

Apache Doris × 字节跳动开源联合 Meetup

8月31日 (周六) 13:30-17:00



Apache Doris 3.0 - 存算分离架 构架构升级 湖仓一体再进化

陈明雨 | 飞轮科技技术副总裁 & Apache Doris PMC Chair



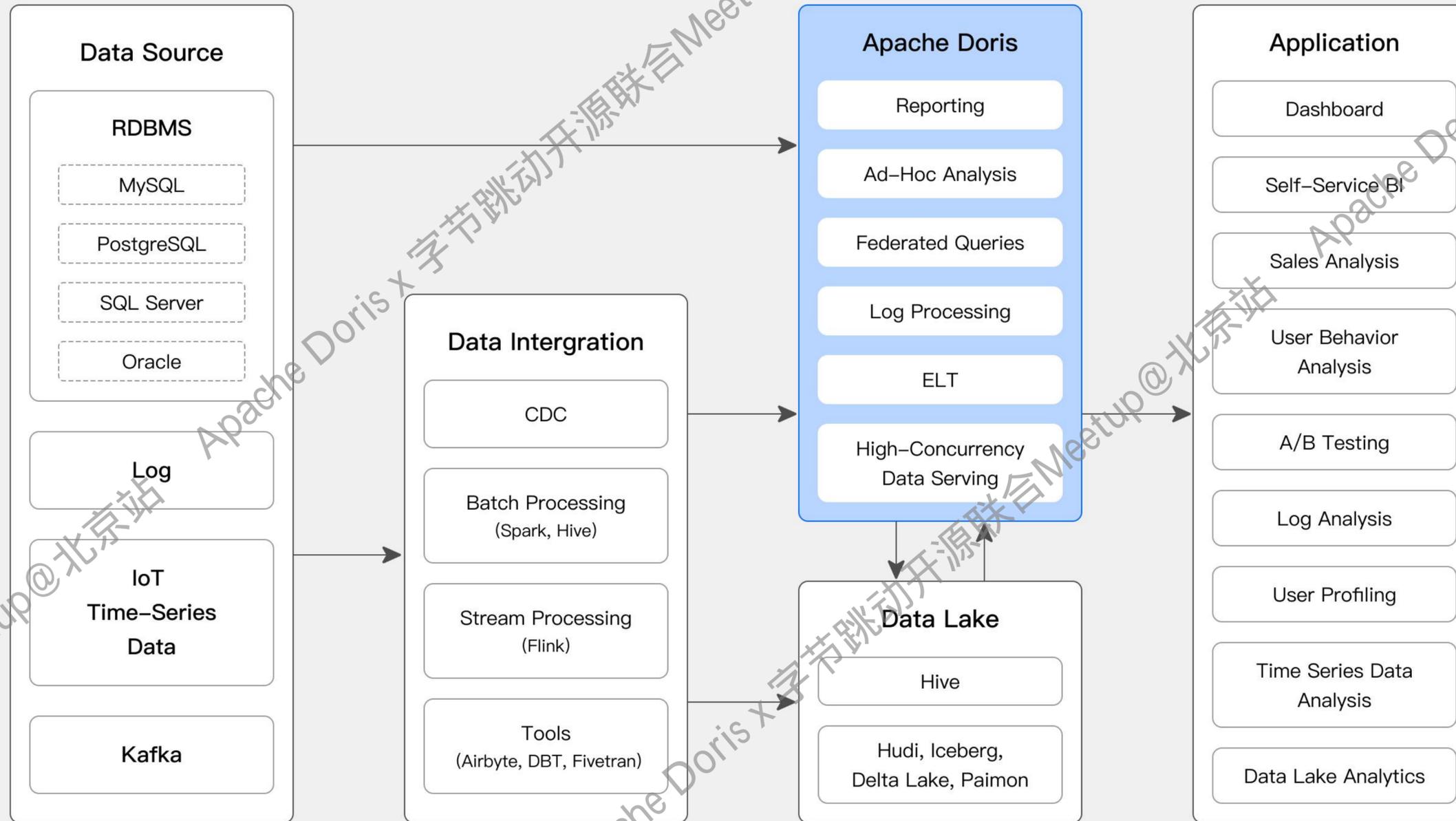
个人介绍



陈明雨

- 飞轮科技技术副总裁、Apache Doris PMC Chair、Apache Member
- 曾担任百度 Doris 团队技术负责人，主导了 Apache Doris 从毕业成为 Apache 基金会顶级项目的全过程
- 有近 10 年分布式数据库架构设计、研发与团队管理经验

