

网易游戏基于 Apache Doris 湖仓融合建设

胡彪

网易游戏 资深大数据开发工程师

分享嘉宾 — 网易游戏



胡彪

网易游戏 OLAP计算引擎负责人

负责 Doris、Trino、统一查询引擎的迭代演进工作

目录

01 架构演进

02 湖仓融合建设

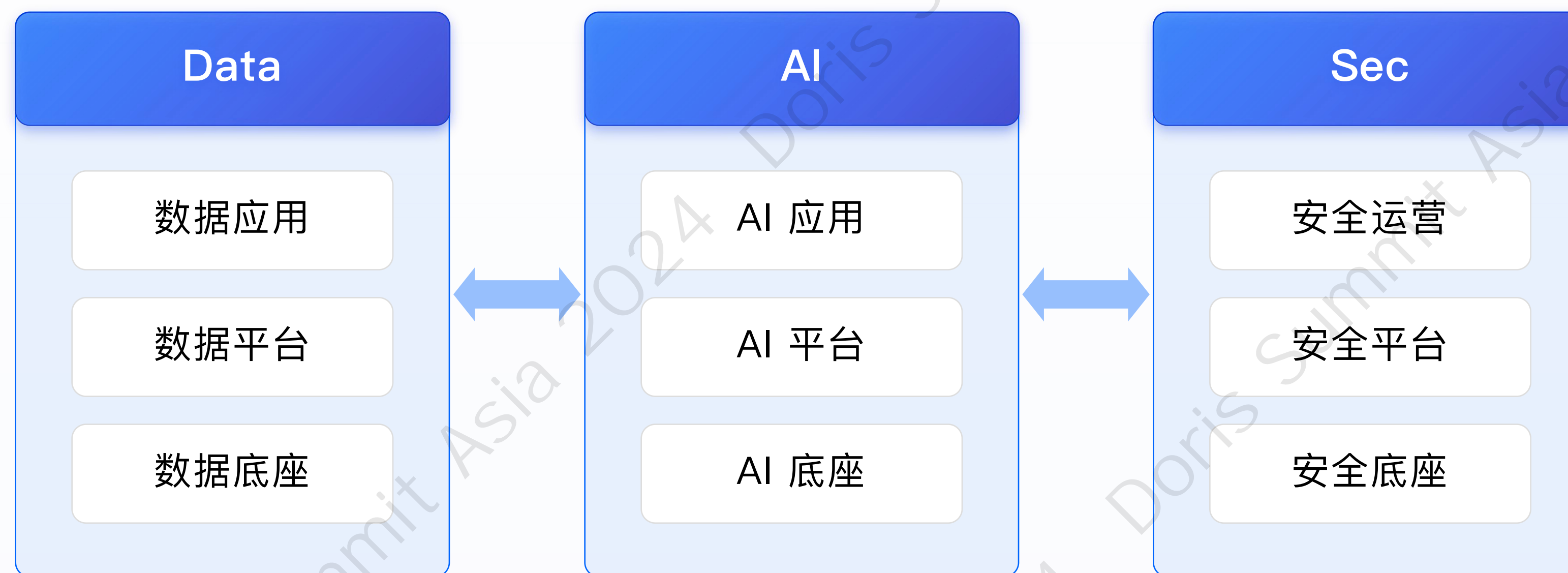
03 应用场景

04 未来规划

01

