

Apache Doris × 讯飞开源联合 Meetup

8月15日 (周四)

13:00-16:15

Hefei

关于 Apache Doris 城市行

Apache Doris 城市行是由飞轮科技发起、面向全国各地大数据、数据库以及实时数据仓库技术爱好者的线下技术交流活动。

Apache Doris 城市行旨在更好的为各地区技术爱好者分享一手的技术知识和实践经验，并搭建一个可以帮助技术爱好者与 Apache Doris 社区技术大咖线下交流、讨论的平台。同时，我们也希望通过线下面对面交互的形式，能够与各地区的社区成员产生紧密的连接，倾听社区成员们的想法，共同建设 Apache Doris 社区。

截止目前，已分别在北京、上海、深圳、杭州、武汉、成都、西安、广州等国内连续成功举办 10 余场次，吸引线上线下近 10w+ 国内技术爱好者的参与、活动曝光量超 100w+，是国内大数据领域参与人数最多、氛围最佳、内容最为丰富的技术交流活动之一。

往期活动展示



Apache Doris 3.0 云原生 存算分离架构的设计与实现

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



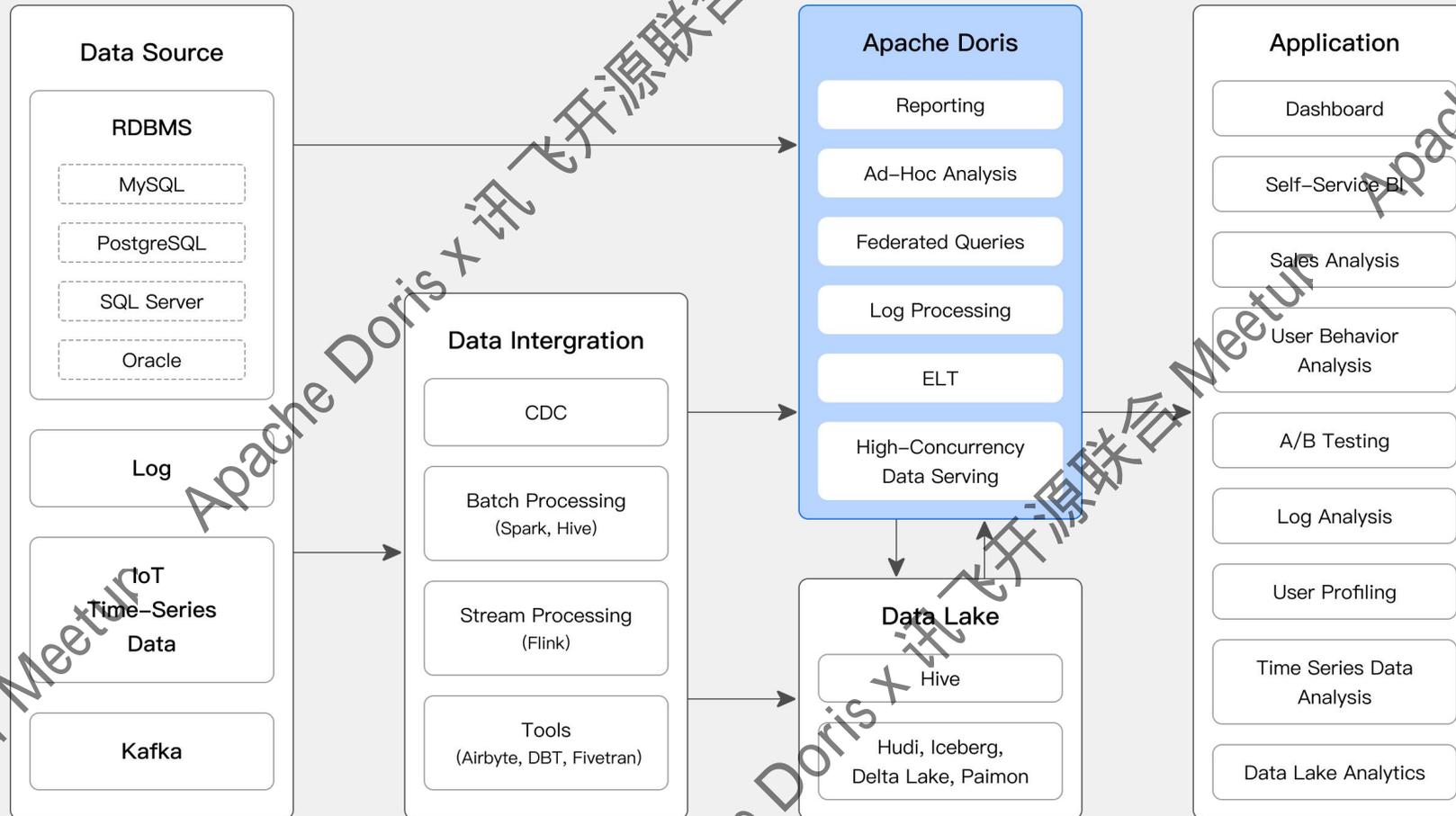
个人介绍



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

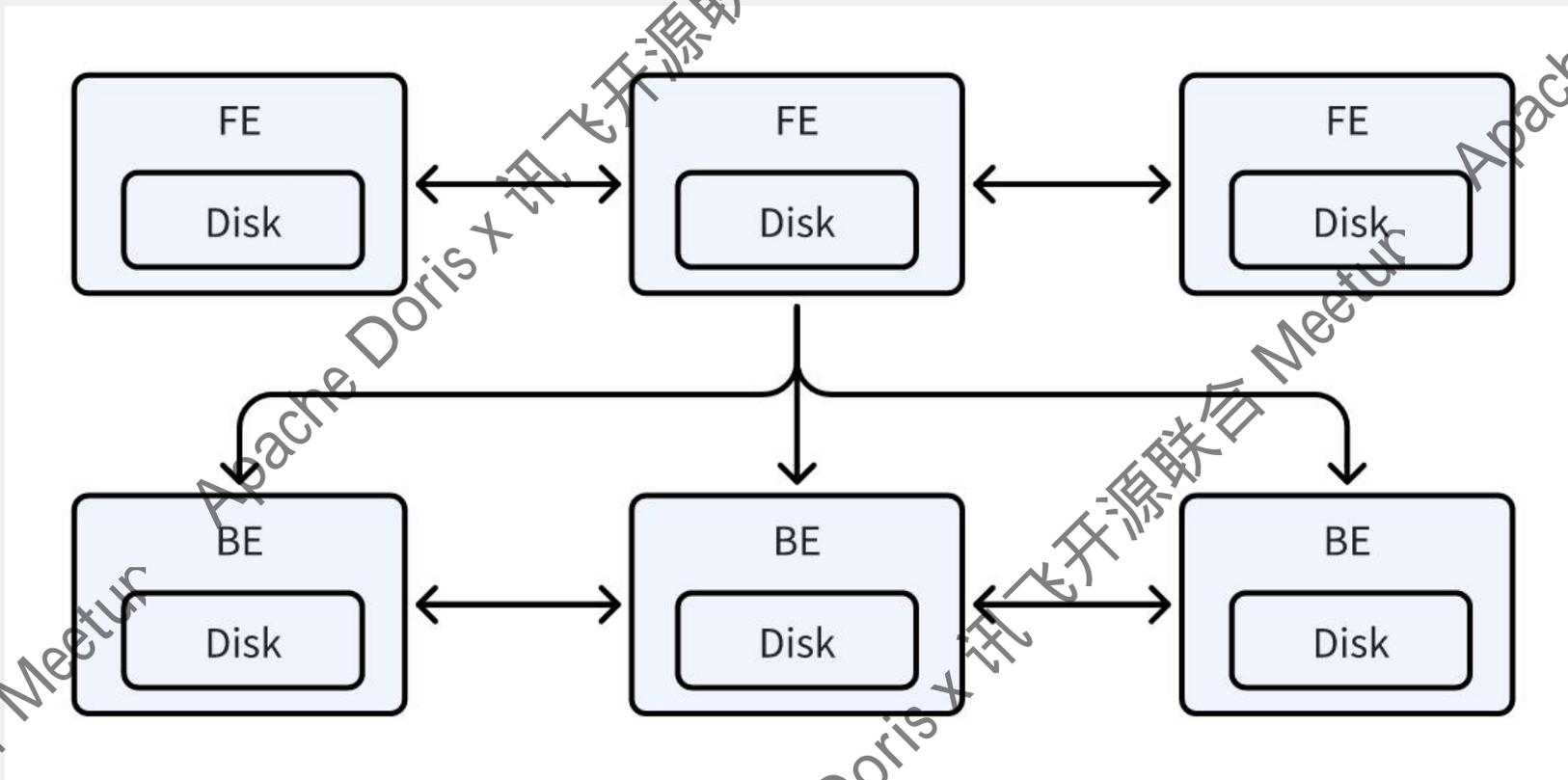
01 存算一体 or 存算分离

02 如何设计面向未来的架构

03 更多 3.0 版本特性揭秘

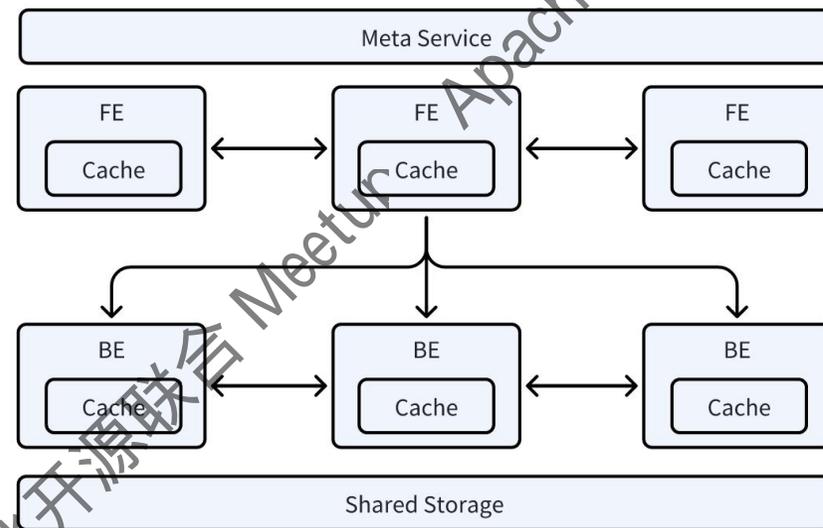
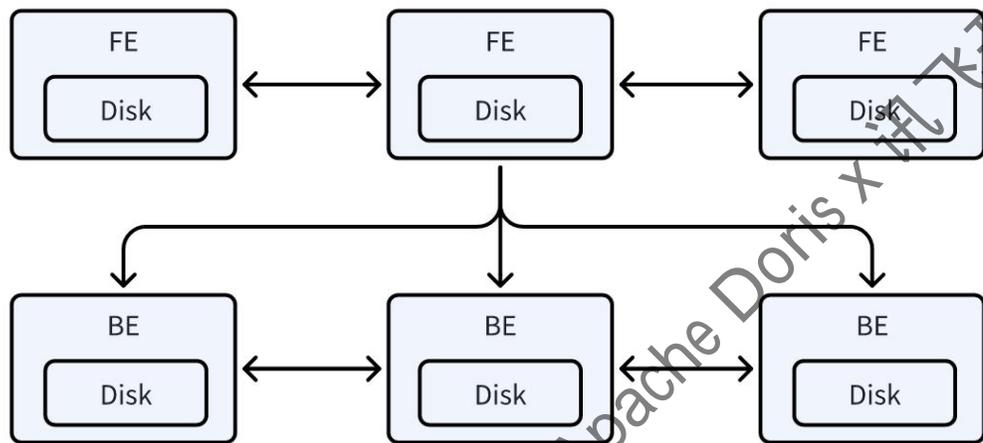
04 社区发展规划

