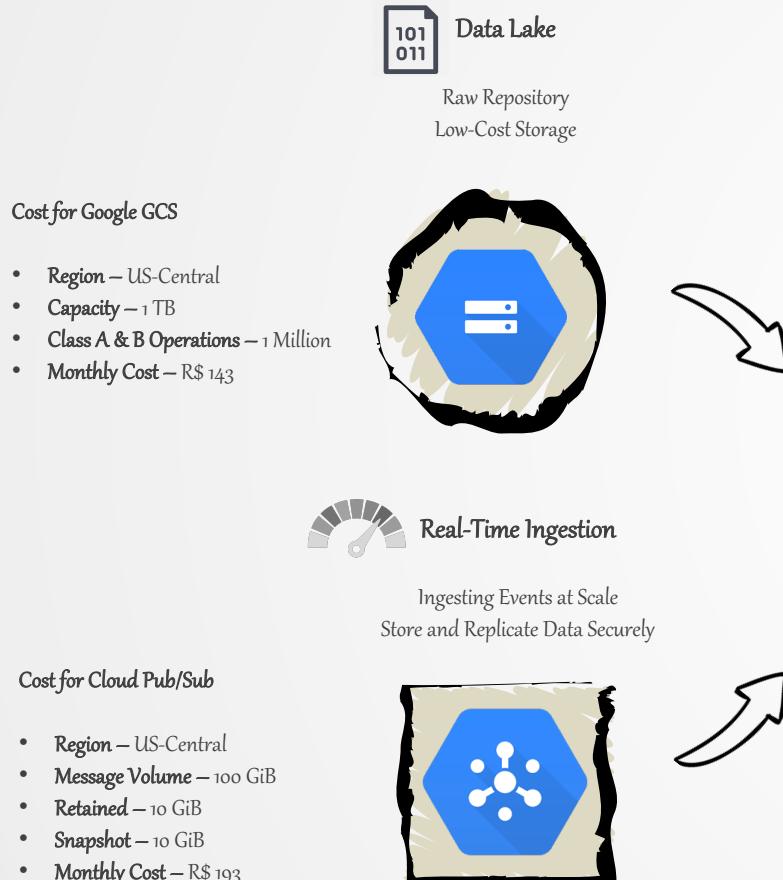


Monitoring

Google Cloud Stackdriver

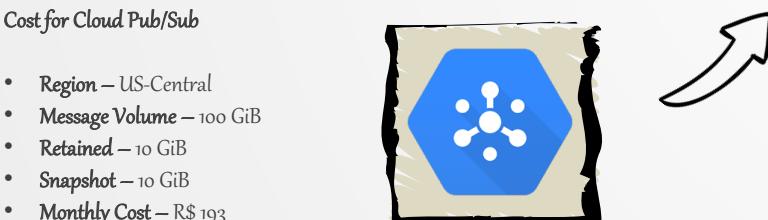


Cost of a Data Pipeline on Google GCP



Total Cost for Data Pipelines on Google GCP

- Storage Layer = R\$ 336
- Data Processing Layer = R\$ 2.156
- Data Serving Layer = R\$ 170
- Total Monthly Cost – R\$ 2.662



Cost for Google DataProc

- Region – US-Central
- VM-Class – Regular
- Instance – N1-Standard-4
- Spec – 4 vCPUs & 15 GB of RAM
- VMs – 3
- Storage – 500 GB SSD
- Monthly Cost – R\$ 2.156

Cost for Google BigQuery

- Region – US-Central
- Storage – 1 TB
- Streaming Inserts – 100 GB
- Queries – 2 TB
- Monthly Cost – R\$ 170

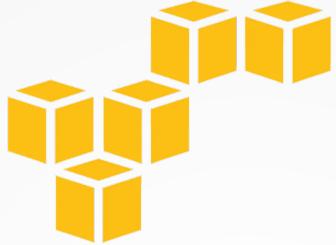


Cost for a Data Pipeline on Cloud Computing Vendors



Total Cost for Data Pipelines on Microsoft Azure

- Storage Layer = R\$ 2.414
- Data Processing Layer = R\$ 6.556
- Data Serving Layer = R\$ 14.580
- Total Monthly Cost – R\$ 23.550