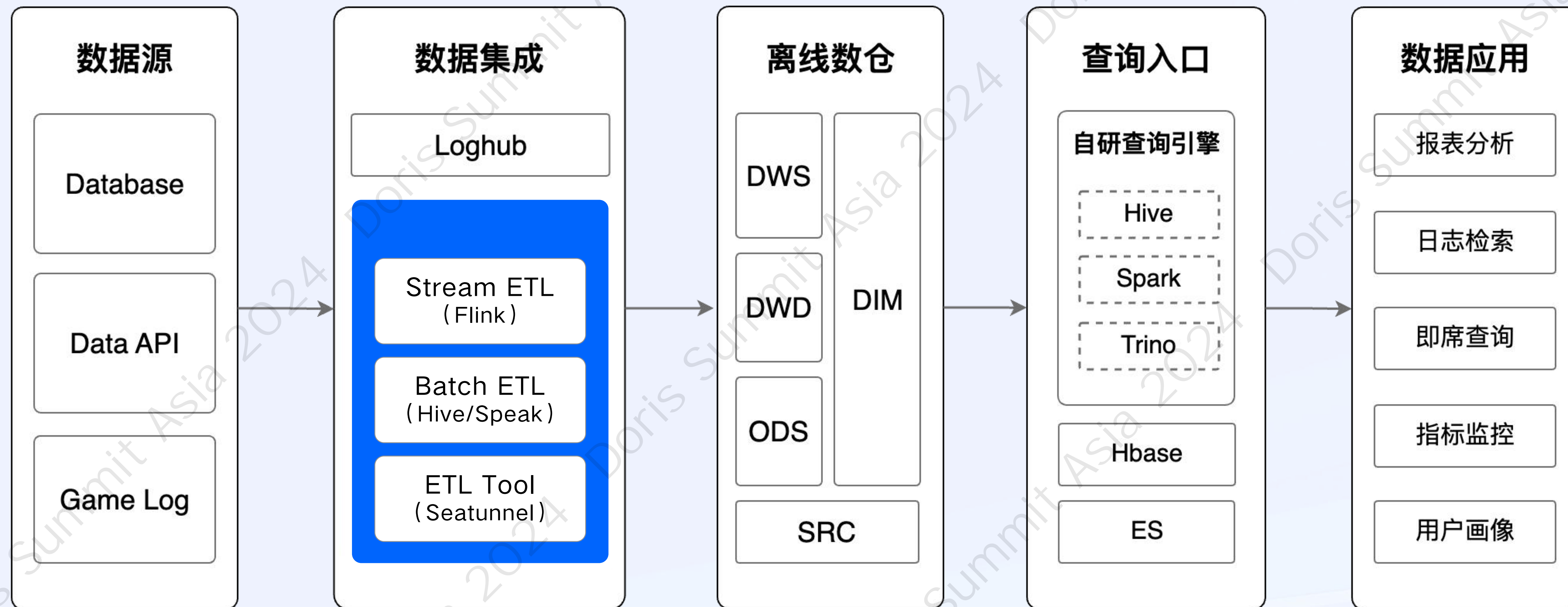
架构演进

背景介绍

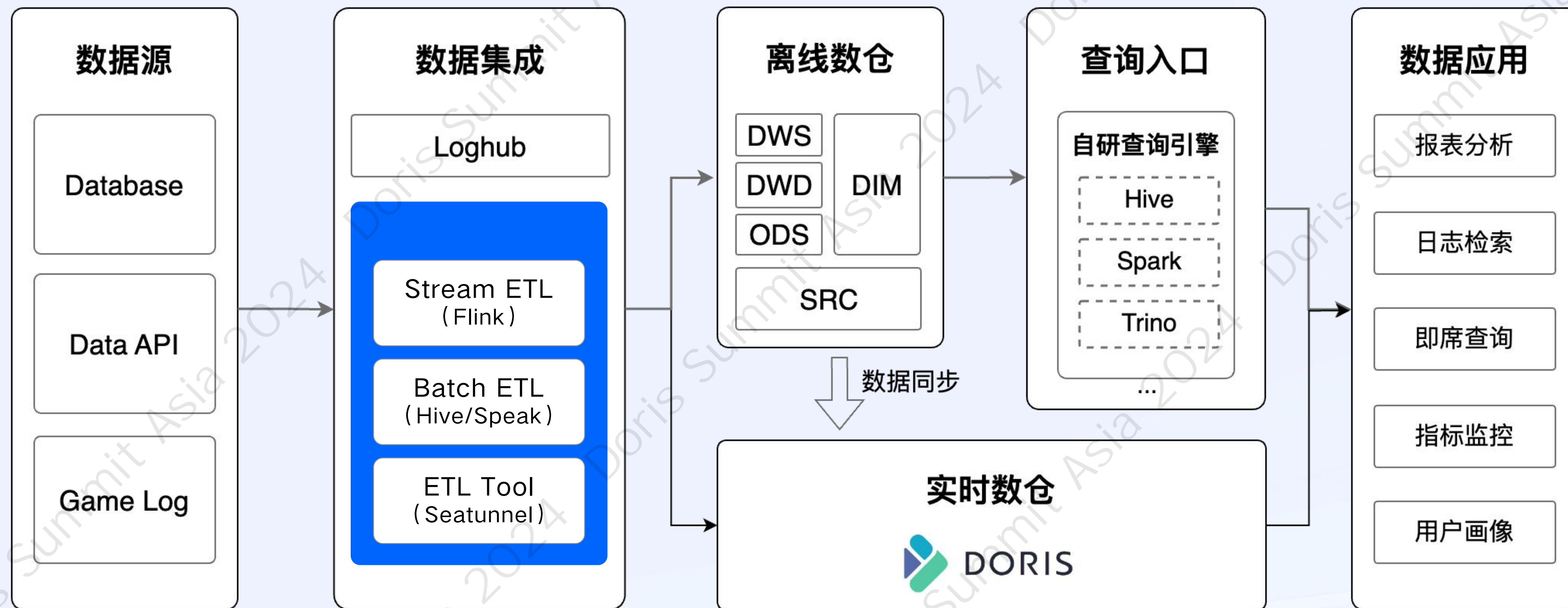


网易游戏-技术中心-效能研发部 的三个重点工作内容是数据、AI和安全。旨在通过数据和AI为公司众多游戏提供运营及决策支持，并且保护网易所有游戏相关的产品、服务和资源的安全。是推动游戏商业成功、品质提升以及渠道优化的重要支撑。