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 更多 3.0 版本特性揭秘

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

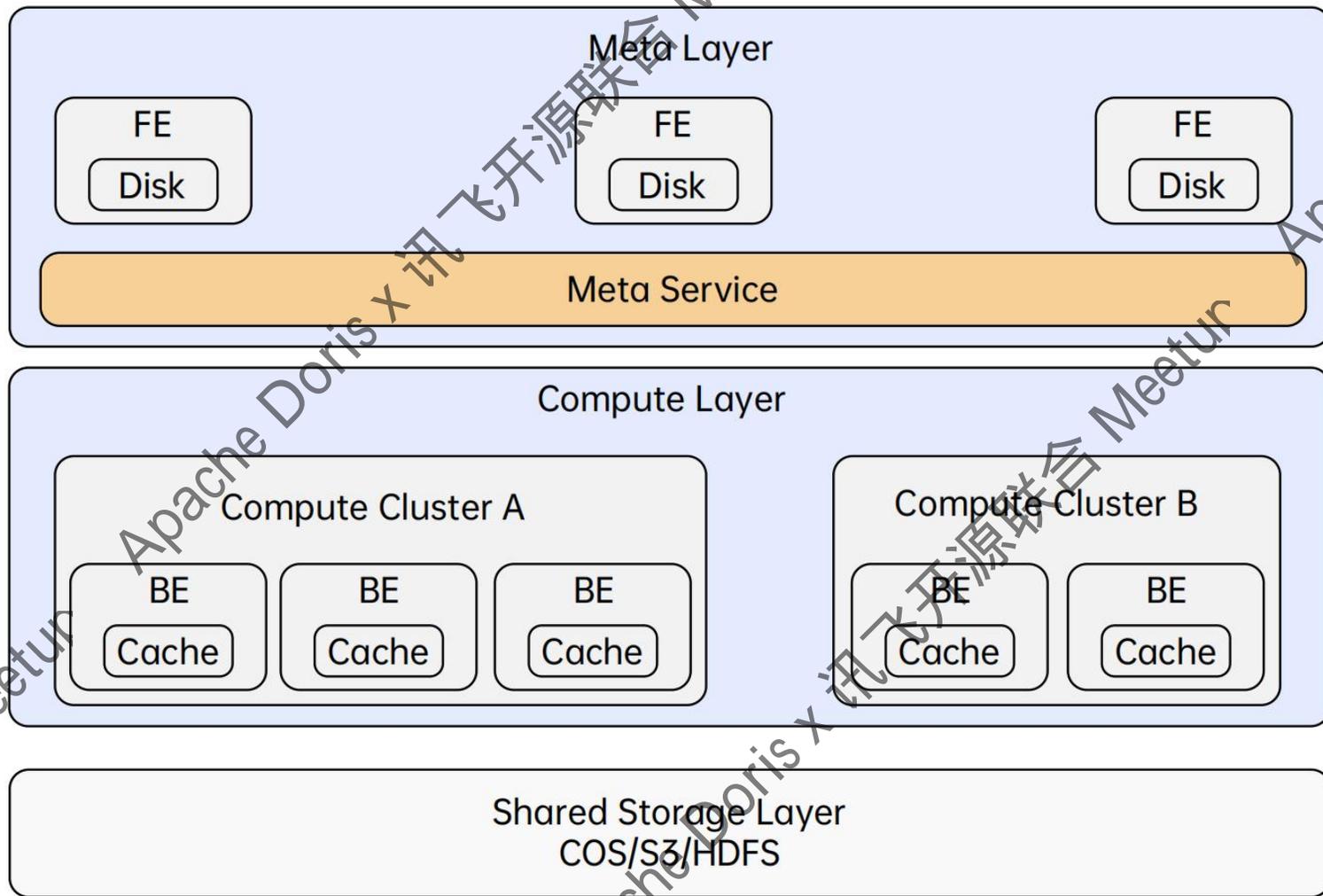


数据共享