What is Apache Doris



目录

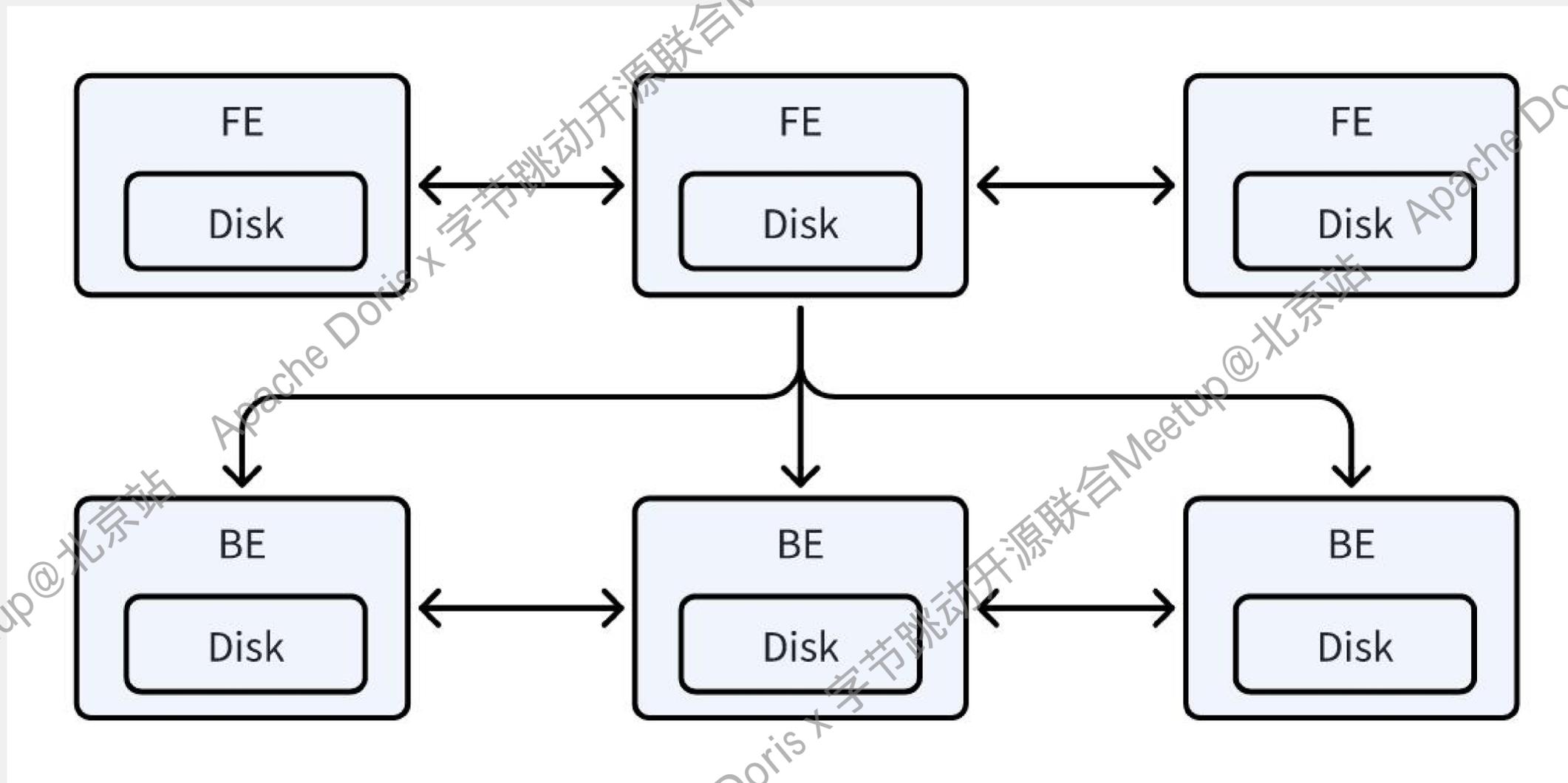
01 存算一体 or 存算分离

02 如何设计面向未来的架构

03 湖仓一体再进化

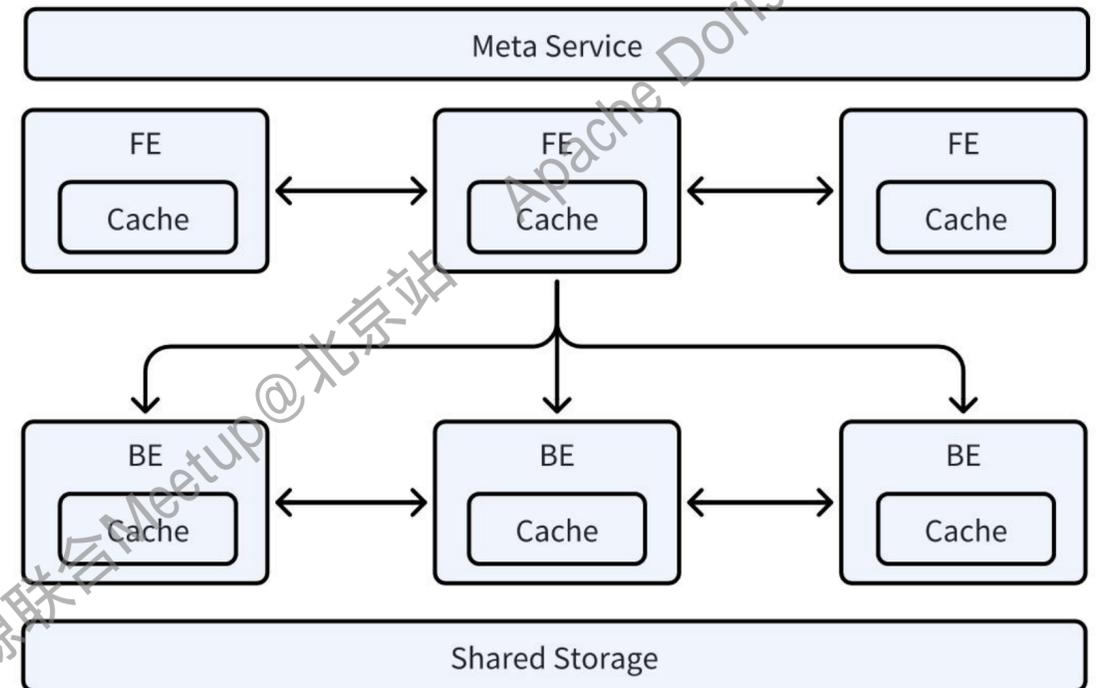
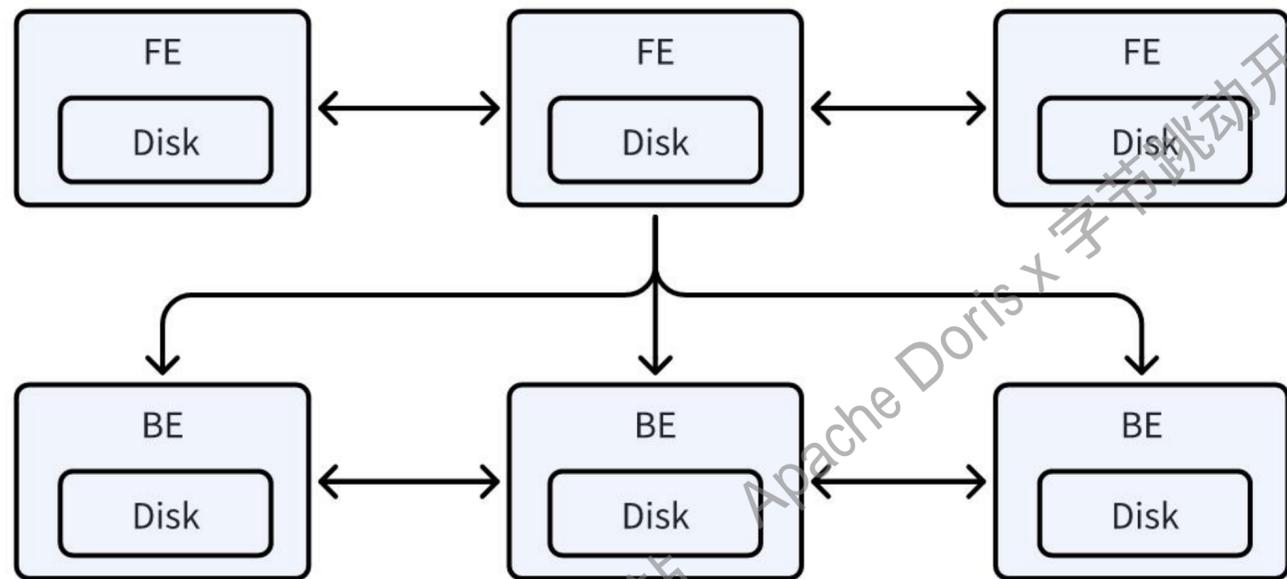
04 社区发展规划

Apache Doris 存算一体模式



在存算一体架构下，BE 节点上存储与计算紧密耦合，数据主要存储在 BE 节点上，多 BE 节点采用 MPP 分布式计算架构。

Apache Doris 存算一体模式



- **部署简单**: 仅FE与BE进程, BE和FE都可以单独扩容
- **稳定可靠**: 不依赖共享存储系统
- **性能优异**: 计算节点访问本地存储

- 为什么要存算分离?