引入 Apache Doris 前的组件架构



引入 Apache Doris 构建实时数仓



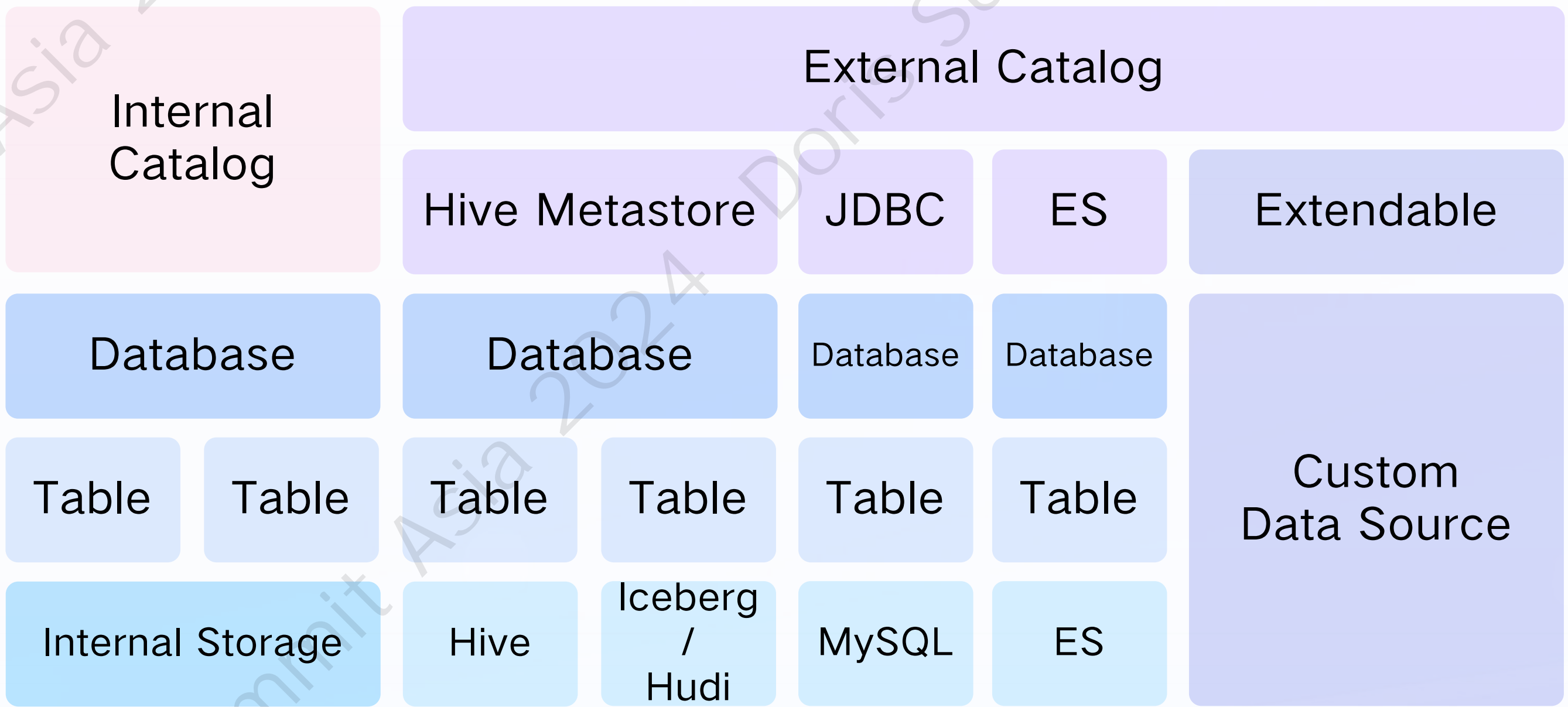
Apache Doris 架构演进

Schema Mapping



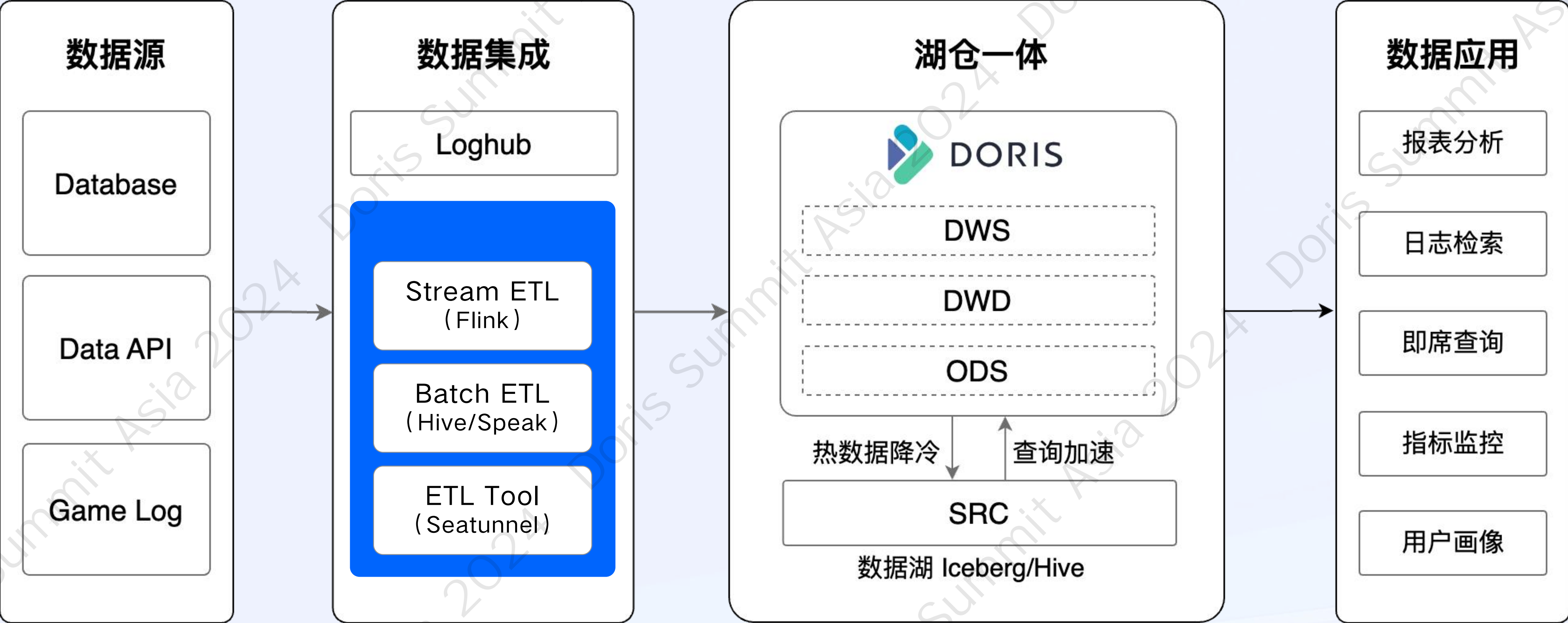
早期 1.0 版本支持通过外部表的方式注册外部数据源为内部表

Metadata