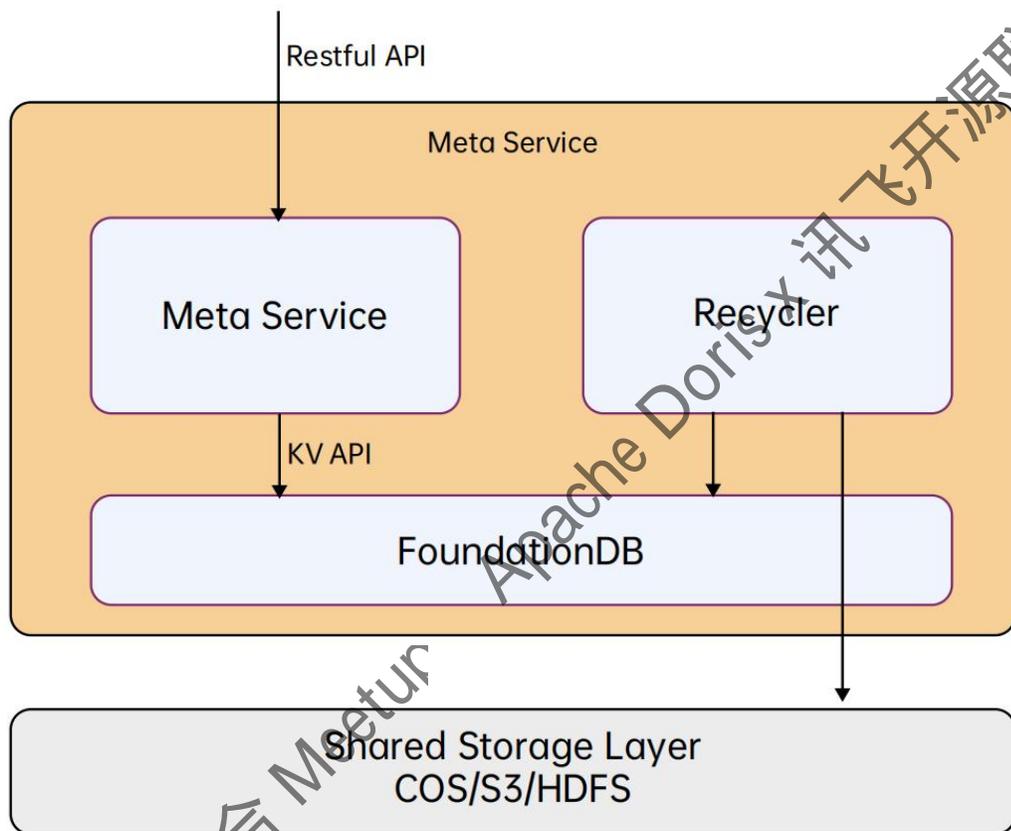
统一元数据服务

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

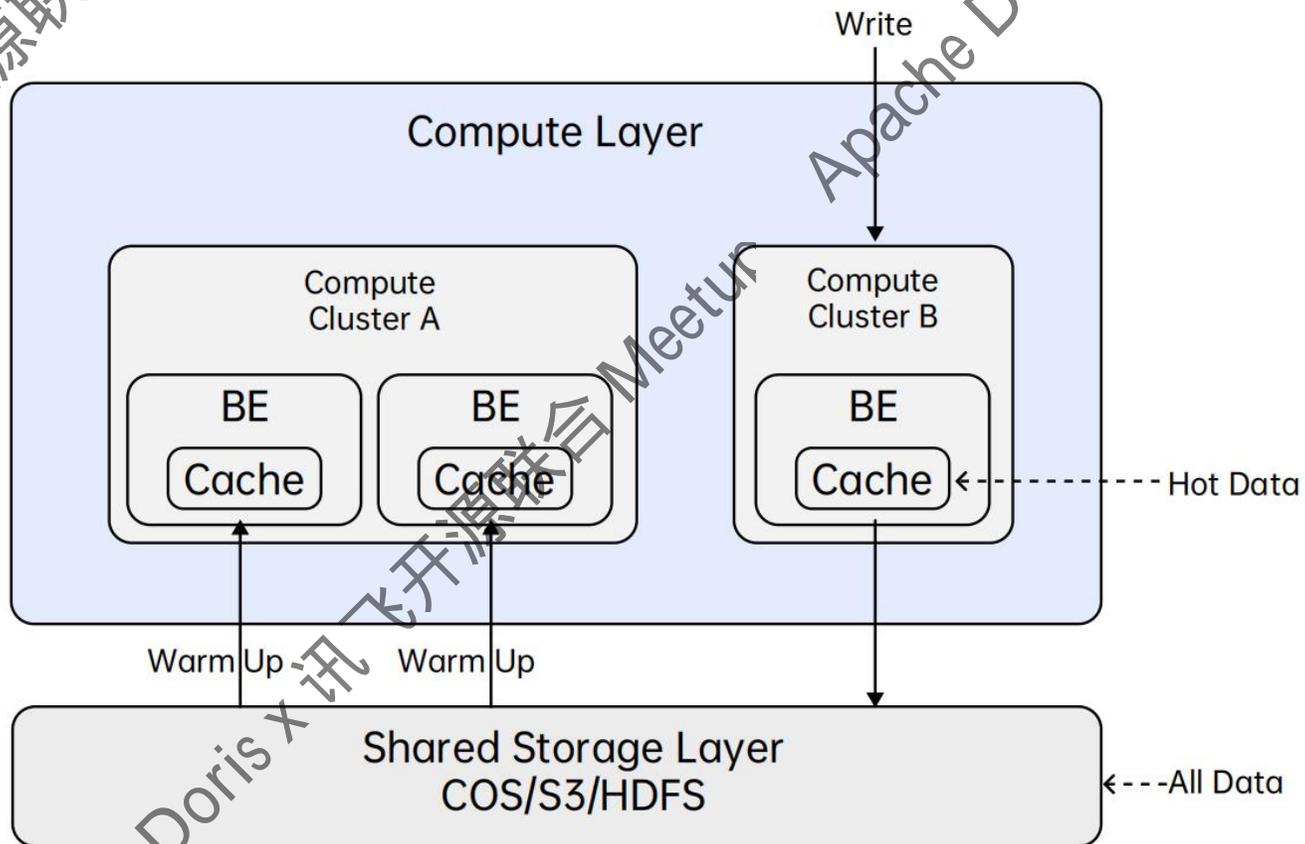
数据存储层

