

# 腾讯数据湖的 元数据治理实践

---

吴怡雯 高级工程师



# 目录 CONTENT

**01** 背景概述

**03** 在线目录

**02** 租户设计

**04** 离线治理

# 01

## 背景概述



# 数据仓库 VS 数据湖

Snowflake + 云厂商

**数据仓库**：面向主题的、集成的、相对稳定的、反映历史的**数据集合**，提供管理决策。(Bill Inmon，数仓之父，1990)

**数据湖**：存储各类**自然格式**数据的系统，提供数据ETL操作。(James Dixon，Pentaho CTO，2010)

优势(DLC为例)：

- **高时效**：表格式(Iceberg)、存储缓存(Alluxio)
- **低成本**：对象存储(COS)、Serverless计算(EKS)
- **易扩展**：存算分离架构、多样化计算引擎(Presto/Spark)

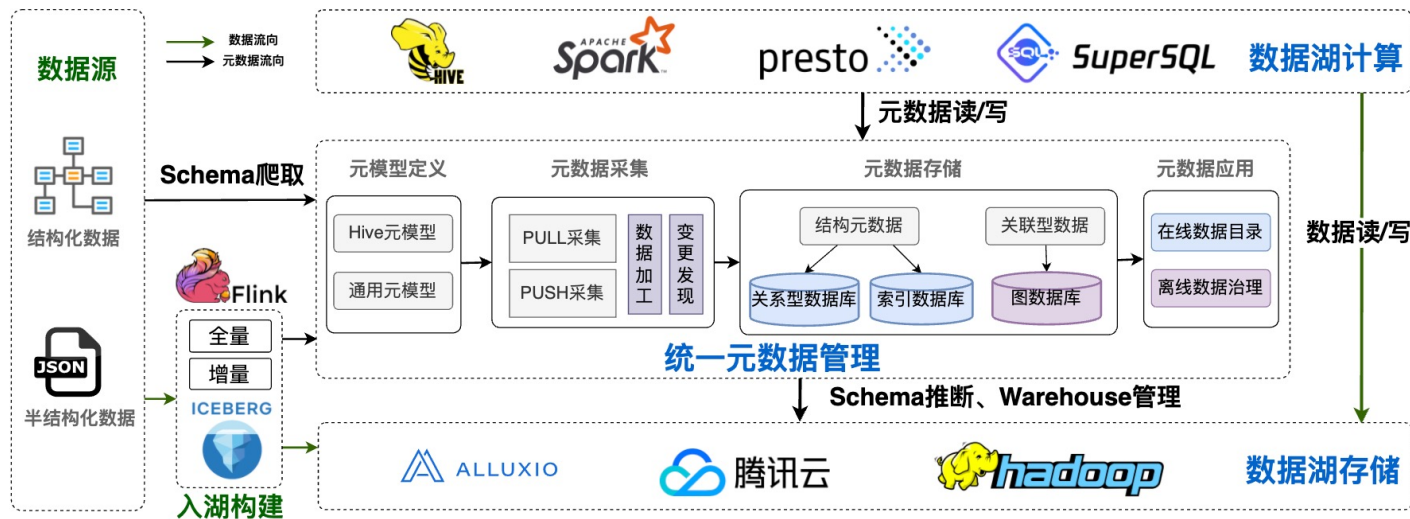
**湖仓一体**：数据湖(灵活性) + 数据仓库(建模/治理)

数据仓库(分而治之)	数据湖(无为而治)
数据模型/数据分层 (ODS、CDM、ADS)	—
结构化数据	结构/半结构化数据
Schema-on-Write	Schema-on-Read
数据质量高	数据质量较低
构建成本较大	构建成本较小