1.2 开始架构升级，定位为统一数据分析网关，加速各类异构数据源的查询和写入

基于 Doris 构建高效易用的湖上实时数仓

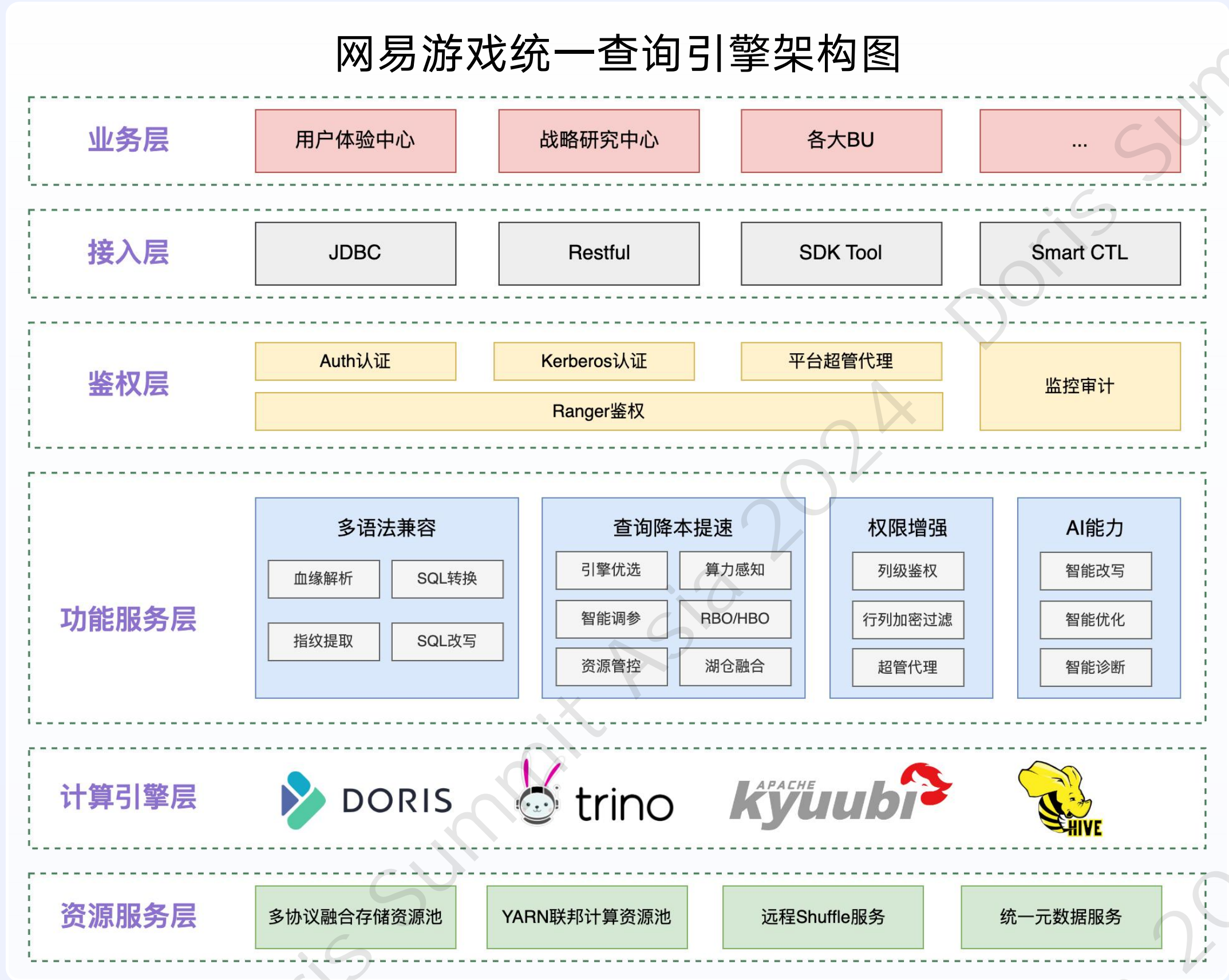


02

湖仓融合建设

网易游戏内部统一查询引擎 - SmartSQL

网易游戏统一查询引擎架构图



SmartSQL主要特点

1. 多引擎语法兼容

以 Hive 为基础扩展设计一套 SQL 语法，在解析层对底层所有引擎的语法进行兼容，自动适配不同计算引擎和数据源语法。