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
- 最高可以节省 90% 以上

灵活的 Cache 管理

- 多优先级淘汰策略：LRU、TTL
- 缓存主动预热：table 粒度、cluster 粒度、导入



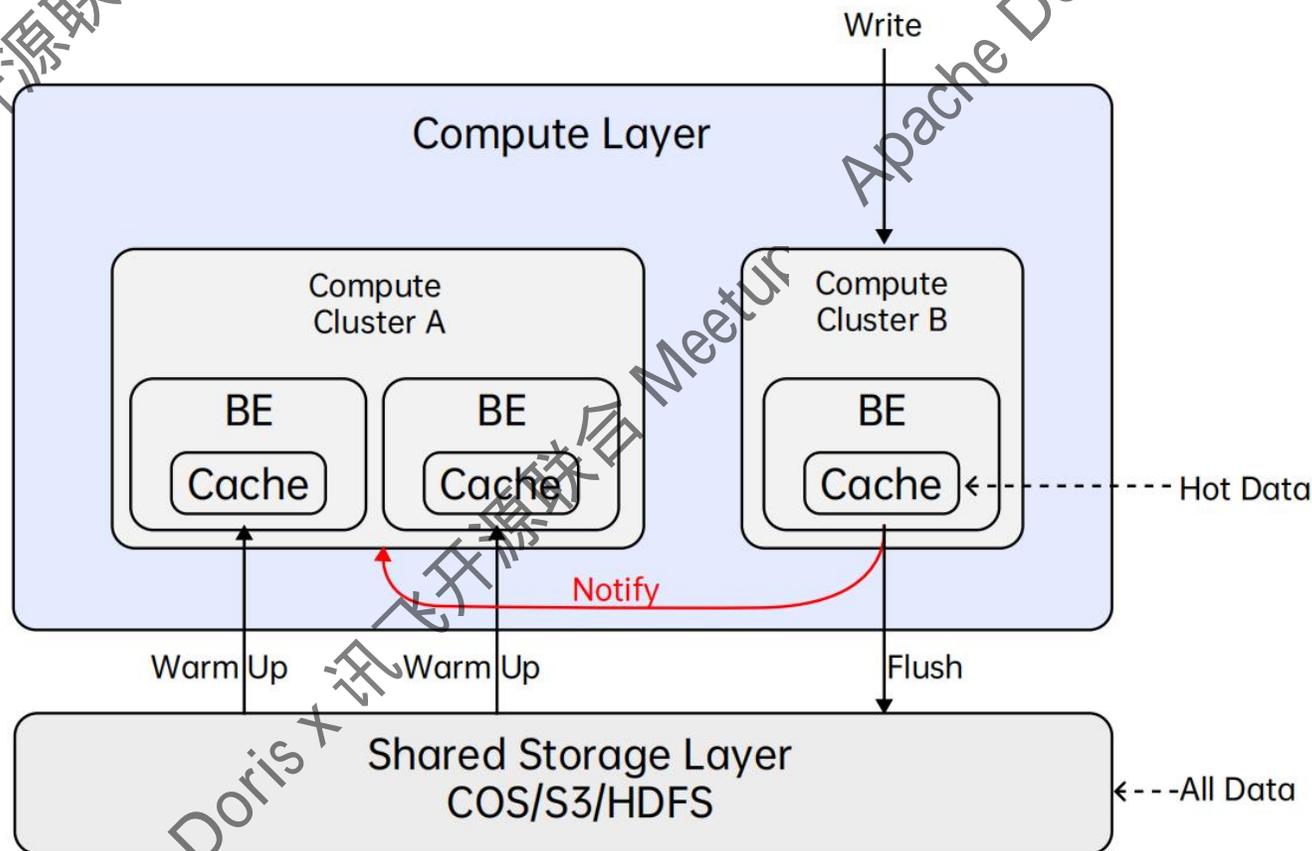
数据计算层-数据导入