为什么需要存算分离

低成本与资源弹性

- 计算和存储解绑，单独扩缩容
- 计算资源波谷波峰，灵活弹性
- 数据存储冷热效应明显

负载隔离

- 读写任务分离
- 更彻底的业务隔离，解决不同业务间的相互影响以及资源抢占问题

数据共享

- 单一数据面向不同的分析负载使用
- 数据快速移动、快速备份恢复
- Single Source of Truth

云基础设施的成熟

- 云上基础设施逐步完善，提供可靠的共享存储
- 完全按量付费，灵活可控

目录

01 存算一体 or 存算分离

02 如何设计面向未来的架构

03 湖仓一体再进化

04 社区发展规划

设计出发点 - 性价比与架构稳定性



如何降低成本

- 引入对象存储节省冷数据资源
- 增加弹性计算能力，按需使用计算资源



不同架构如何迭代

- 绝大多数用户已采取存算一体架构
- 升级过程中需要保证对已有架构的兼容

设计目标



负载隔离

读写分离

业务隔离

内部负载隔离



低成本

存储成本大幅下降

计算和存储可以独立弹性

使用业务的波峰波谷调整计算资源

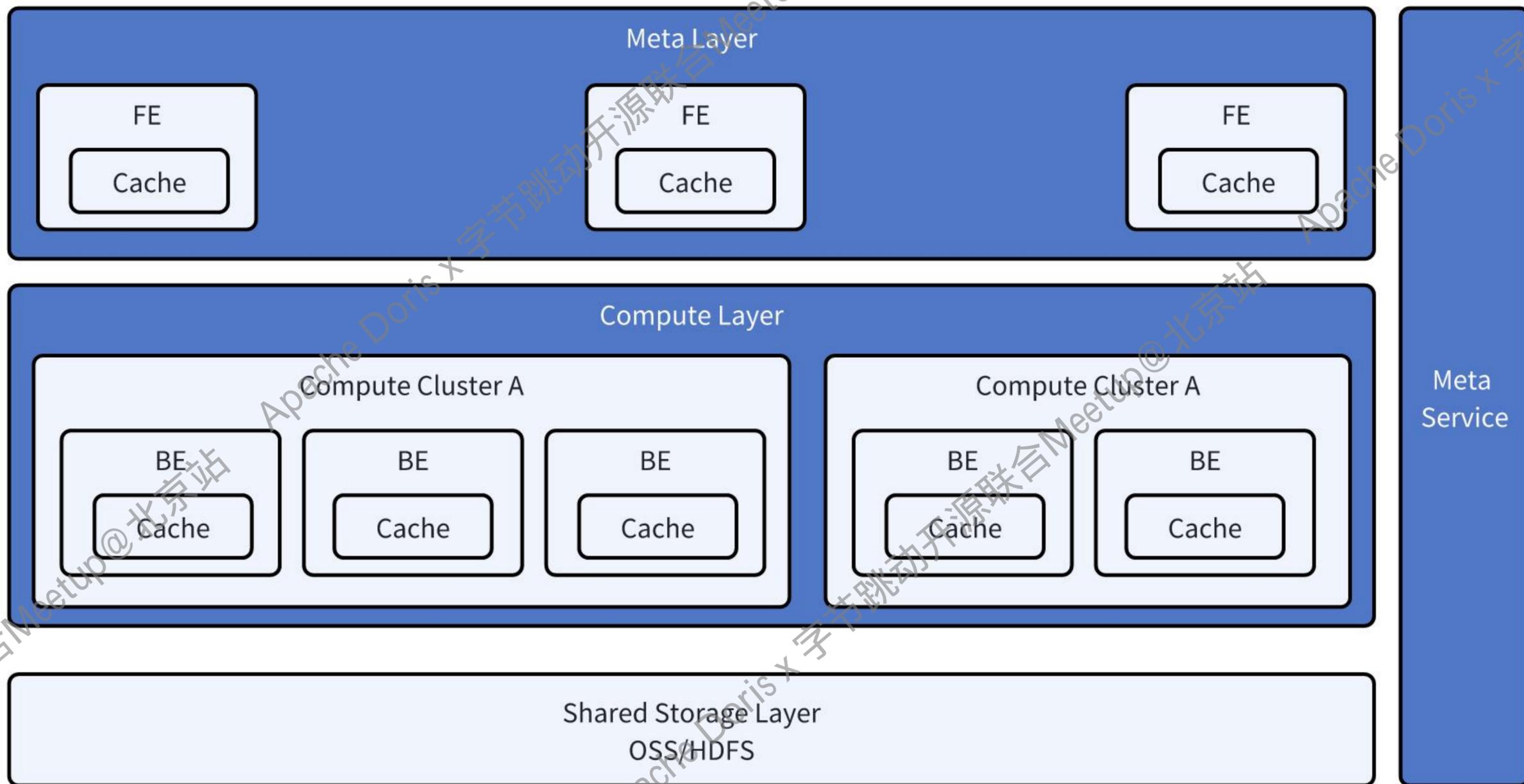


数据共享

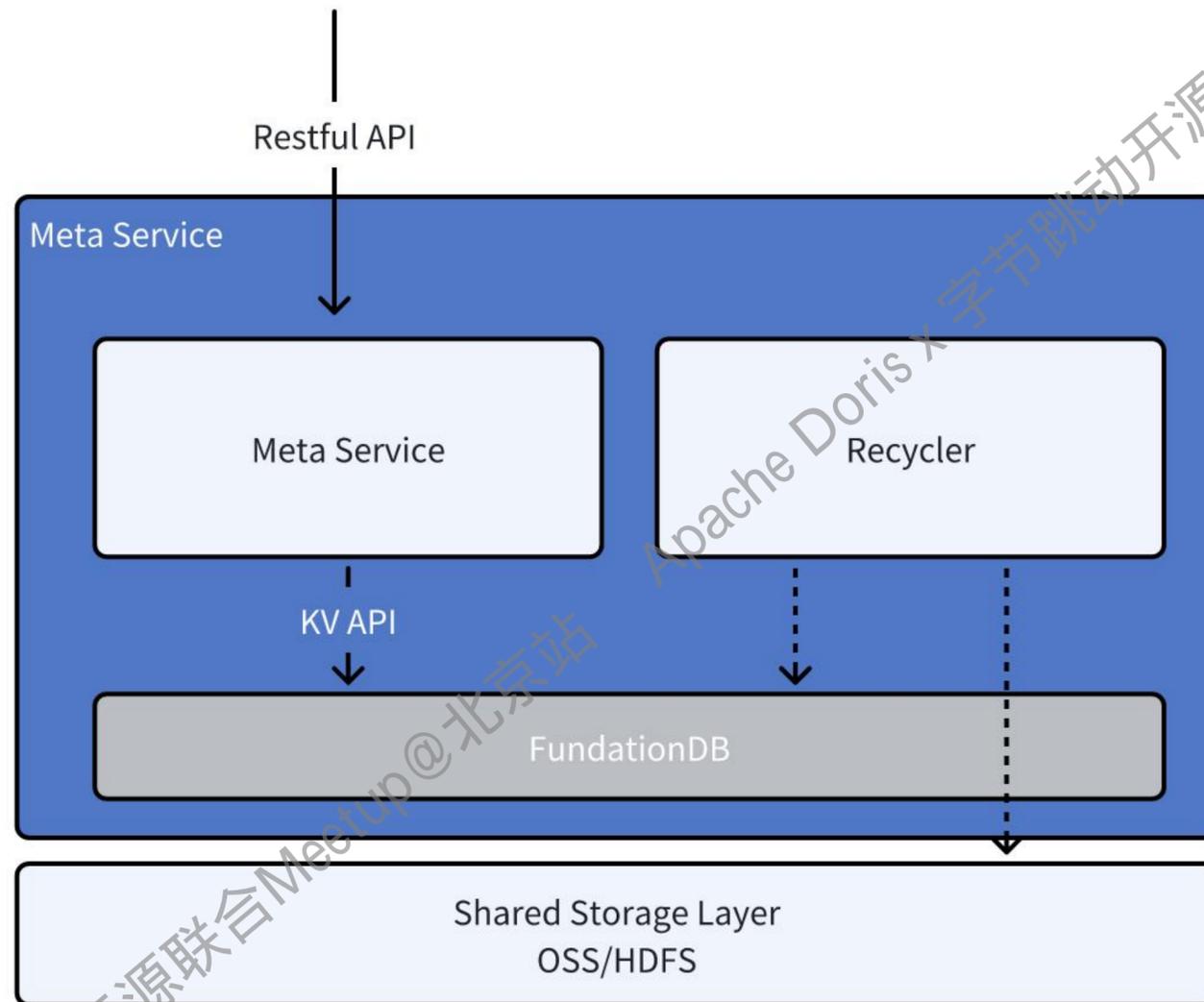
统一元数据服务

统一存储