2. 多引擎融合查询优化作业效率和成本

采用 CBO、RBO 和 HBO 策略，实现智能路由、动态资源配置及智能调参，结合动态压力感知技术，显著提升作业效率并优化运行成本

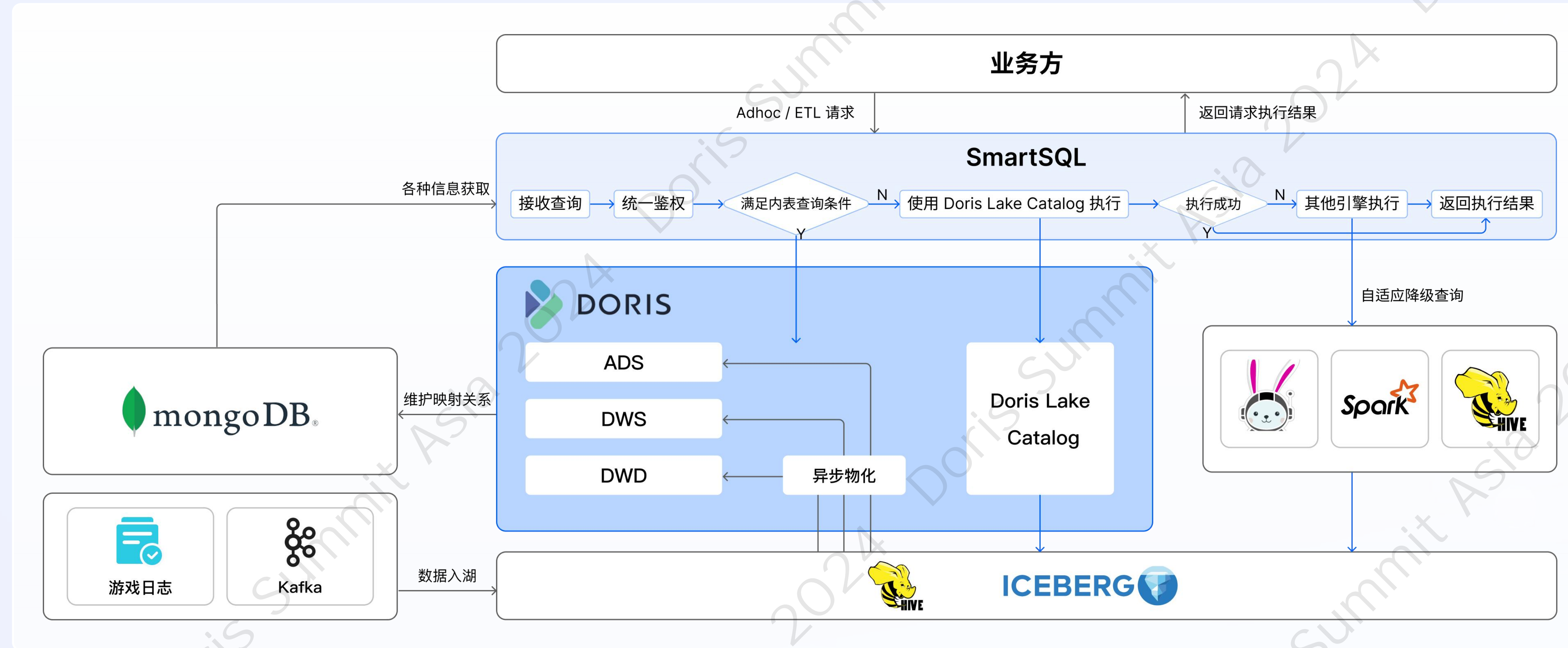
3. 多引擎统一权限管控与行列加密

结合 Ranger 实现了底层引擎无关的复杂鉴权动态脱敏，行列加密过滤、函数鉴权等复杂权限特性