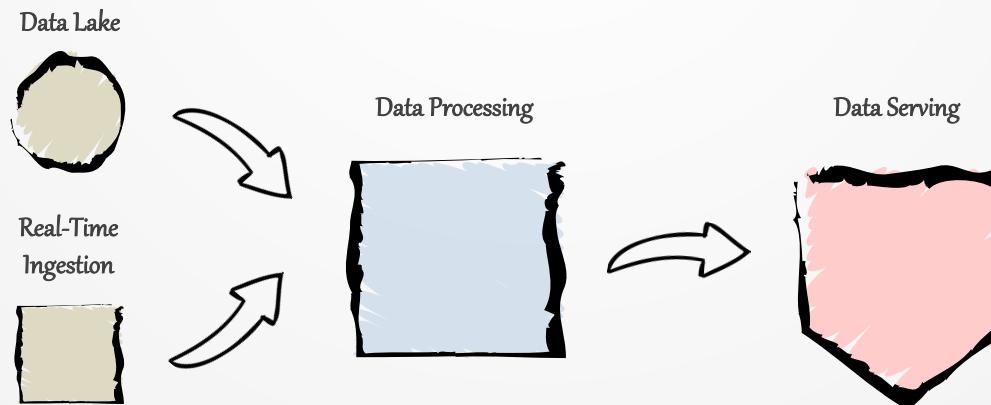
Total Cost for Data Pipelines on Amazon AWS

- Storage Layer = R\$ 2.634
- Data Processing Layer = R\$ 5.791
- Data Serving Layer = R\$ 13.238
- Total Monthly Cost – R\$ 21.663



Total Cost for Data Pipelines on Google GCP

- Storage Layer = R\$ 336
- Data Processing Layer = R\$ 2.156
- Data Serving Layer = R\$ 170
- Total Monthly Cost – R\$ 2.662

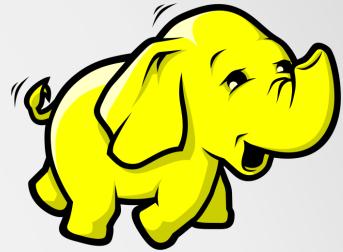




A goal without a
plan is just a wish.

Antoine de Saint-Exupéry

OSS Big Data Products on [Spotlight]



Apache Kafka

80% of ALL Fortune 100 Companies Trust. Ingest and Process Data Effortlessly



Apache Pulsar

Messaging & Streaming Platform. Pulsar Functions, Persistent Storage, Multi-Tenancy with Low-Latency



Apache Spark

PySpark, Spark SQL, Java, Scala, R, .NET. Most Used Big Data Product



Apache Airflow

Programmatically Author, Schedule & Monitor Workflows using Python. Newest 2.0 Version Out



Apache Pinot

Real-Time Distributed OLAP Data Store, Designed for Low-Latency Queries at Scale



Trino

Data Processing Engine Unleashing SQL at Scale & Providing Data Virtualization Process Layer



Dremio

Next-Generation Data Lake Engine for Interactive Query in a Blazing Fast Speed



YugaByteDB

Cloud-Native Database Platform using Different APIs – Redis, Postgres & Cassandra

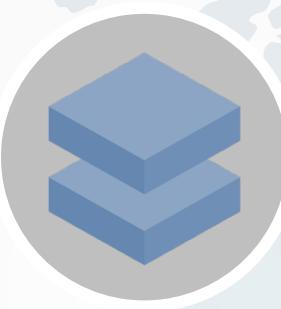
Azure Big Data Products on [Spotlight]



Azure Purview

Unified Data Governance with Data Discovery, Sensitive Data Classification & End-to-End Data Lineage

1. Data Discovery, Classification and Mapping
2. Data Catalog: Searching & Web-Based Experience
3. Data Governance: Enabling Key Insights and Understanding of Data Quality Rules



Azure Databricks

Fast, Easy & Collaborative Apache Spark Based Analytics Service Providing Fast Deployment Process

1. Databricks Runtime ~ Optimized for Cloud Storage
2. Managed Delta Lake
3. Integrated Workspace – GitHub
4. Production Jobs & Workflows
5. Enterprise Security
6. Integrations using ODBC & JDBC
7. SQL Analytics – Redash + Delta Lake Engine

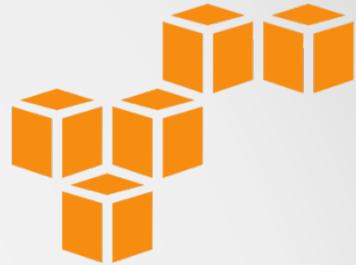


Azure Synapse Analytics

MDW with Limitless Analytics Service with Unmatched Time to Insight – PaaS & SaaS Approaches

1. Serverless & Dedicated Options
2. Data Lake Exploration
3. Code-Free ETL & ELT
4. Deeply Integration with Apache Spark & SQL Engines
5. Languages – T-SQL, Python, Scala, Spark SQL and .NET
6. Cloud-Native HTAP with Azure Synapse Link ~ CosmosDB
7. AI & BI