存算分离整体架构



元数据服务层



- 统一语义层：Restful API
- 高性能分布式KV存储：FoundationDB
- 正向垃圾回收：Recycler

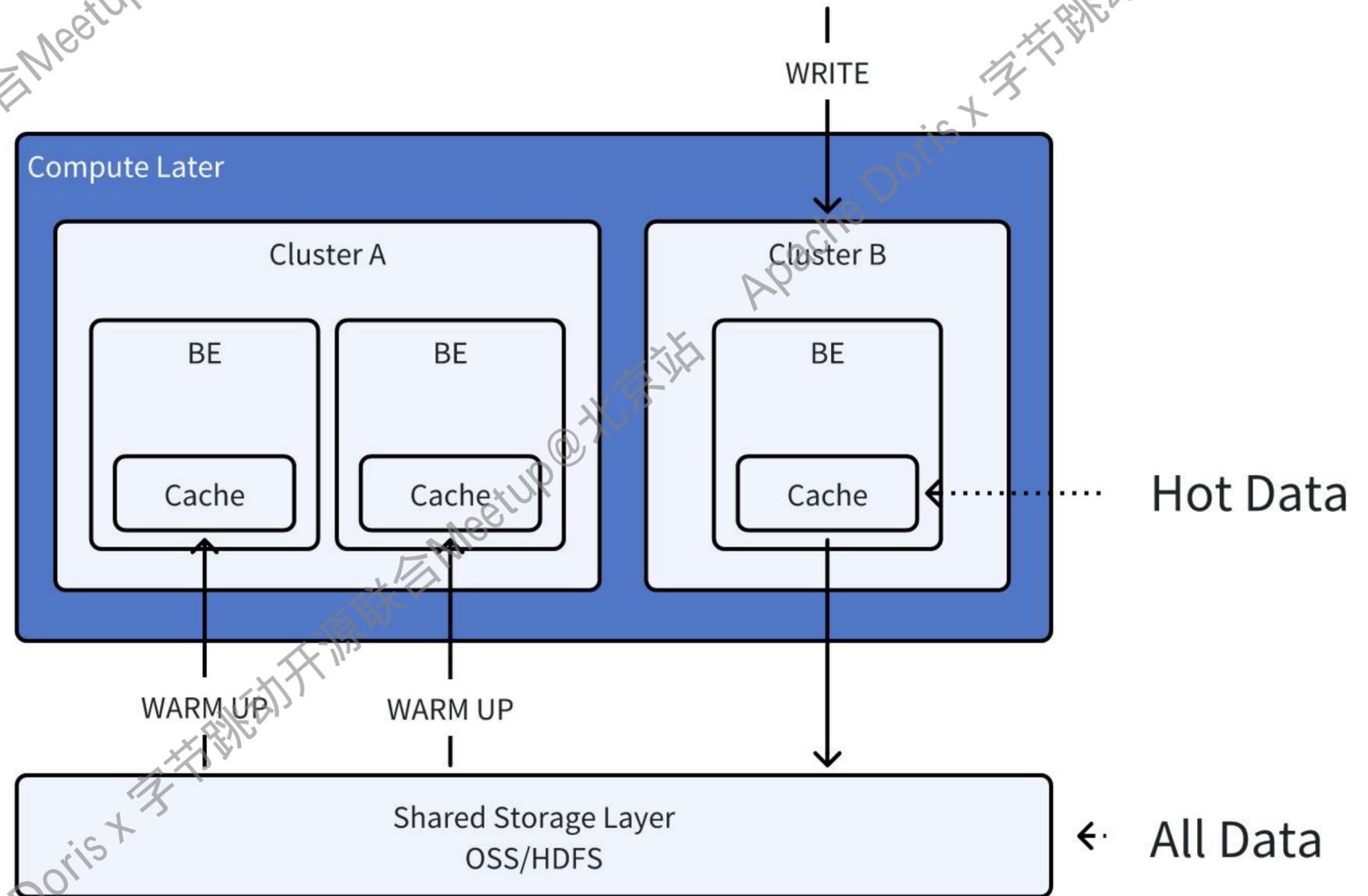
数据存储层

成本最高降低90%

- 存算一体
- 全量数据 * 3 * 块存储价格
- 存算分离
- 热数据 * 1 * 块存储价格 + 全量数据 * 对象存储价格
- 最高可以节省90%以上

灵活的Cache管理

- LRU, TTL
- 缓存预热

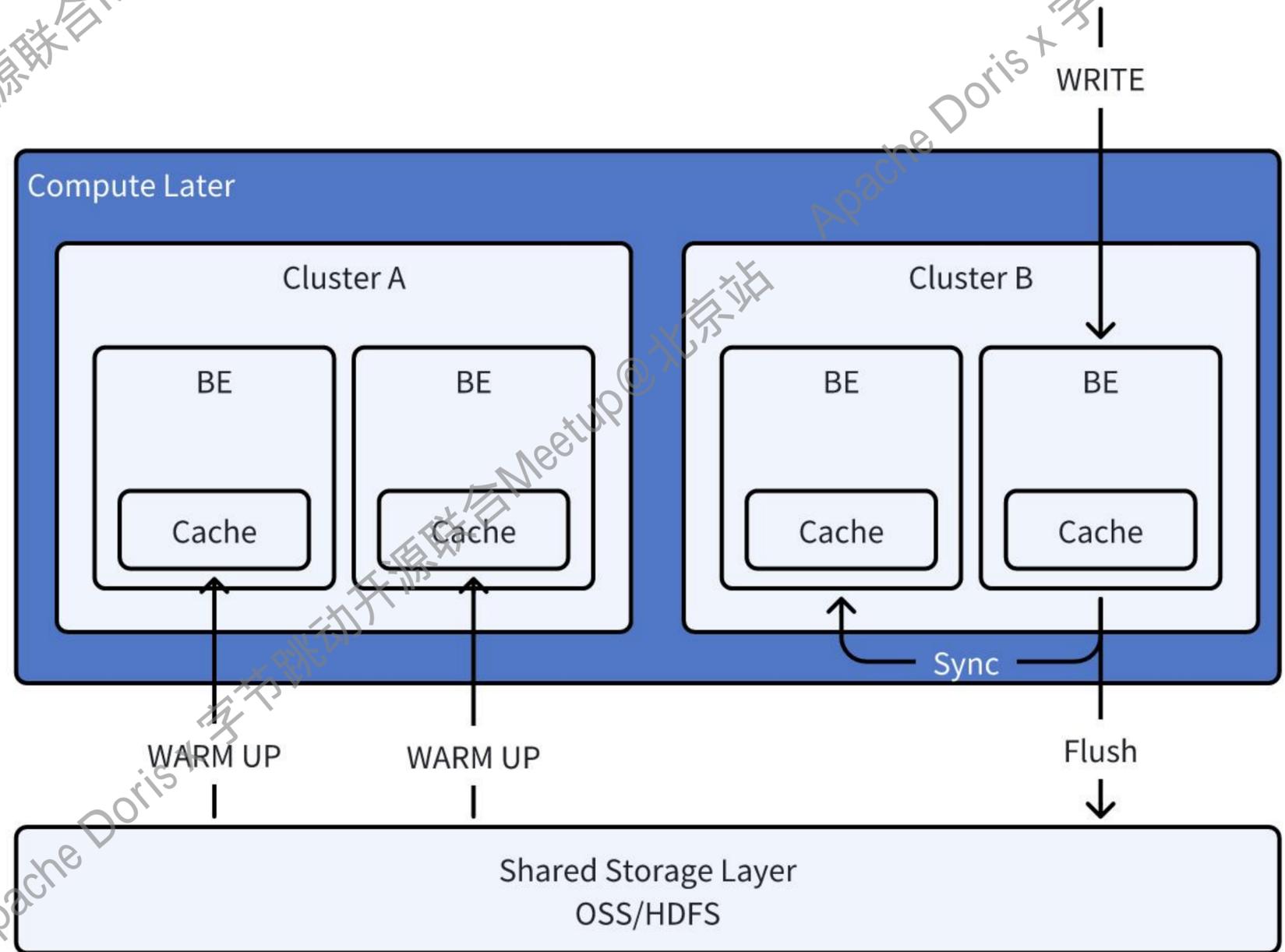


数据计算层-数据导入

1. 数据进入协调者 BE
2. 数据分发到多个 BE
3. 数据写入 Cache
4. 数据写入 S3
5. 读写分离 Cluster 预热 Cache

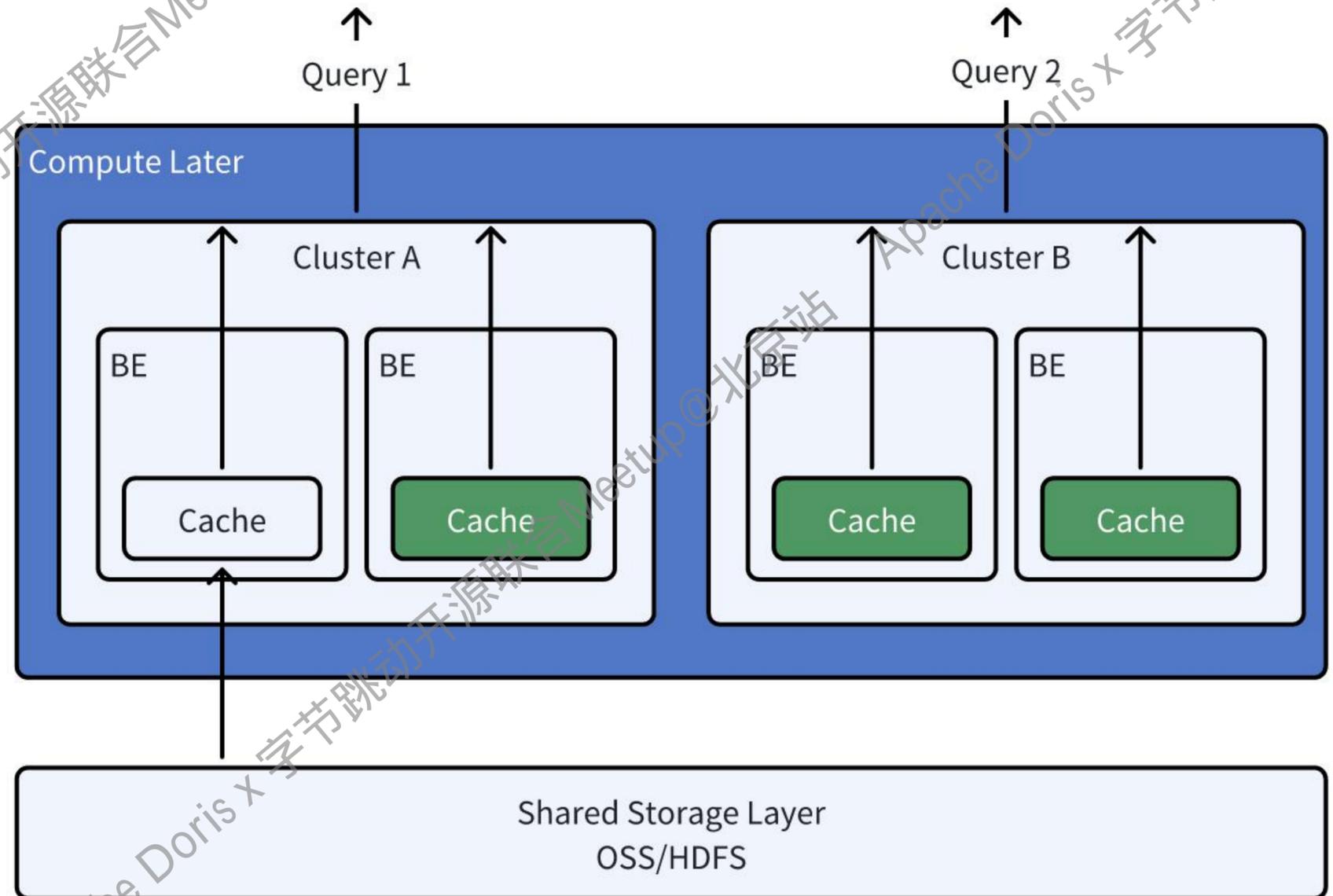
数据导入效率更高

- 只需处理单副本数据
- 数据和BE没有固定的关系
- 没有 Publish 阶段，写入流程更短

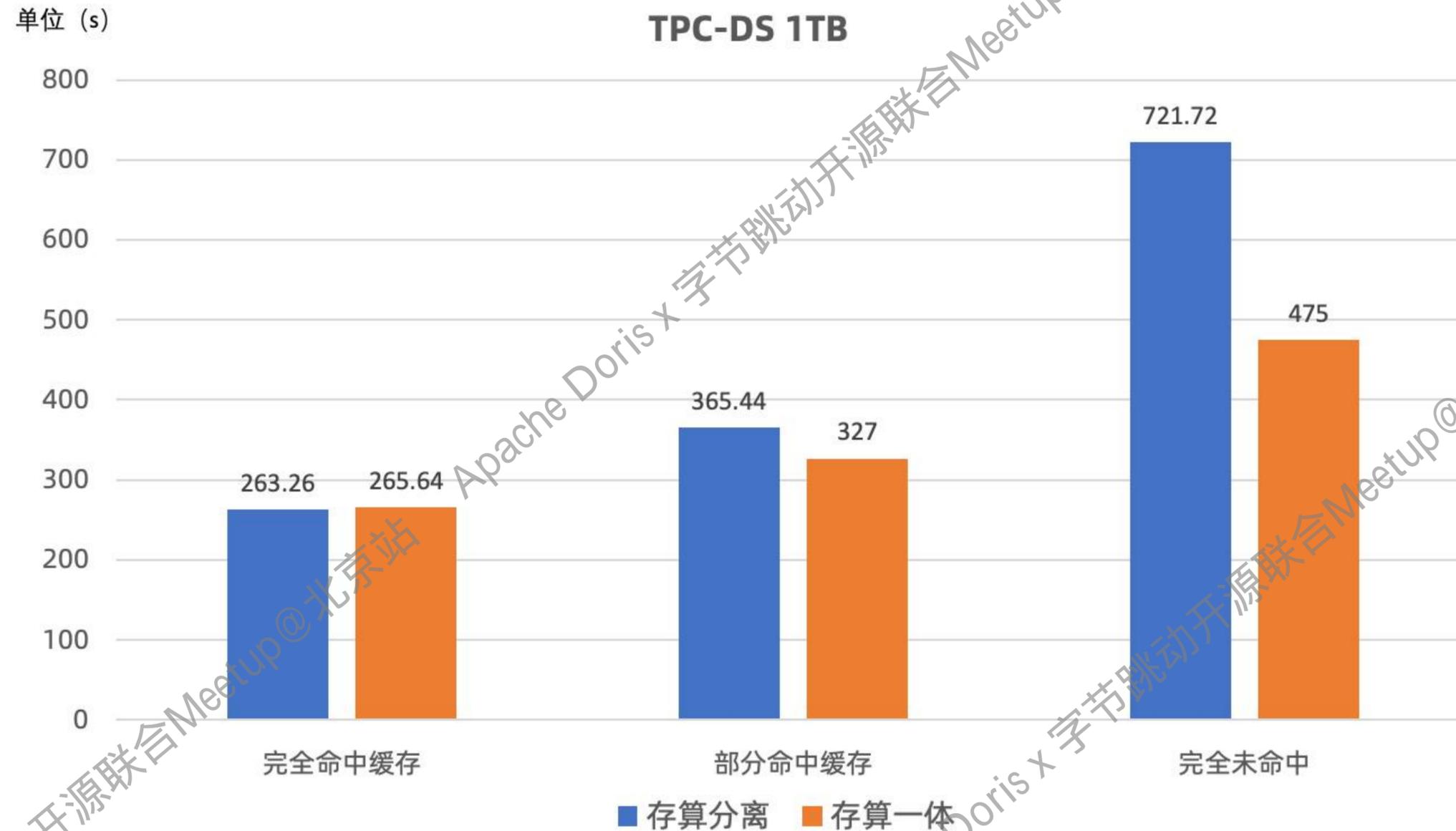


数据计算层-数据查询

- 多个 cluster 独立
- 不命中 Cache 时从 S3 读数据
- 命中时从本地 Cache 读数据