# 腾讯数据湖架构

统一元数据：数据湖的基石和枢纽

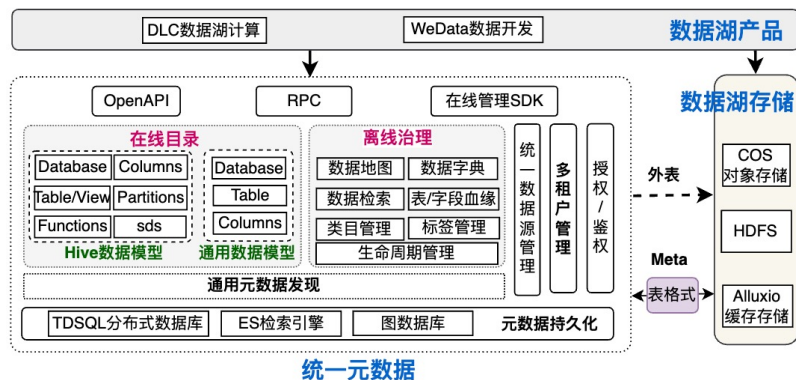
避免陷入数据沼泽、可打通业务产品的数据孤岛，提供统一数据资产视图



3+2架构

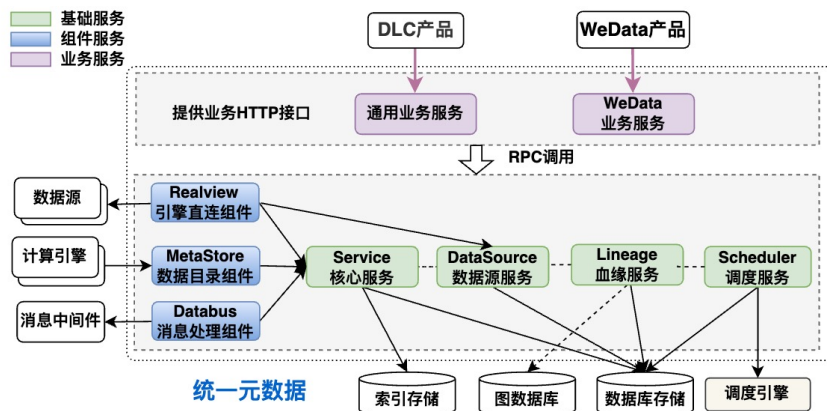
# 腾讯统一元数据架构

## 逻辑架构：在线+离线



常用Schema管理：Hive Metastore、AWS Glue  
Hive数据模型，借鉴原生Hive设计  
多租户管理：实现混合云场景的通用租户设计

## 服务架构：分层微服务+K8S+CI/CD



服务分层解耦业务、便于团队协作

# 02

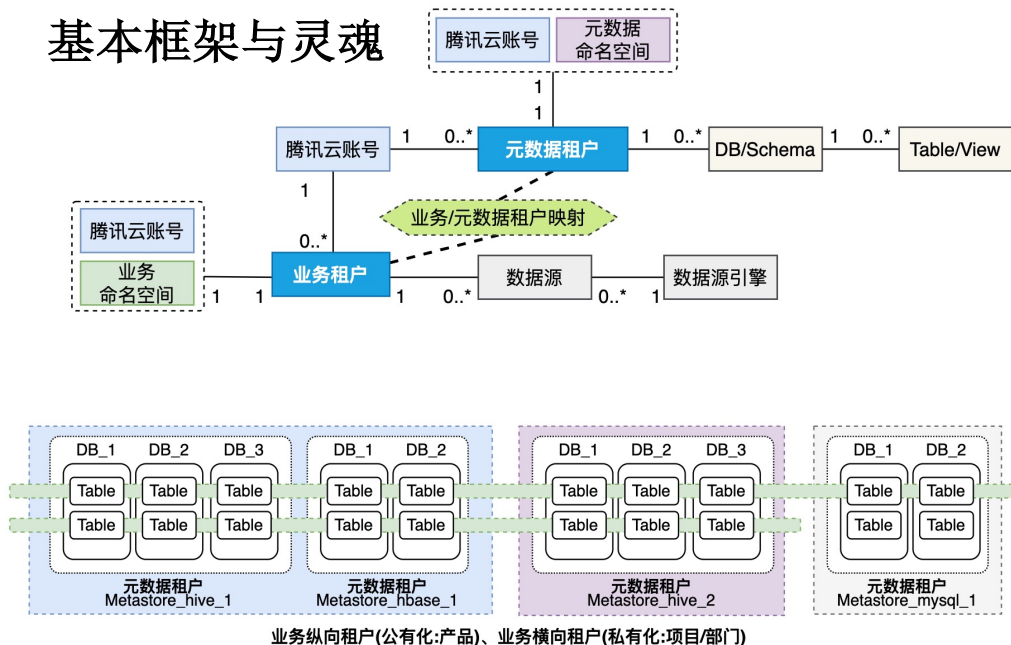
