

Doris Summit Asia 2024

Doris Application Within Vehicle Data

Ford China – Biao Jiang, Yong Qing Zhang



databricks Lakehouse Platform

Data
Warehousing

Data
Engineering

Data
Streaming

Data Science
and ML

Unity Catalog

Fine-grained governance for data and AI

Delta Lake

Data reliability and performance

Cloud Data Lake

All structured and unstructured data



Microsoft Azure



Google Cloud

Facing Problems

High Cost

- The cost of the machine
- The cost of the DBU

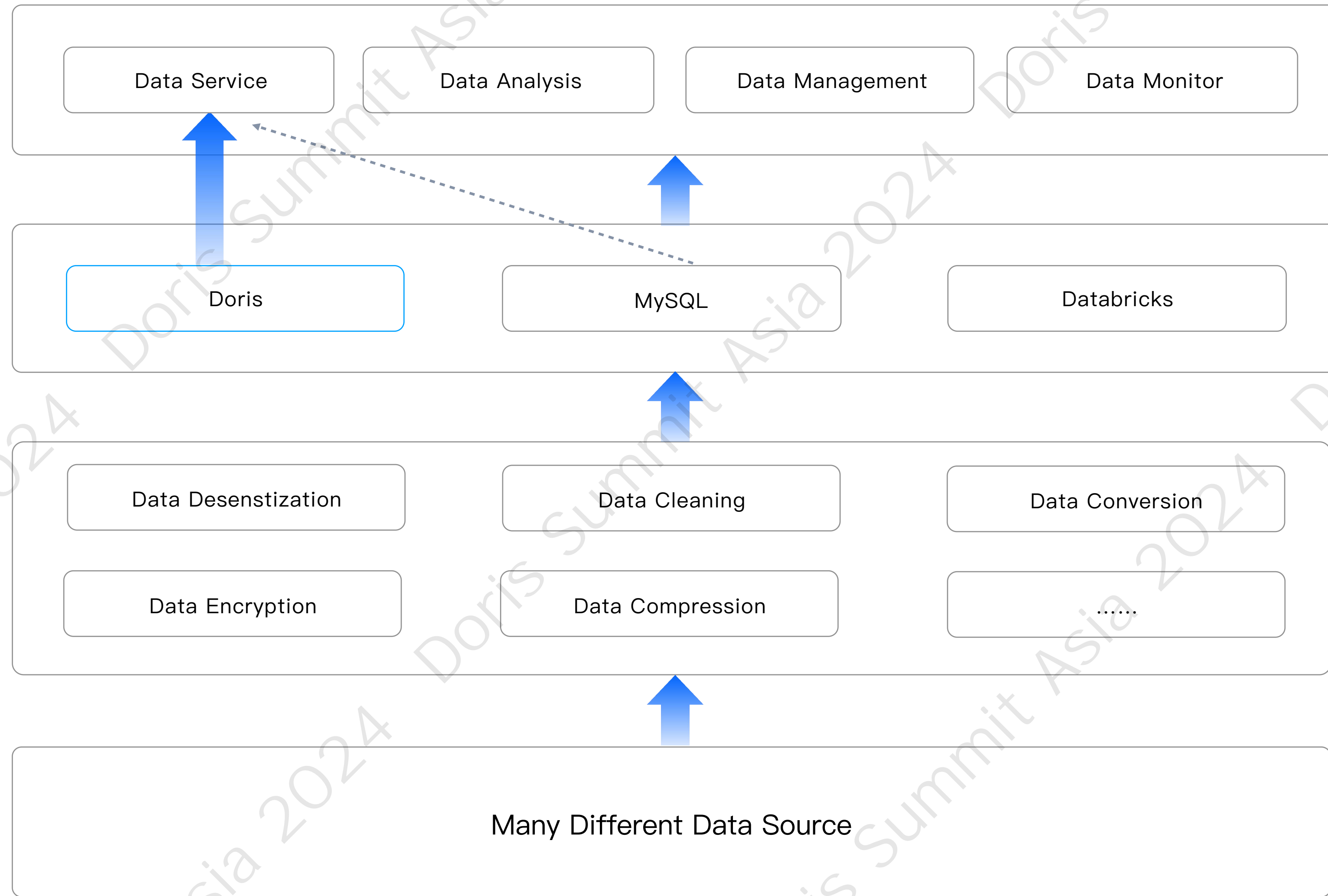
Real-Time Query Capability

- Real time query capability is very weak when providing data services
- Especially the ability to perform concurrent queries, because the query technology is based on Apache Spark

Real-Time Hot Data

- Use Apache Doris to provide the ability to query real-time hot data

Architecture





Service

Doris Follower Service

Doris BE Service

Stateful Set

Doris Follower Cluster

Doris BE Cluster

Data Storage

PersistentVolumeClaim

PersistentVolumeClaim

PersistentVolumeClaim

1

Spark Doris Connector

2

Doris Streamloader

3

Many Common Functions

4

Materialized View

Thanks for Watching!