- 弹性资源大幅降低成本



查询性能对比

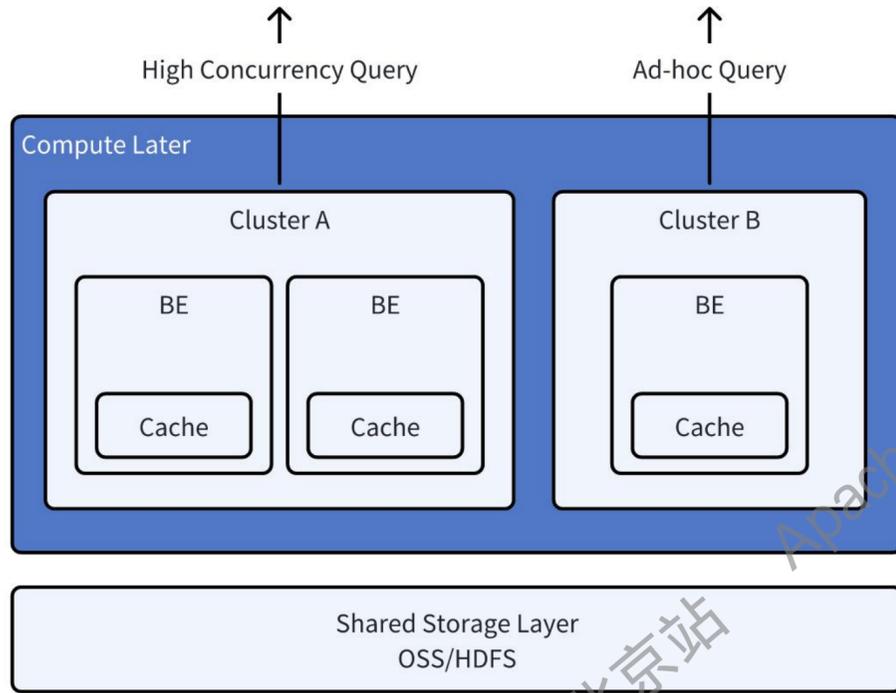


缓存层级

- Remote Storage
- Local File Cache
- System Page Cache
- Doris Page Cache

完全命中缓存时查询性能完全持平，部分命中缓存时有10%的性能损耗、随测试进行数据逐渐加载进缓存，性能随之提升；极端情况下（完全未命中任何缓存）性能损耗约30%。

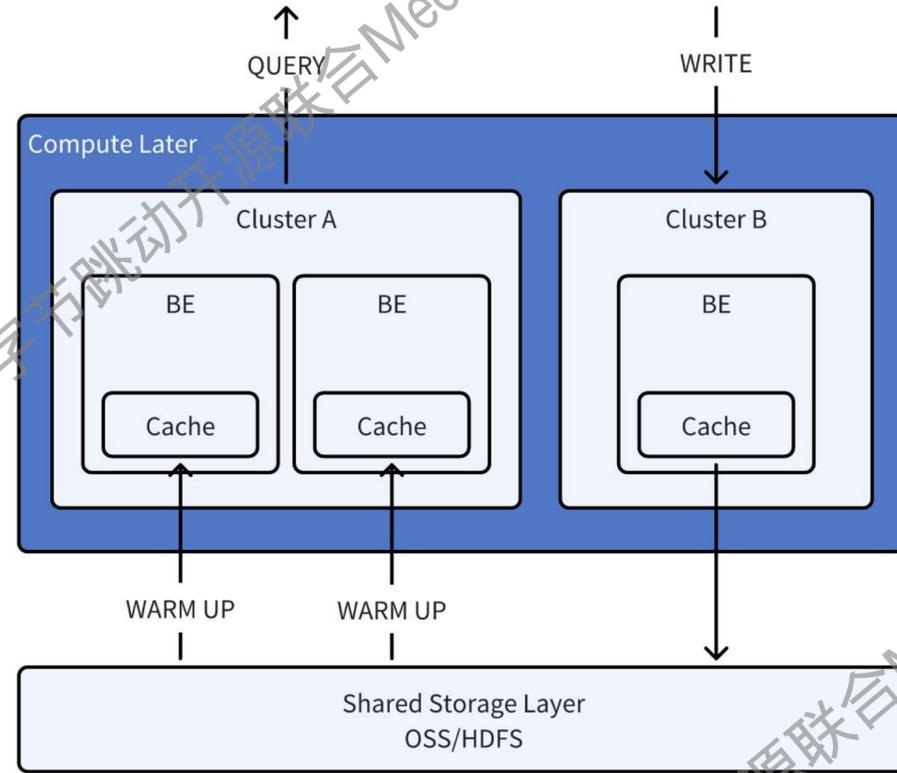
最佳实践



- **读读隔离**

- 高优查询和普通查询

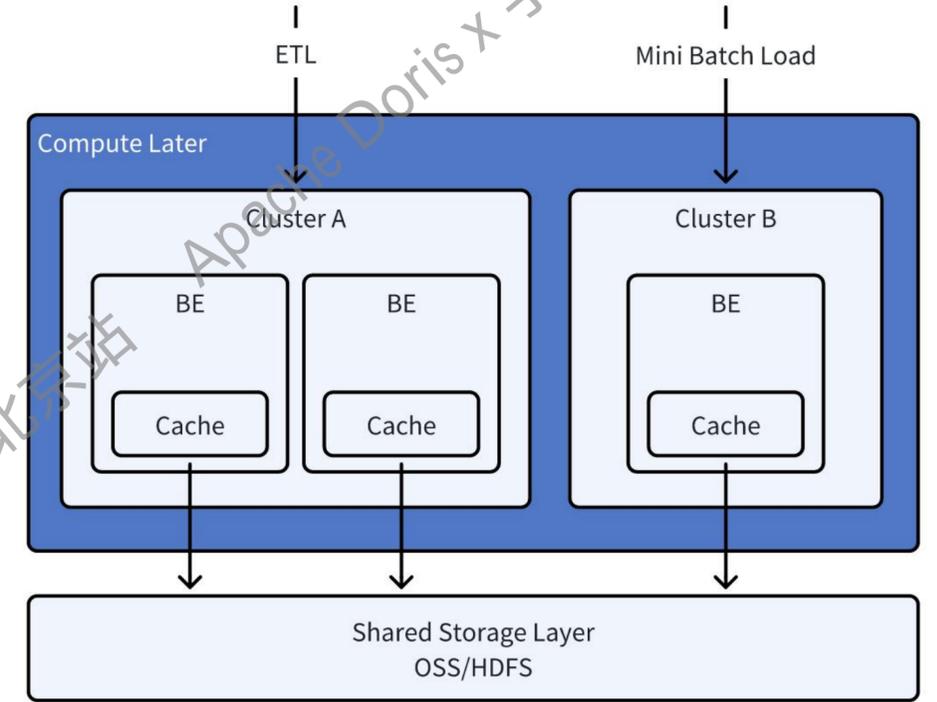
- 高并发点差和即席分析



- **读写隔离**

- 实时同步

- 自动预热



- **写写隔离**

- 高频导入和ETL

目录

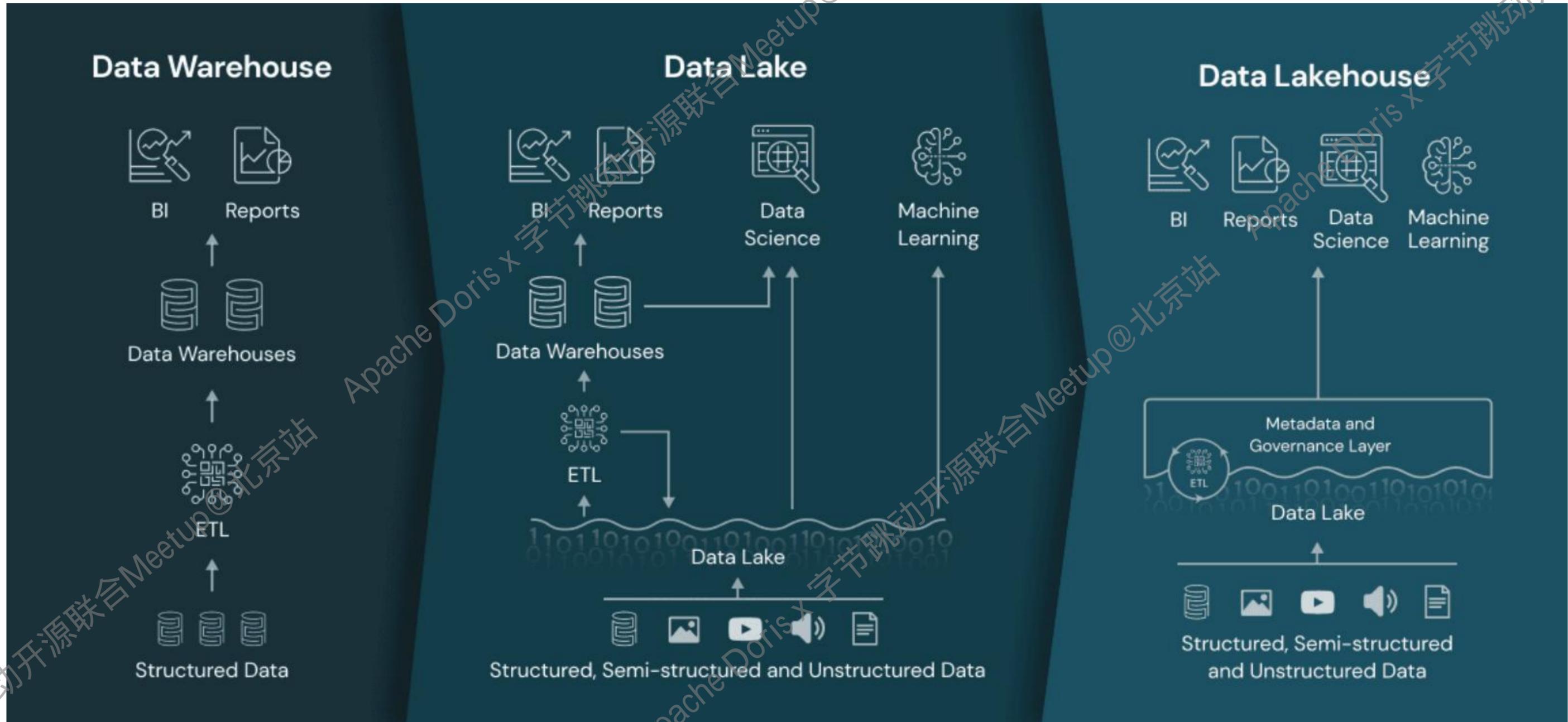
01 存算一体 or 存算分离

02 如何设计面向未来的架构

03 湖仓一体再进化

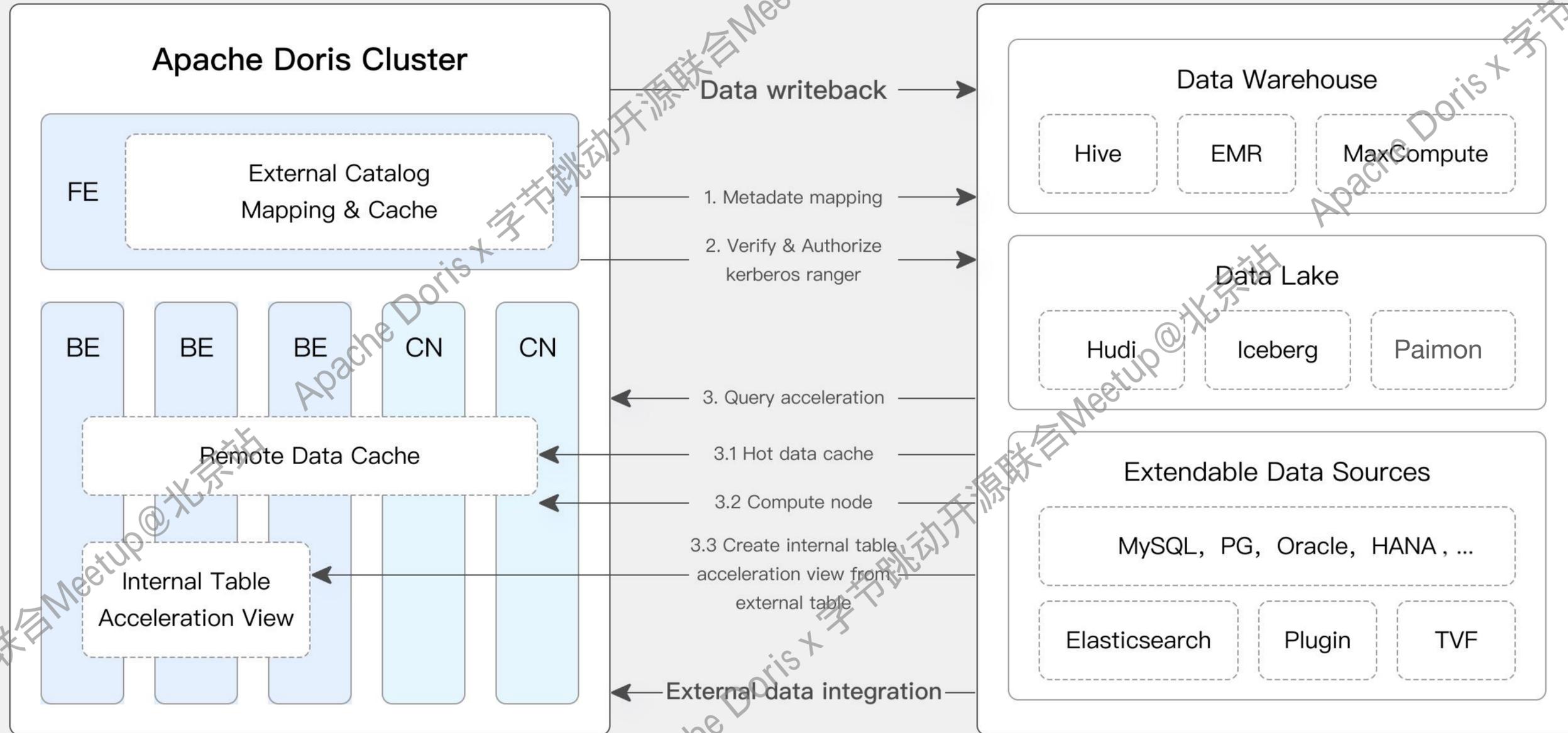
04 社区发展规划

从 Lakehouse 概念起源说起