AWS Big Data Products on [Spotlight]



Managed Streaming for Apache Kafka [MSK]

Fully Managed Service to Build and Run Applications ~ Apache Kafka to Process Streaming Data Effortlessly

1. Amazon MSK Runs and Manages Apache Kafka, Maintain Open-Source Compatibility, MirrorMaker, Apache Flink, and Prometheus
2. VPC Network Isolation, AWS IAM for Control-Plane API Authorization, Encryption at Rest, TLS Encryption & In-Transit



Amazon Glue & DataBrew

Serverless Data Integration Service for ETL, ELT, Catalog, Lineage & Transformations for Cleaning and Enriching Data

1. Discover, Prepare, & Combine Data for Analytics, Machine Learning, and Application Development
2. DataBrew is New Visual Data Preparation Tool ~ Clean and Normalize Data for Analytics and Machine Learning

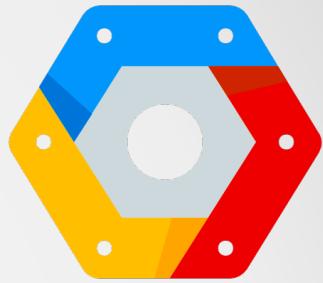


Amazon Managed Workflows for Apache Airflow [MWAA]

Managed Orchestration Service for Apache Airflow, Operate End-to-End Data Pipelines in Cloud at Scale with Minimum Effort & Configuration

1. Data Secured by Default Running in an Isolated and Secure Cloud Environment using VPC, Data is Automatically Encrypted using KMS
2. Connect ~ AWS or On-Premises Resources Required for Workflows Including Athena, Batch, Cloudwatch, DynamoDB, DataSync, EMR, Fargate, EKS, Firehose, Glue, Lambda, Redshift, SQS, SNS, Sagemaker & S3

GCP Big Data Products on [Spotlight]



Cloud Data Fusion

Fully Managed, Cloud-Native Data Integration at Any Scale using
Ephemeral DataProc Cluster Underneath

1. Code-Free ETL & ELT Deployment of Data Pipelines
2. Library of 150+ Configured Connectors & Transformations
3. Built with OSS Core CDAP for Pipeline Portability



Cloud Dataflow

Unified Stream and Batch Data Processing Serverless, Fast, and Cost-Effective
using Beam Framework

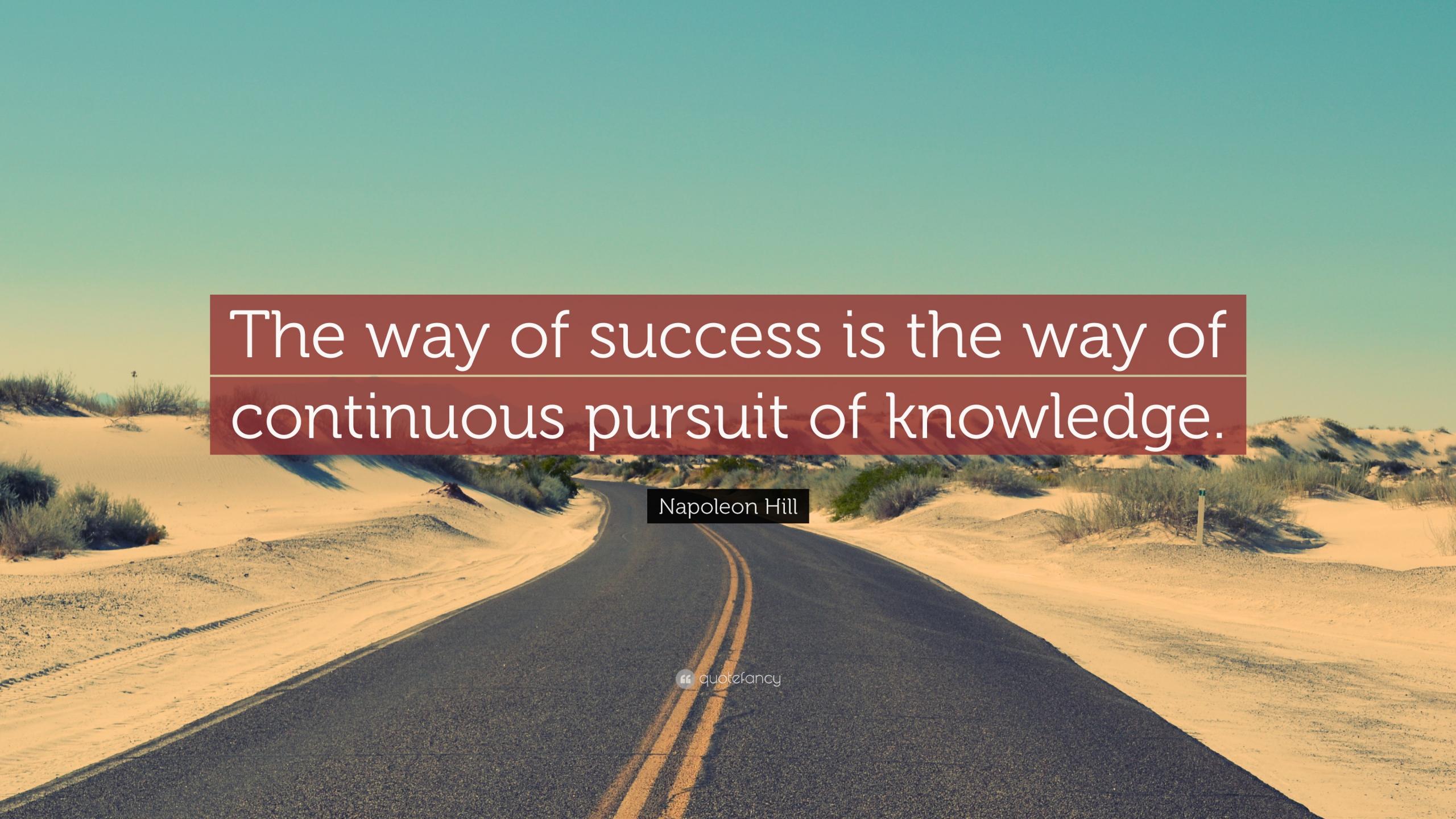
1. Automated Provisioning and Management of Processing Resources
2. Horizontal Autoscaling of Worker Resources
3. OSS Community-Driven Innovation with Apache Beam SDK
4. Reliable and Consistent Exactly-Once Processing