## 租户设计





# 多层次租户设计

## 基本框架与灵魂



## 元数据租户

- 最小租户粒度，可类比Hive Metastore
- 元数据租户=腾讯云账号+命名空间(别名)
- 支持不同元数据类型：Hive、MySQL等



## 业务租户

- 解耦通用元数据与具体业务
- 业务租户承担具体业务场景的关联
- 灵活的中间映射表维护





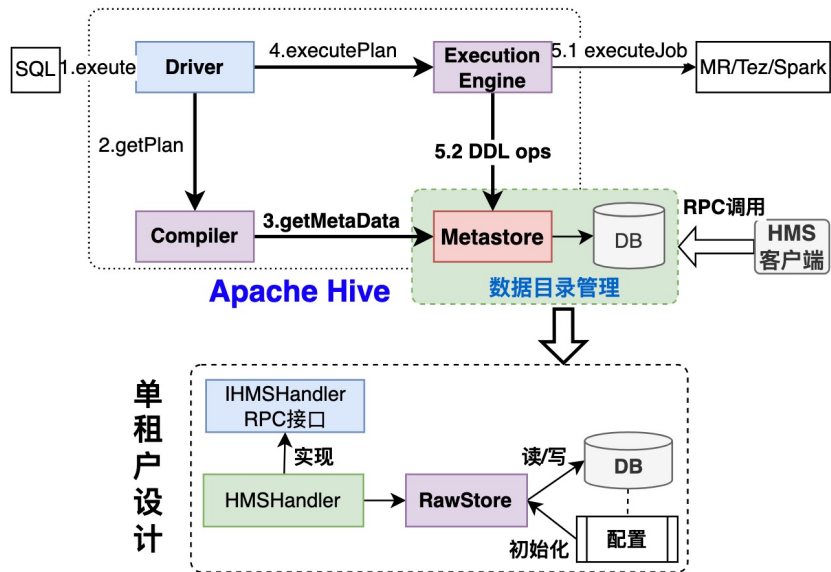
# 03

## 在线目录

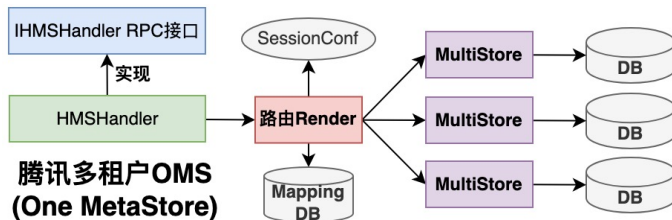


# 业界方案

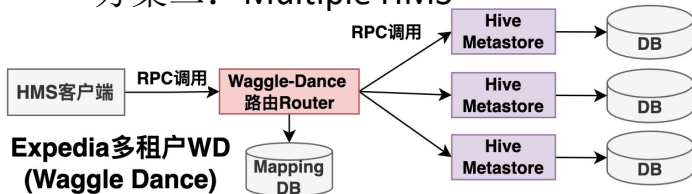
Hive Metastore：为SQL on Hadoop提供通用Schema管理  
能与计算引擎无缝对接