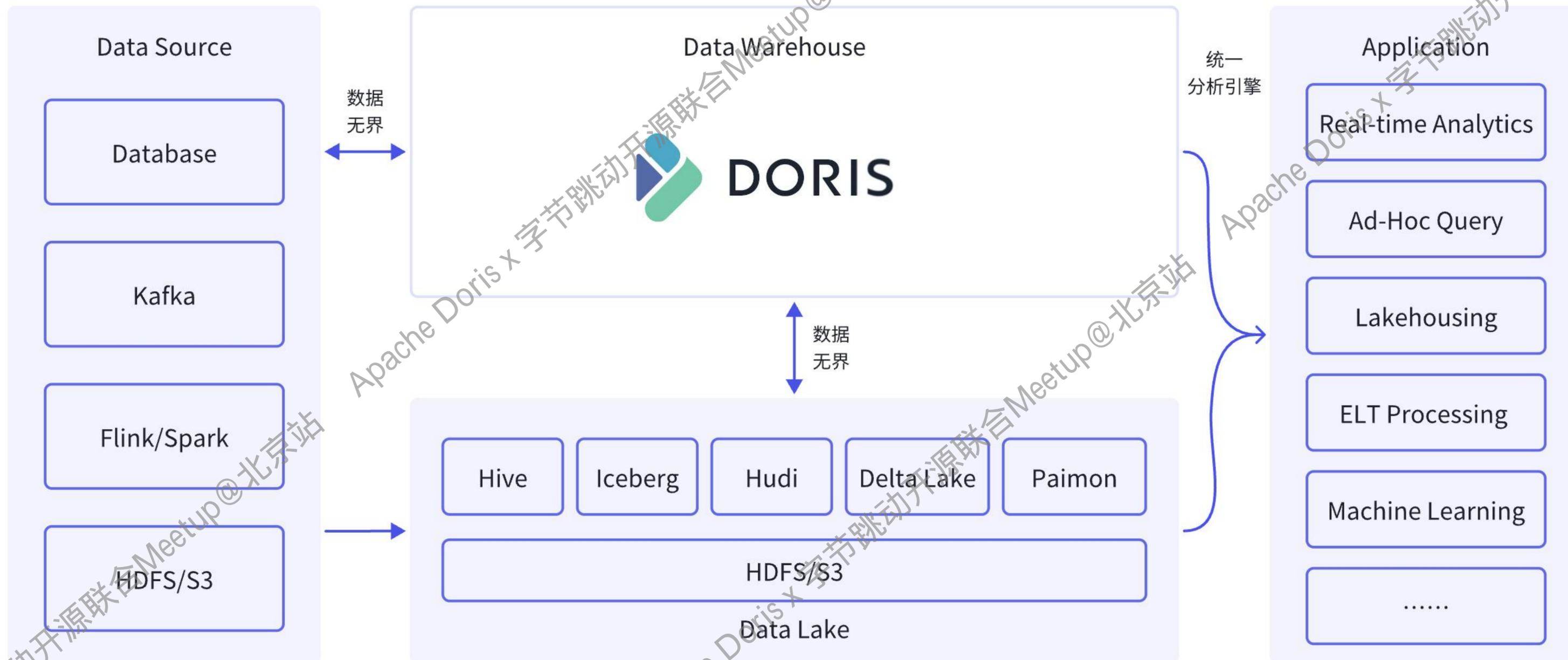


Lakehouse 诞生的出发点是将数据湖的低成本、高扩展性与数据仓库的高性能相融合，实现对各类数据的高效、安全、质量可控的存储和处理分析。本质是一种融合架构的解决方案，而非具体的单一产品。

Apache Doris Lakehouse 解决方案



数据无界 湖仓无界 - Lakehouse 构建之道

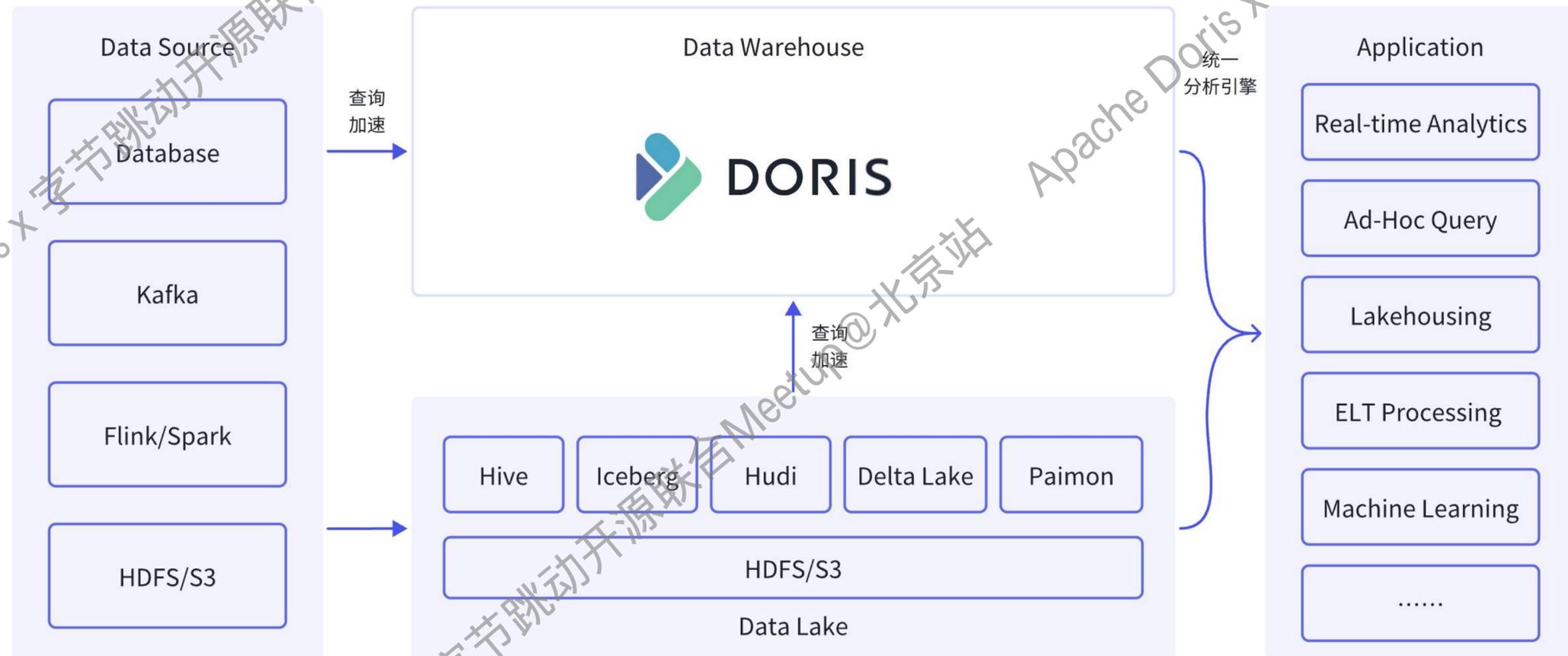


利用 Apache Doris 的高性能作为统一计算引擎，并打破数据在不同系统间的屏障，在数据仓库、数据湖乃至数据流、本地数据文件等所有数据源端都能提供一致性的极速分析处理体验。

场景一：极致分析性能、助力湖仓查询加速

核心技术：

- 高性能查询执行引擎
- 丰富的数据源连接能力
- 外部数据源读取优化
- 弹性计算节点 Compute Node
- 物化视图和透明改写

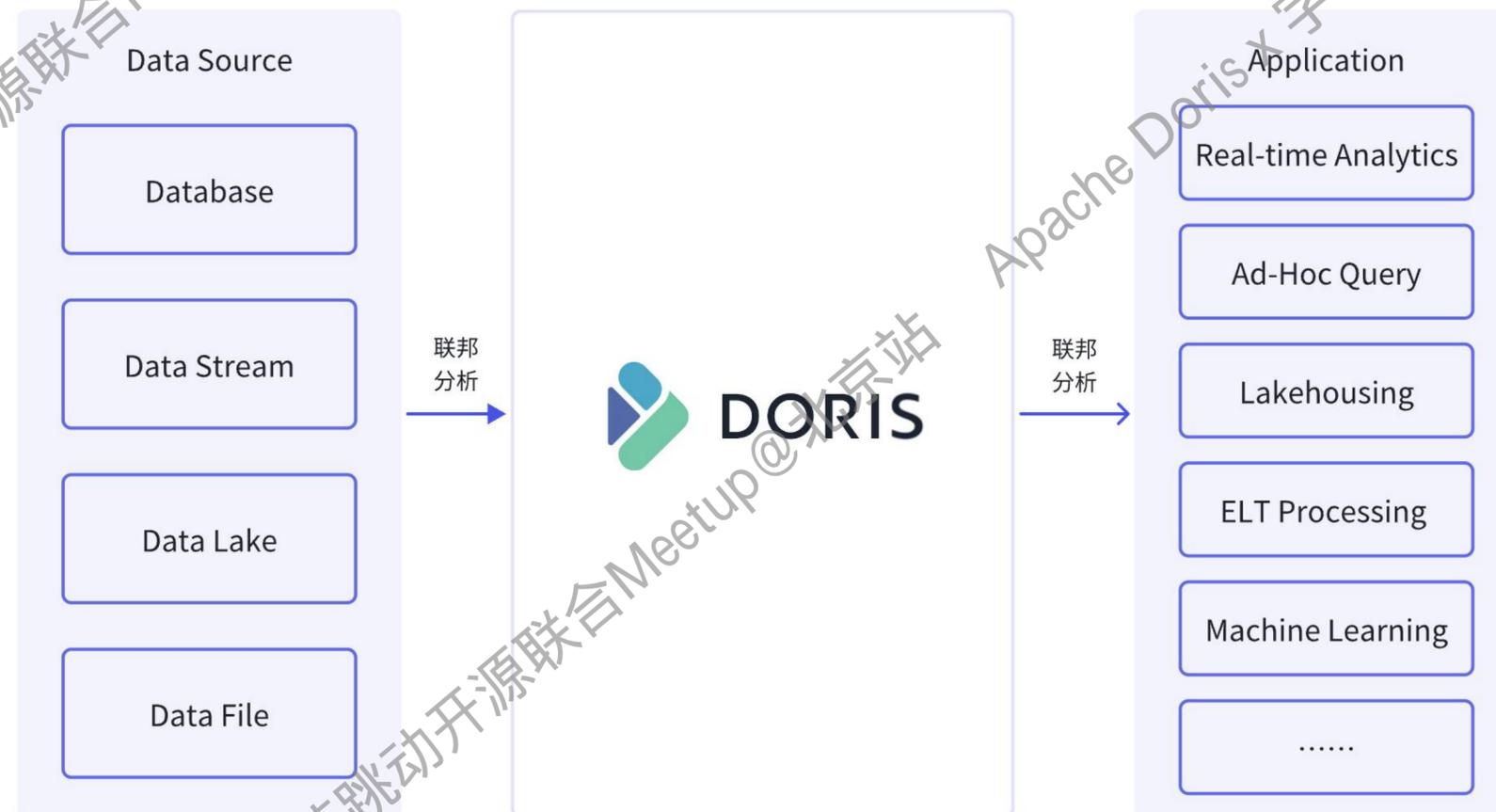


相较于 Trino/Presto, Apache Doris 的查询性能提升超过 400%, 相较于 Spark 性能提升接近 10 倍。基于 Apache Doris 优异的数据湖分析性能, Apache Doris 已成为业界领先的湖仓查询加速引擎。