BigQuery

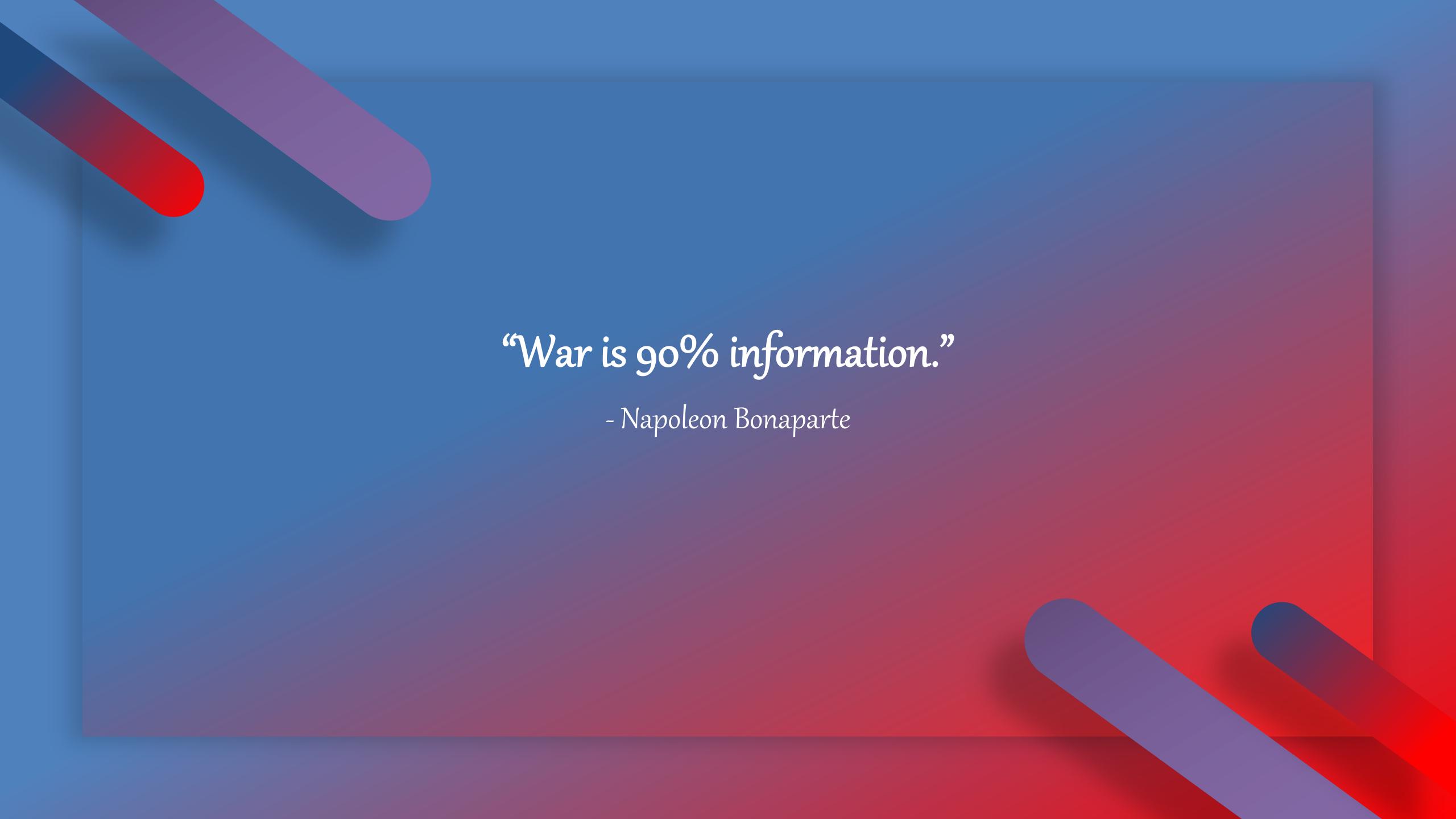
Serverless, Highly Scalable, and Cost-Effective Multi-Cloud Data
Warehouse Designed for Business Agility