4. 复杂场景统一接入方案

通过 SmartSQL 集成和兼容不同数据环境，实现业务零配置低门槛接入湖仓环境

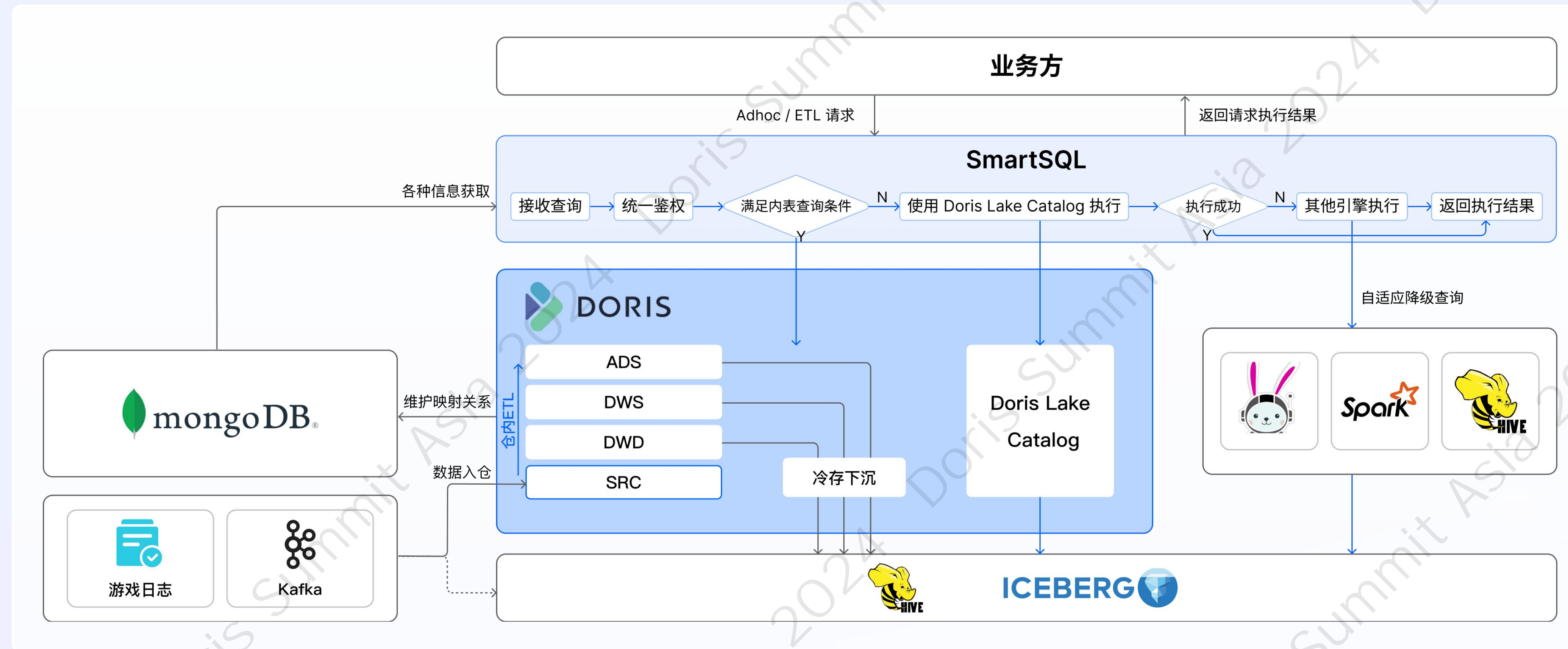
方案一：湖上建仓



特点

1. 业务数据原本的链路就是入湖、改动非常小；
2. 数据入湖的 checkpoint 较长，因此延迟在 5-10 min 左右；
3. 物化视图不宜频繁刷新，适用于离线分析为主的场景。