场景二：多源联邦分析，消除数据孤岛

核心技术：

- Multi-Catalog 多源数据目录
- 引入 TVF 表值函数，支持直接分析 S3/HDFS 上的数据文件
- Trino Connector 扩展了 Apache Doris 可连接的数据源范围。

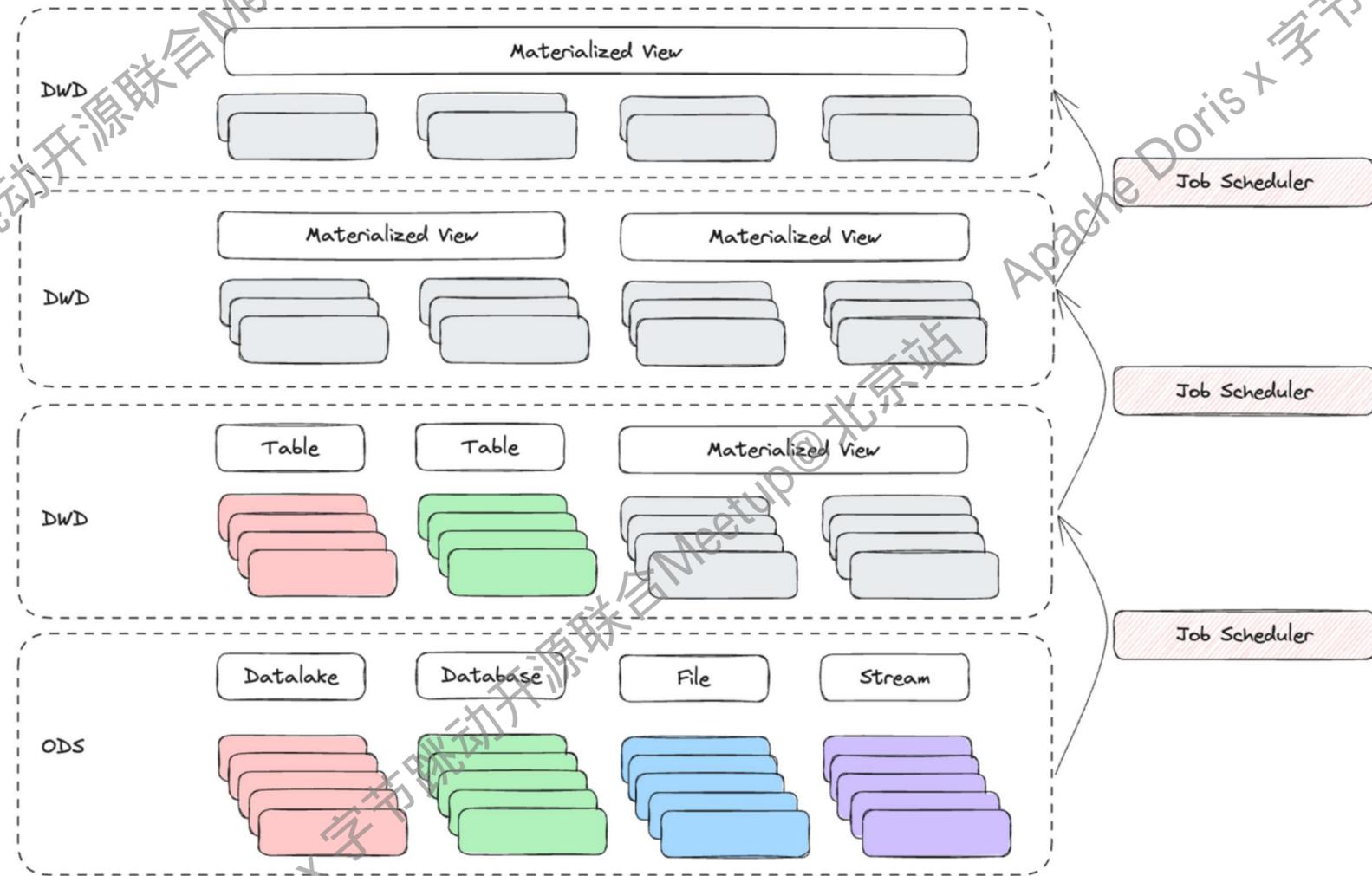


通过 Apache Doris 即可实现多个异构数据源的统一分析，既可以直查外部表以及存储文件、也可以执行内表和外表之间以及外表与外表之间的关联分析，消除数据孤岛、提供全局一致性的数据洞察分析。

场景三：湖仓数据无缝集成、自由流转

核心技术：

- 引入 Job Scheduler 功能，不依赖外部调度组件即可实现外部数据的定期自动同步，并能够灵活调整导入任务的调度策略
- 引入了多表物化视图，可在数据湖之上构建物化视图，将外表查询结果保存在数据仓库内部，实现数据湖与数据仓库之间的无缝集成。

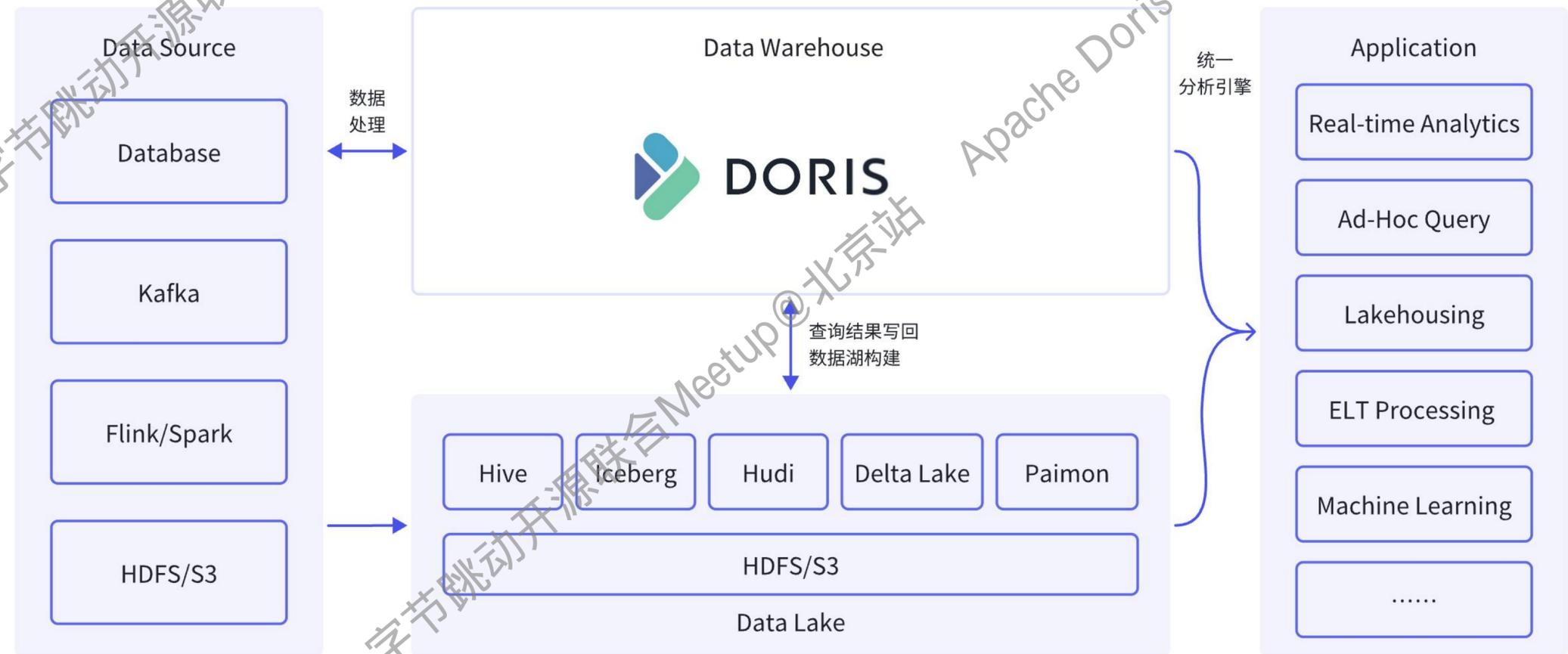


从数据湖到数据仓库的数据集成、业务建模、ETL 任务编排可以最大化减少外部组件依赖、仅凭 SQL 就可以维护整套数据加工逻辑，数据与数据之间的调度凭借系统自身即可运转，真正实现了数据湖与数据仓库之间的无缝集成和自由流转。

场景三：湖仓数据无缝集成、自由流转

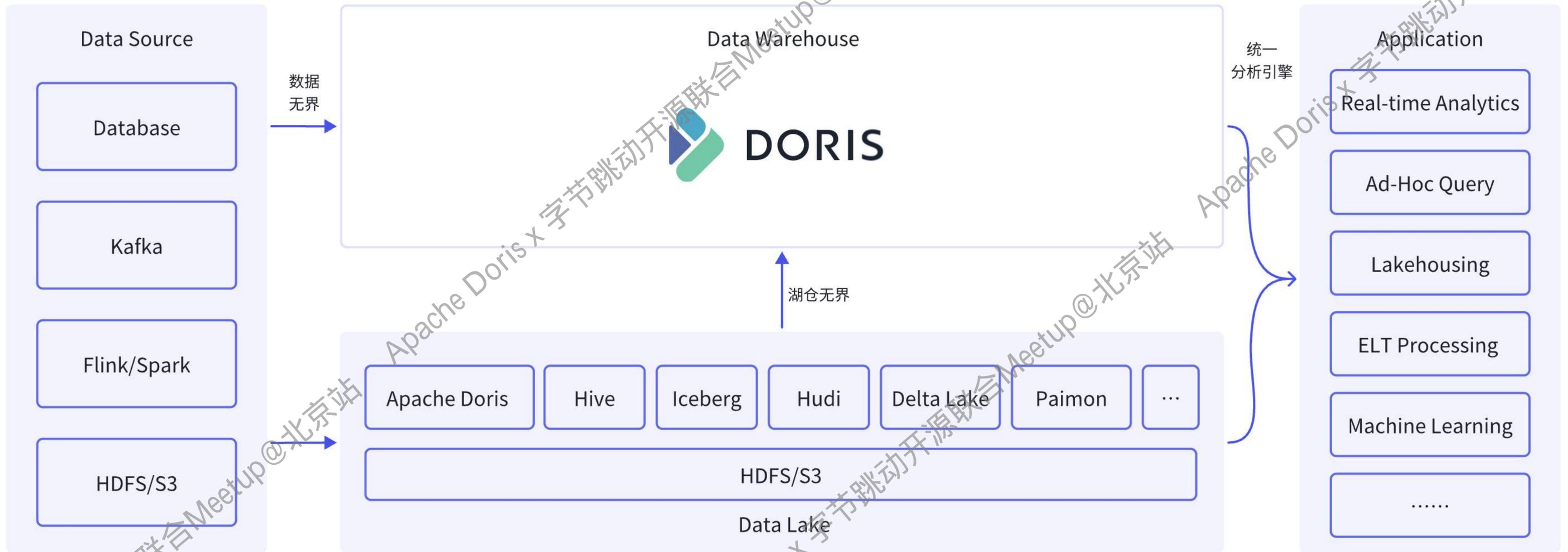
核心技术：

- 多 SQL 方言支持，直接使用对应的 SQL 方言查询 Doris 中的数据；
- 支持数据湖写回，实现多个数据源之间的数据分析、共享、处理、存储操作。
- 中间结果落盘（实验性功能）