1. Analyze Petabytes of Data Using ANSI SQL at Blazing-Fast Speeds,
with Zero Operational Overhead
2. Democratize Insights with a Trusted and Secure Platform Scales
3. Gain Insights from Data Across Clouds with a Flexible, Multi-Cloud
Analytics Solution ~ Omni

A photograph of a paved road curving through a desert environment. The road is dark asphalt with yellow double lines, set against light-colored sand dunes and sparse green vegetation. The sky is clear and blue.

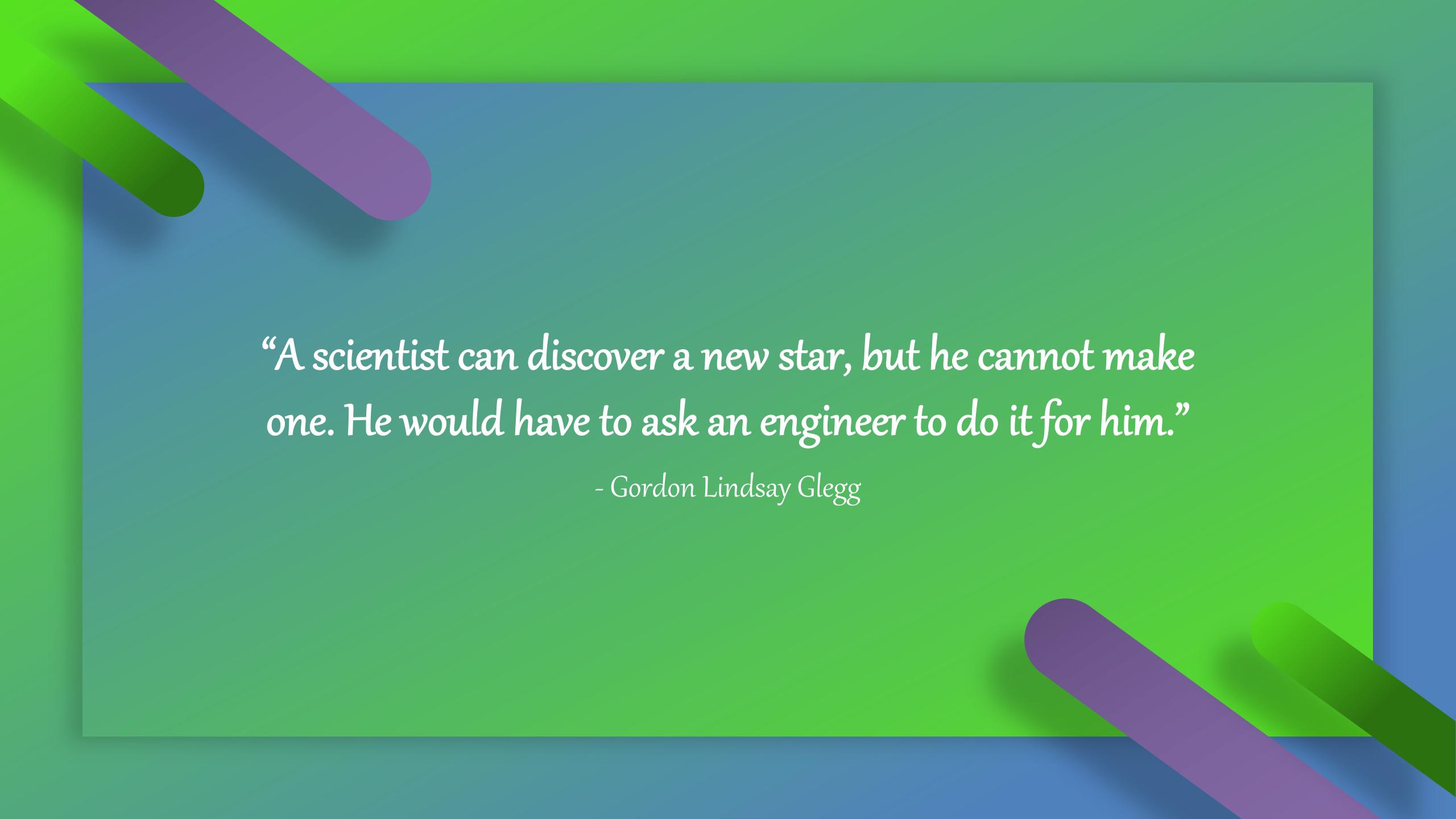
The way of success is the way of
continuous pursuit of knowledge.