方案二：仓湖融合

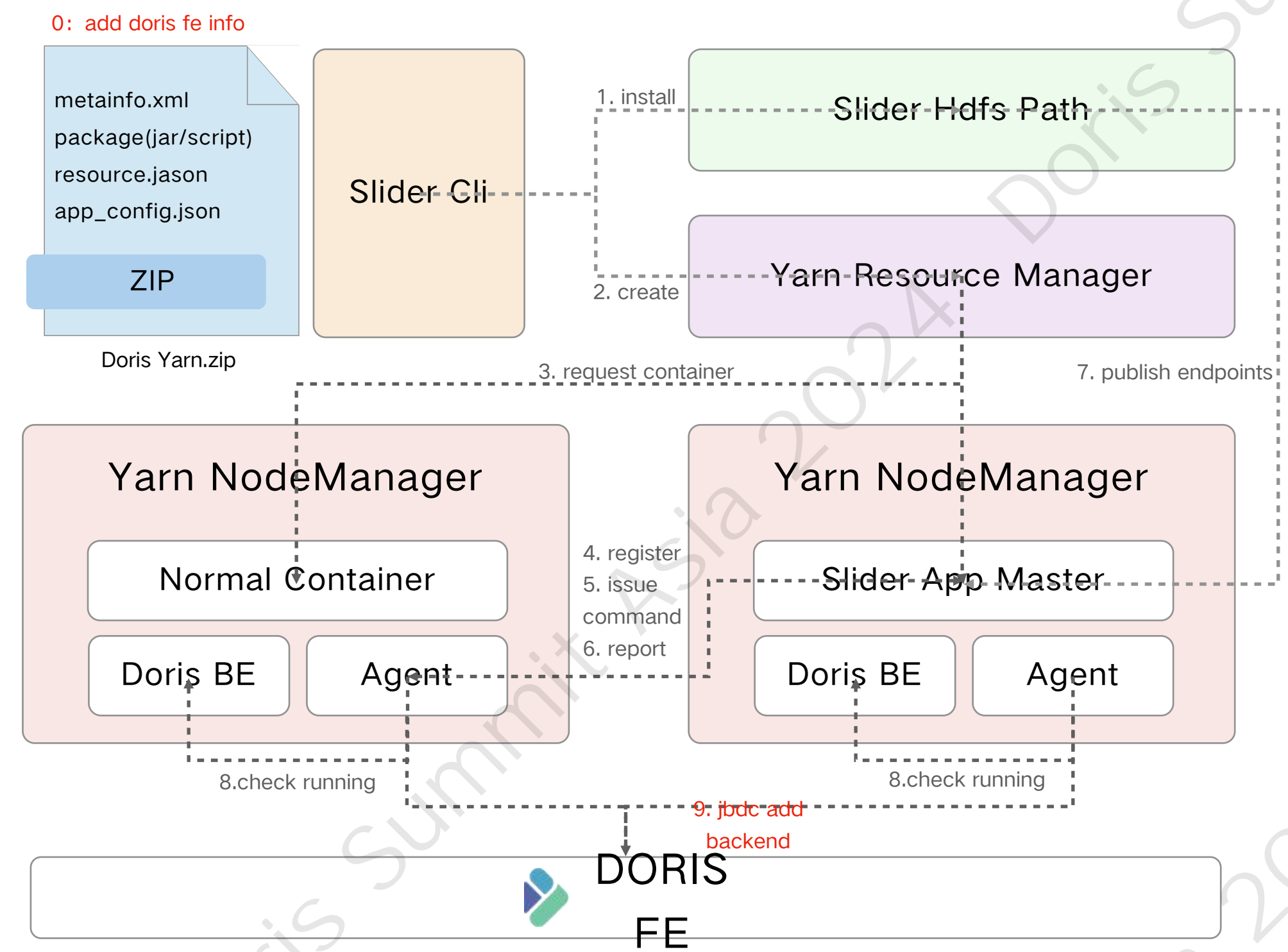


特点

1. 所有数据优先入仓、在当前版本数据量过大可能难以 ETL，因此适合于中小体量的业务；
2. 数据入湖的 checkpoint 较短，一般 30s 左右数据可见；
3. 以仓为主，更适合于实时分析为主的场景。

Doris on Yarn 实现弹性计算节点

Apache Slider拉起集群原理



Apache Slider 实现 Doris On Yarn 部署

- Doris 1.2 开始支持计算节点特性，核心配置 `be.conf`: `be_node_role=computation`
- Apache Slider 预定义了一套创建、启停、弹性、卸载的流程，简化 Deploy on Yarn 部署流程
- 定义 `doris-be` 服务单元，安装包、配置文件、启动脚本统一打包上传到 HDFS 上，启动时下发
- 快速弹性伸缩，按需使用：30s 拉起一批计算节点，5s 完成计算节点资源回收

Apache Slider 相关命令

- **安装配置**: `bin/slider package --install --name xxx - package xxx.zip`
- **创建应用**: `bin/slider create xxx`
- **卸载应用**: `bin/slider destroy xxx --force`
- **弹性启动**: `bin/slider start xxx`
- **弹性停服**: `bin/slider stop xxx`
- **伸缩节点**: `bin/slider flex xxx --component BE ±N`

Doris Workload Group 实现兼容 Trino 参数

Trino参数	含义	Doris实现
query.max-scan-physical-bytes	控制单个查询扫描的最大数据量的限制，避免单个大规模查询占用过多资源。	be_scan_rows : 统计某个查询在单BE上的扫描数据量 create workload policy canel_scan_1g_query conditions(be_scan_bytes > 1073741824) actions(cancel_query);
query.max-memory	限制了单个查询可以使用的最大内存量。超过此限制的查询将被终止。	自己实现，通过BE汇报给FE的数据限制
query.max-memory-per-node	定义了每个节点针对每个查询可以使用的最大内存量，用于控制单个节点的内存压力。	query_be_memory_bytes : 统计某个查询在单BE上的使用内存 create workload policy cancel_10m_query conditions(query_be_memory_bytes > 10240) actions(cancel_query);
query.max-execution-time	限制了查询的最大执行时间。包括排队时间和执行时间。超时的查询将被取消。	query_timeout : 指定用户的查询超时
query.low-memory-killer.policy	定义了在内存不足时，系统应采取的策略来终止查询，以释放内存	workload group : 当这个Workload Group的内存使用超过配置的硬限值时，会通过杀死查询的方式释放内存，避免进程内存被打满。
query.queue-config-file	指定了一个包含查询队列配置的文件路径，用于管理查询的排队策略。	max_concurrency : 最大查询并发数 max_queue_size : 查询排队队列的长度