结合新版本中的中间结果落盘、异步物化视图等能力，Apache Doris 可以作为企业统一的数据湖构建和计算引擎，简化系统架构的同时实现更高效的数据处理能力，帮助用户更好的管理湖、仓、数据库中的数据。

构建 Lakehouse 的最后一块拼图 - 存算分离



在实现存算分离后，Apache Doris 计算资源按实际负载灵活弹性扩缩容、全量数据存储到成本更低且极其可靠的共享存储中、存储成本最高下降至原先的 1/10，解决了传统数据仓库扩展性有限和存储成本较高的问题，支持在数据湖之上构建更简易的湖仓架构。

3.0 更多特性

- Runtime Filter 自适应能力增强
- 查询算子落盘
- 显式事务支持
- Routine Load 写入优化
- Variant 性能和易用性增强
- 倒排索引性能进一步提升

目录

01 存算一体 or 存算分离

02 如何设计面向未来的架构

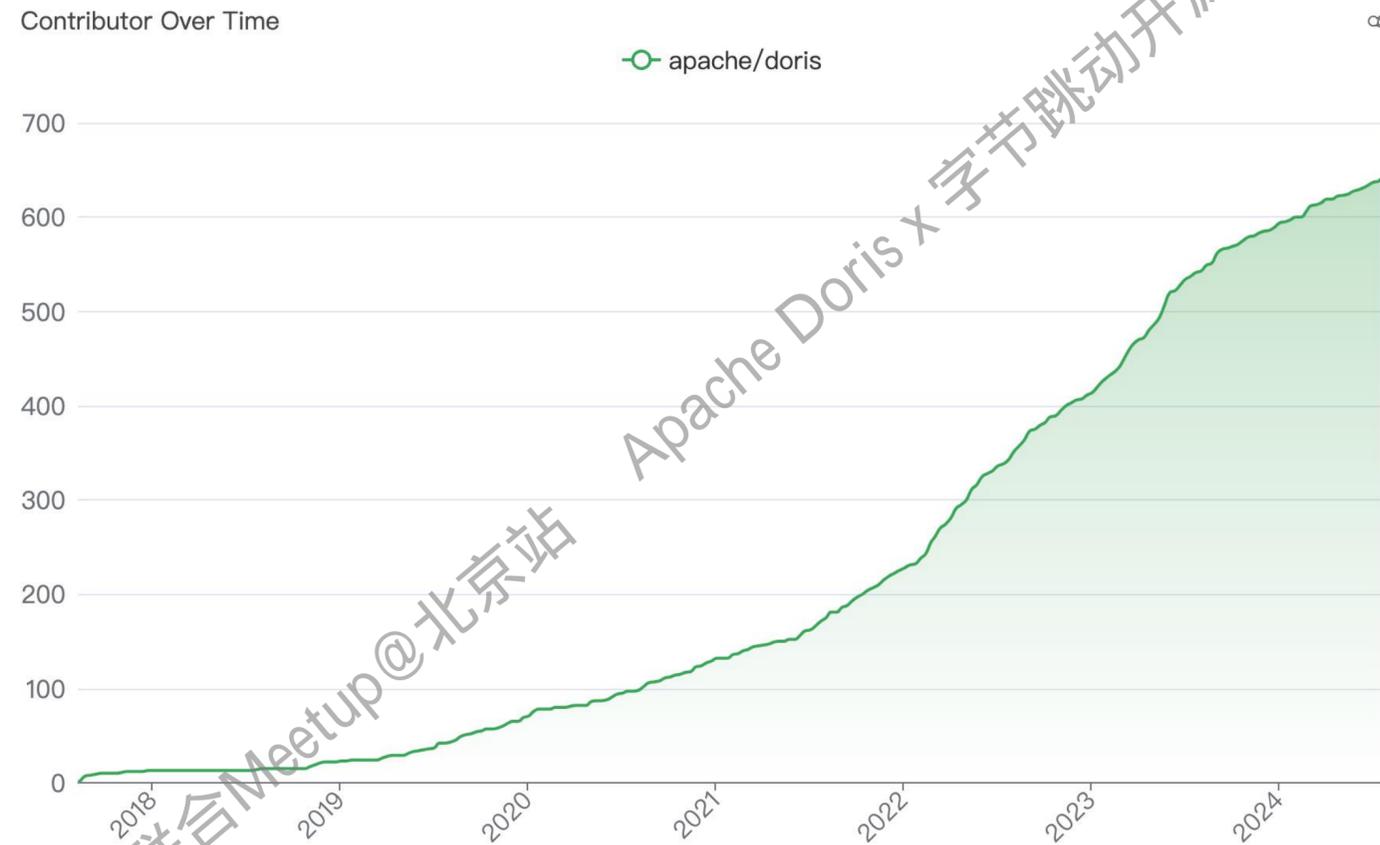
03 湖仓一体再进化

04 社区发展规划

全球大数据和数据库领域最活跃的开源社区之一

*统计时间: 截至2024年3月

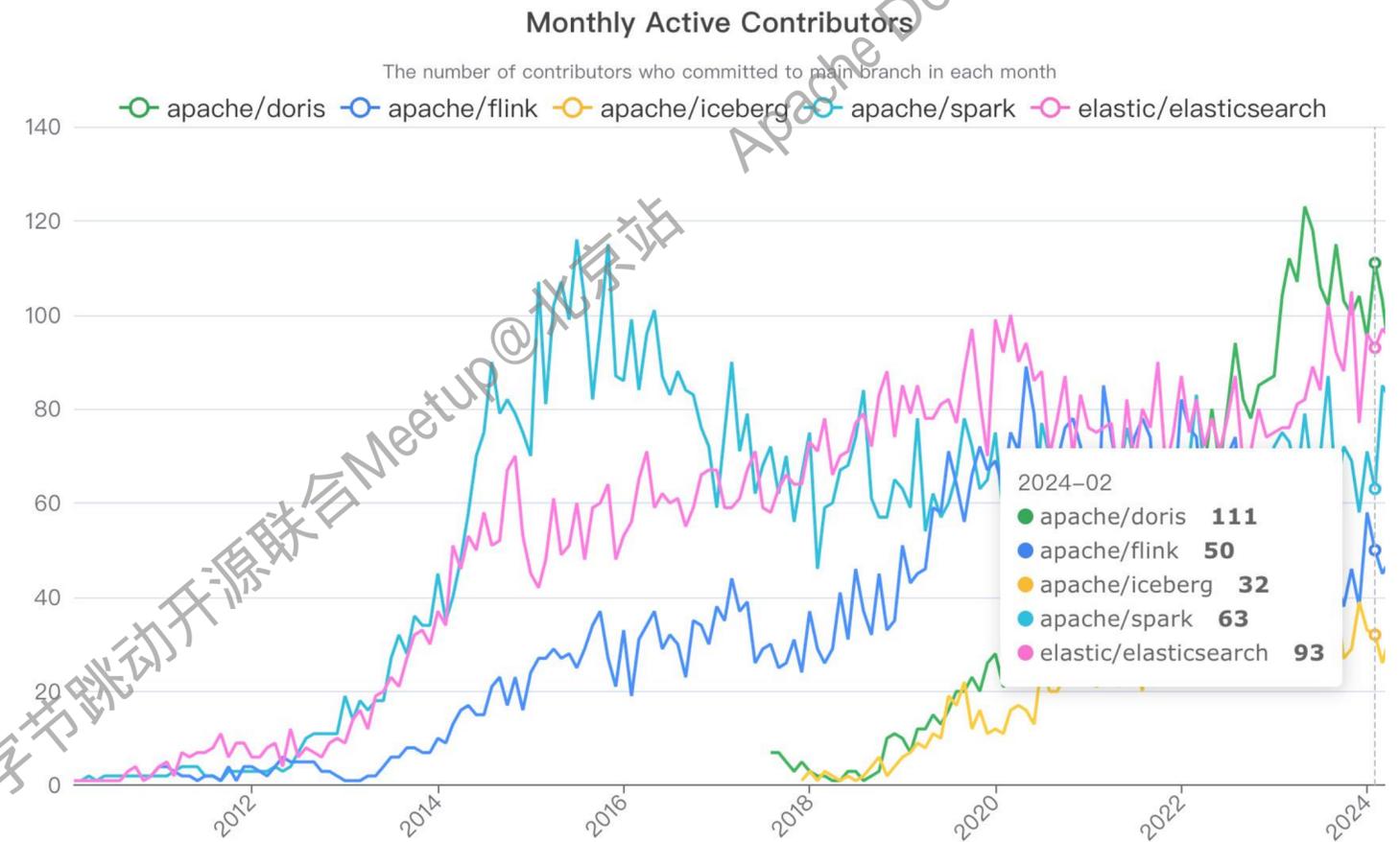
累计贡献者



650+

累计贡献者**已经超过650人**，距上一年度新增贡献者**超过150**，并仍处于持续上升的态势。

活跃贡献者



Top1

自2022年7月起至今，一直稳居在全球大数据开源项目排行中**活跃贡献者数Top1**

获得全球超过5000家中大型企业的信赖，广泛应用于核心线上分析场景

金融



互联网



电信



游戏



交通物流



零售快消



能源制造



加入 Apache Doris 社区任何时候都不晚

- 订阅开发者邮件组

订阅社区开发者邮件组 dev@doris.apache.org 并参与讨论

- Doris 小助手

如想进入社区用户社群请备注 **加群**



Doris 小助手



扫一扫上面的二维码图案，加我为朋友。