Napoleon Hill



“War is 90% information.”

- Napoleon Bonaparte



“A scientist can discover a new star, but he cannot make one. He would have to ask an engineer to do it for him.”

- Gordon Lindsay Glegg



“In God we trust. All others must bring data.”

- W. Edwards Deming

Data Engineer Technical Skills

Data Engineer Career - Part 1



OS & Programming Language

- Linux
- SQL
- Python
- Scala



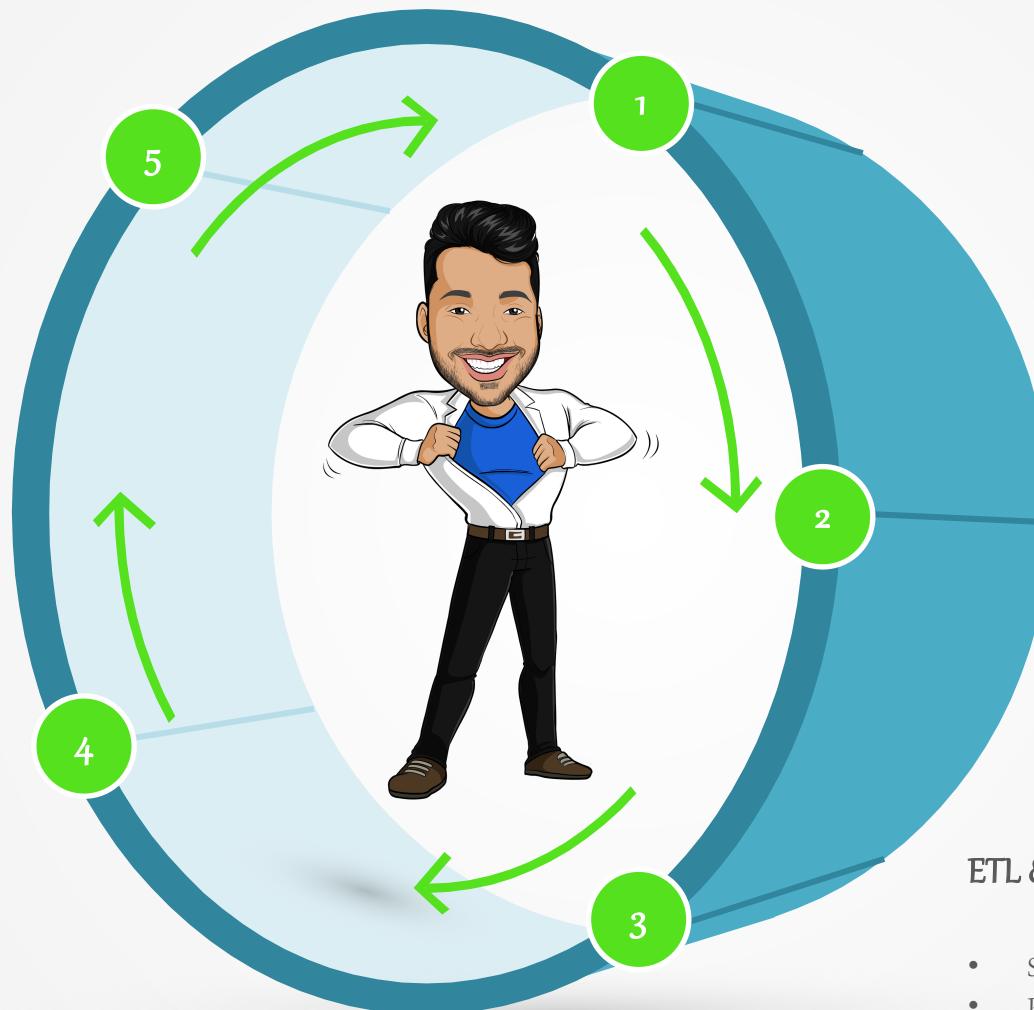
DBMS & NoSQL

- SQL Server
- Oracle
- PostgreSQL
- MySQL
- MongoDB
- Cassandra
- Redis Cache



ETL & DW

- SSIS & ODI
- PowerCenter
- Talend
- Pentaho
- Oracle Exadata
- Sybase IQ



Data Pipelines & Cloud Computing



- Lambda & Kappa
- Google GCP
- Amazon AWS
- Microsoft Azure

Distributed Systems & Big Data Frameworks



- Apache Hadoop [HDFS]
- Apache Spark
- Apache Kafka
- Apache Airflow

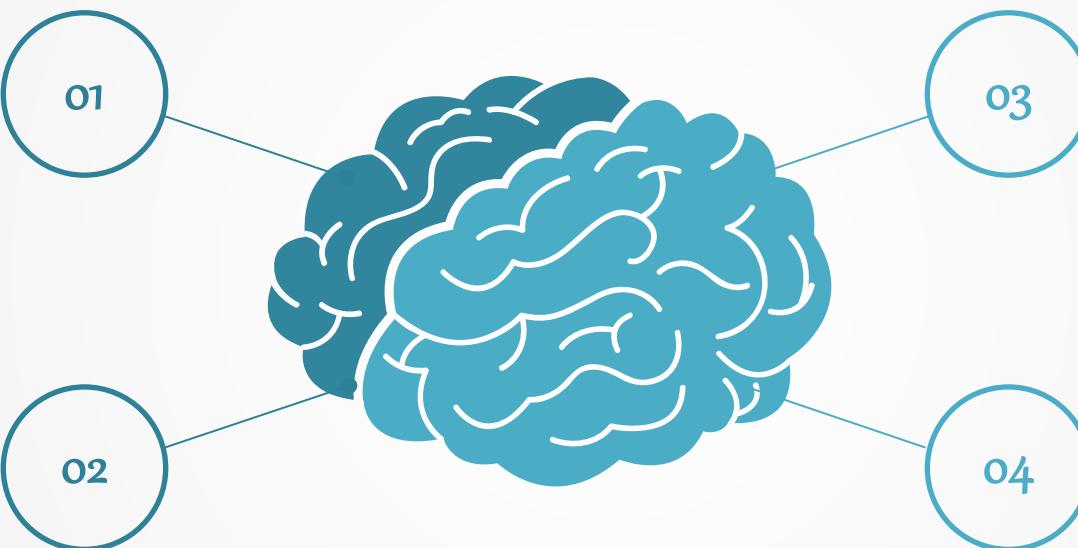
Data Engineer Business Skills

Data Engineer Career - Part 2



Creative Problem-Solving

approaching data organization challenges with a clear eye on what is important; employing the right approach/methods to make the maximum use of time and human resources.



Effective Collaboration

carefully listening to management, data scientists and data architects to establish their needs.

Intellectual Curiosity

exploring new territories and finding creative and unusual ways to solve data management problems.

Industry Knowledge

understanding the way your chosen industry functions and how data can be collected, analyzed and utilized; maintaining flexibility in the face of big data developments.

Data Engineer Certifications

Data Engineer Career - Part 3



Amazon Web Services (AWS)
Certified Big Data – Specialty

the aws certified big data – specialty certification is intended for individuals who perform complex big data analysis with at least two years of experience using aws technology.



Google Professional Data Engineer

professional data engineer enables data-driven decision making by collecting, transforming, and publishing data.



Microsoft Certified: Azure Data Engineer
Associate

azure data engineers design and implement the management, monitoring, security, and privacy of data using the full stack of azure data services to satisfy business needs.



Databricks Certified Associate Developer
for Apache Spark 3.0

validates your knowledge of the core components of the dataframes api and confirms that you have a rudimentary understanding of the spark architecture.