## 方案一：Multiple Store



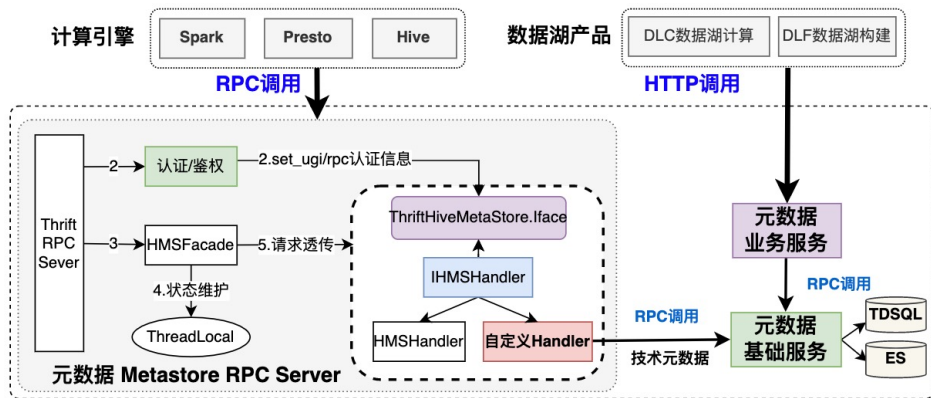
## 方案二：Multiple HMS



问题：大量数据连接维护、资源浪费、强绑定Hive

# 架构及实现

- 版本: Hive 2.3.7
- 总接口167, 已实现79<sub>(73+6)</sub>
- 无缝适配多种引擎

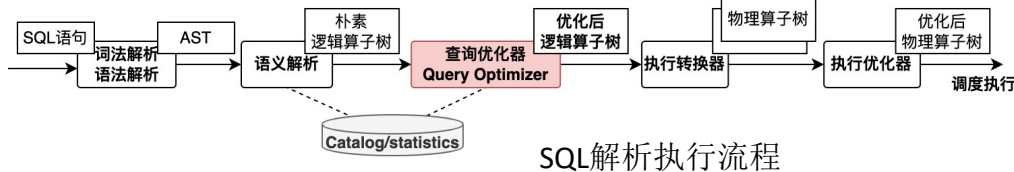


实现方案: 重新实现Hive Metastore RPC接口

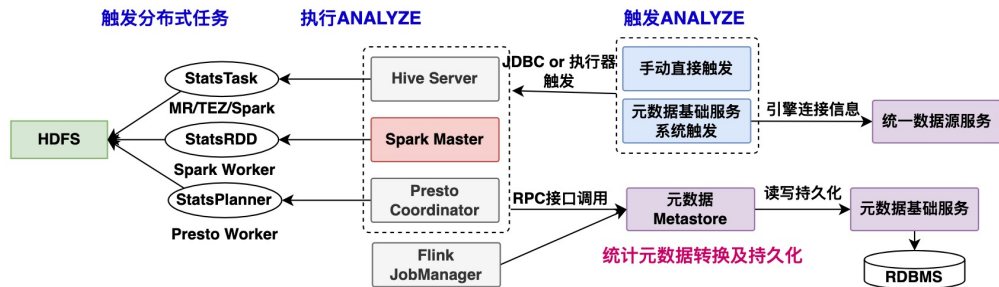
引擎	Database	Table	Partition	Function	Statistics	Lock	Role
PrestoDB(0.267)	✓	✓	✓	—	✓	—	✓
Spark(3.2.x)	✓	✓	✓	✓	—	—	—
Flink(1.14.0)	✓	✓	✓	✓	✓	—	—
Iceberg(0.12.1)	✓	✓	—	—	—	✓	—
Alluxio(2.7.1)	✓	✓	✓	—	✓	—	—
实现接口数	6	19	28	6	10	4	—



# CBO统计元数据



- 查询优化器：代数优化最短执行路径
- CBO：感知数据
- CBO要素=统计信息+代价模型
- 多引擎通用



多引擎通用的统计元数据

统计级别	统计信息	备注
表/分区	numFiles	数据文件个数
	numRows	数据行数
	rawDataSize	数据大小
	totalSize	总数据大小
字段	num_distincts	字段distinct数
	num_nulls	字段空值统计
	avg_col_len	字段平均长度
	max_col_len	字段最大长度
	histogram	字段等高直方图
	XXX_high_value	字段的最大值
	XXX_low_value	字段的最小值