融合统一工作



Hive UDF 兼容

UDF 类型

通用 UDF

业务 UDF

Doris UDF 特性

1. 兼容 Hive，极少改动
2. 支持动态加载 UDF
3. 支持远程加载 UDF
4. 支持库 UDF 和全局 UDF

Sql Convert 语法兼容

主要语法类型

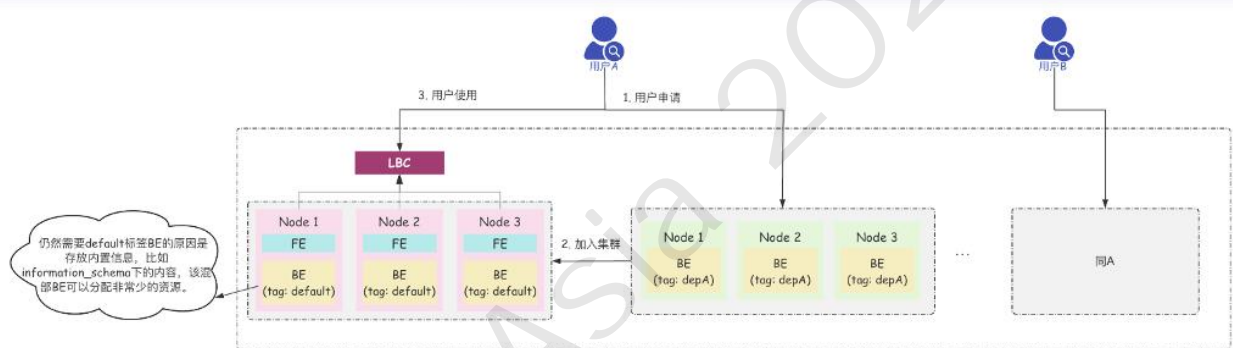
Trino

Spark

基于业务旧的知识启用不同语法

1. 对于新增新业务，要求使用 Doris 语法开发
2. 旧业务全局启用 Trino 语法，兼容历史 Trino 查询
3. 对于特定业务，支持独立指定 Spark 语法

BE存储隔离



问题描述

基于 Tag 的方式未隔离存储资源



内部增强

WaggleDance兼容

```
Create Catalog Action
| |--> catalogIf.getDbNames();
| |--> makeSureInitialized();
| |--> init();
| |--> getFilteredDatabaseNames();
| |--> listDatabaseNames();
| |--> metadataOps.listDatabaseNames();
| |--> client.getAllDatabases();
```

问题描述

Database -> Table -> Column 逐级缓存元数据
WaggleDance的getAllDatabases无法获取所有DB

Doris Manager支持外部物化视图管理

问题描述

1. 支持可视化创建外部物化视图；
2. 支持按SQL片段检索物化视图；
3. 支持查看物化视图的详细task；
4. 支持立即触发、暂停、删除物化视图；
5. 统计物化视图的命中次数；

03

应用场景

网易游戏内部应用场景

场景	网易游戏相应业务	Apache Doris相关特性
大宽表场景	替代 CK，如问卷业务、CDN 业务等	底层实现与 CK 基本一致、性能相当，但运维难度小于 CK
AdHoc 分析引擎库	替代 Trino，所有业务场景 Adhoc 的第一选择	向量化引擎性能更优，存储引擎相当于一层缓存，结合sql-convert可实现兼容大部分SQL语法
用户行为分析引擎库	网站点击行为、付费事件跟踪、用户画像等业务	在 bitmap 位图索引上做了大量优化，bitmap 函数丰富，并且拥有light schema change的能力
单表高并发点查引擎库	游戏交易网站首页商品缓存	Doris 2.0 采用了行列混存数据存储模式、支持短路径查询
日志检索引擎库	各类大数据组件底层日志存储、游戏日志存储	Doris 用 Loki 为代表的轻量索引/无索引架构，设计了倒排索引和全文检索特性
实时数仓	所有实时场景、对接各类看板	导入、存储、查询全线优化，导入支持事务性、多种表模型及各类索引和物化视图支撑数仓分层加速查询
统一网关平台	Doris/Mysql/Hive/Iceberg 等联邦查询	Doris 自带了联邦查询模块，可以有效平替 Presto/Trino 的应用场景，性能还要优于 Presto 2-3 倍

Doris 作为元数据存储底座在 AI 层面的应用



04

未来规划

未来规划

数据湖解决方案推广

将基于 Apache Doris 的数据湖解决方案在更多业务部门和业务场景进行推广，助力用户降本增效

实现智能物化

物化视图是加速湖查询的一大利器，基于用户作业对热点 SQL 片段物化将是后续的重点工作

3.0 版本跟进

3.0 是基于云原生存算分离的全新架构，在数据湖场景提供了更多业务模式的可能性

内部Manager建设

内部 Doris Manager 对于 2.1 及 3.0 版本新特性的管理运维支持，尤其是对用户最关心的功能迭代开发

Thanks for Watching!