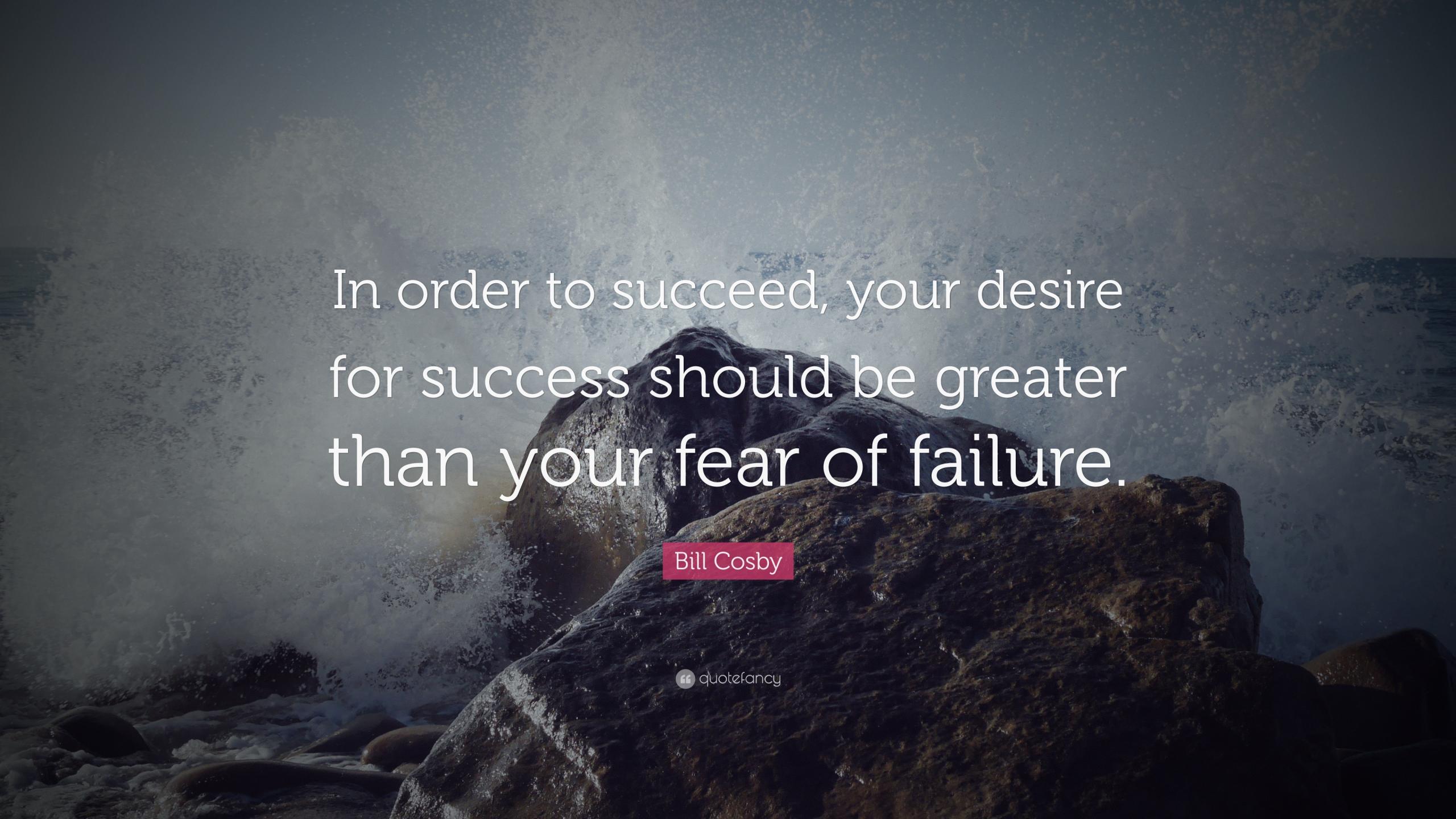
Confluent Certified Developer
for Apache Kafka (CCDAK)

this examination is based upon the most critical job activities that a confluent developer performs.

Data Engineer Study

Data Engineer Career - Part 4





In order to succeed, your desire
for success should be greater
than your fear of failure.

Bill Cosby



quotefancy

Luan Moreno M. Maciel



YouTube
luanmorenommaciel



LinkedIn
Luan Moreno Medeiros Maciel



Facebook
Luan Moreno Medeiros Maciel



Instagram
engenhariadedados



Podcast
engenhariadedadoscast



Thank You



One Way Solution



ONE WAY
SOLUTION