# 04

## 离线治理

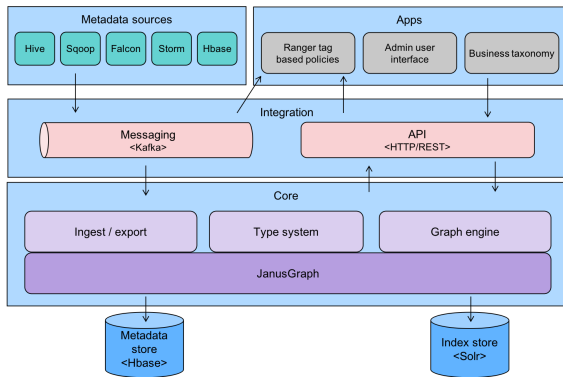


# 业界方案

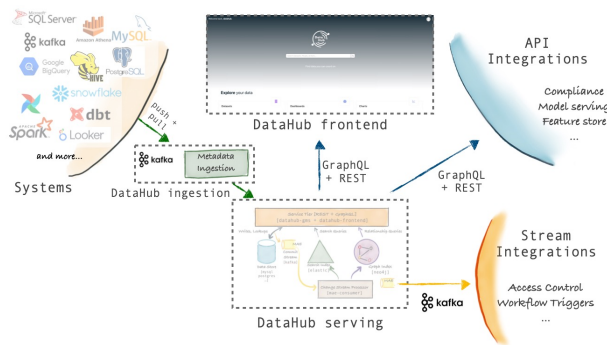
开源方案层出不穷

基本模块：元模型定义、元数据采集、元数据加工及存储、元数据应用

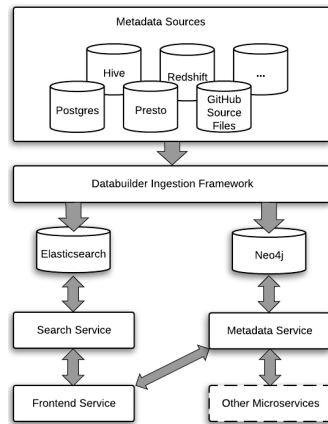
基础组件：关系型数据库、索引数据库、图数据库、消息中间件、调度引擎



Apache Atlas



LinkedIn DataHub(由Warehouse重构)



Lyft Amundsen

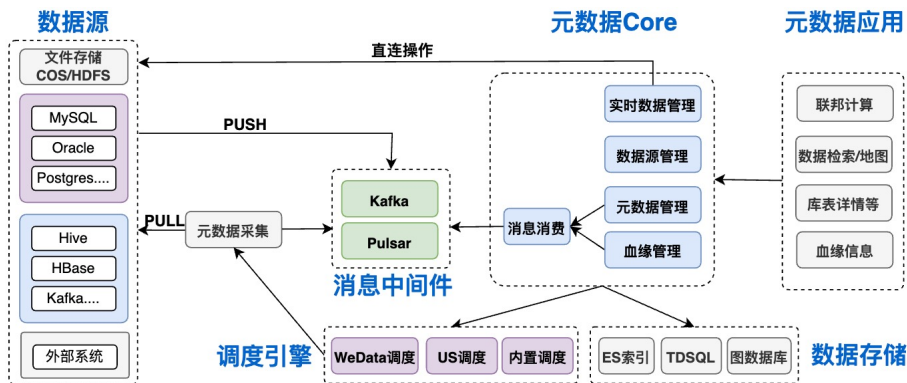


# 架构与实现

与业务整合需要深度的二次开发

功能	Atlas	DataHub	Amundsen	Metacat	Hybris
Schemas	✓	✓	✓	✓	✓
数据采集	✓	✓	✓	✗	✓
数据血缘	✓	✗	✗	✗	✓
数据预览	✗	✗	✓	✗	✓
库表管理	✗	✗	✗	✓	✓
数据库	HBase	MySQL	✗	MySQL	RDBMS
图数据库	JanusGraph	Neo4j	Neo4j	✗	Gremlin适配
数据索引	Solr/ES	ES	ES	ES	ES
调度引擎	✗	Airflow	Airflow	✗	可适配
编程语言	Java	Java	Python	Java	Java

与开源治理系统对比



# 非常感谢您的观看

---

DataFunCon 2021