数据导入流程

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

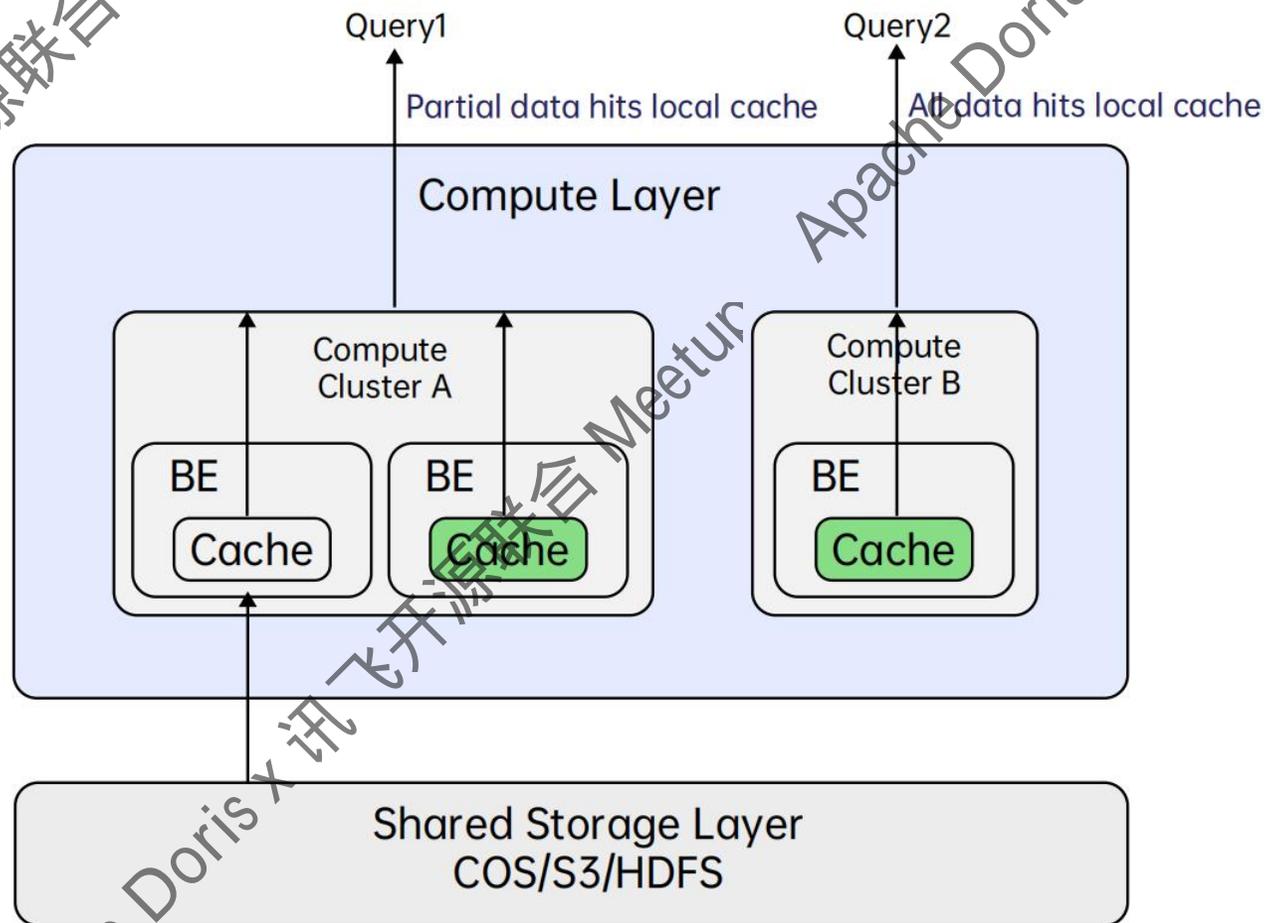
数据导入效率更高

- 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

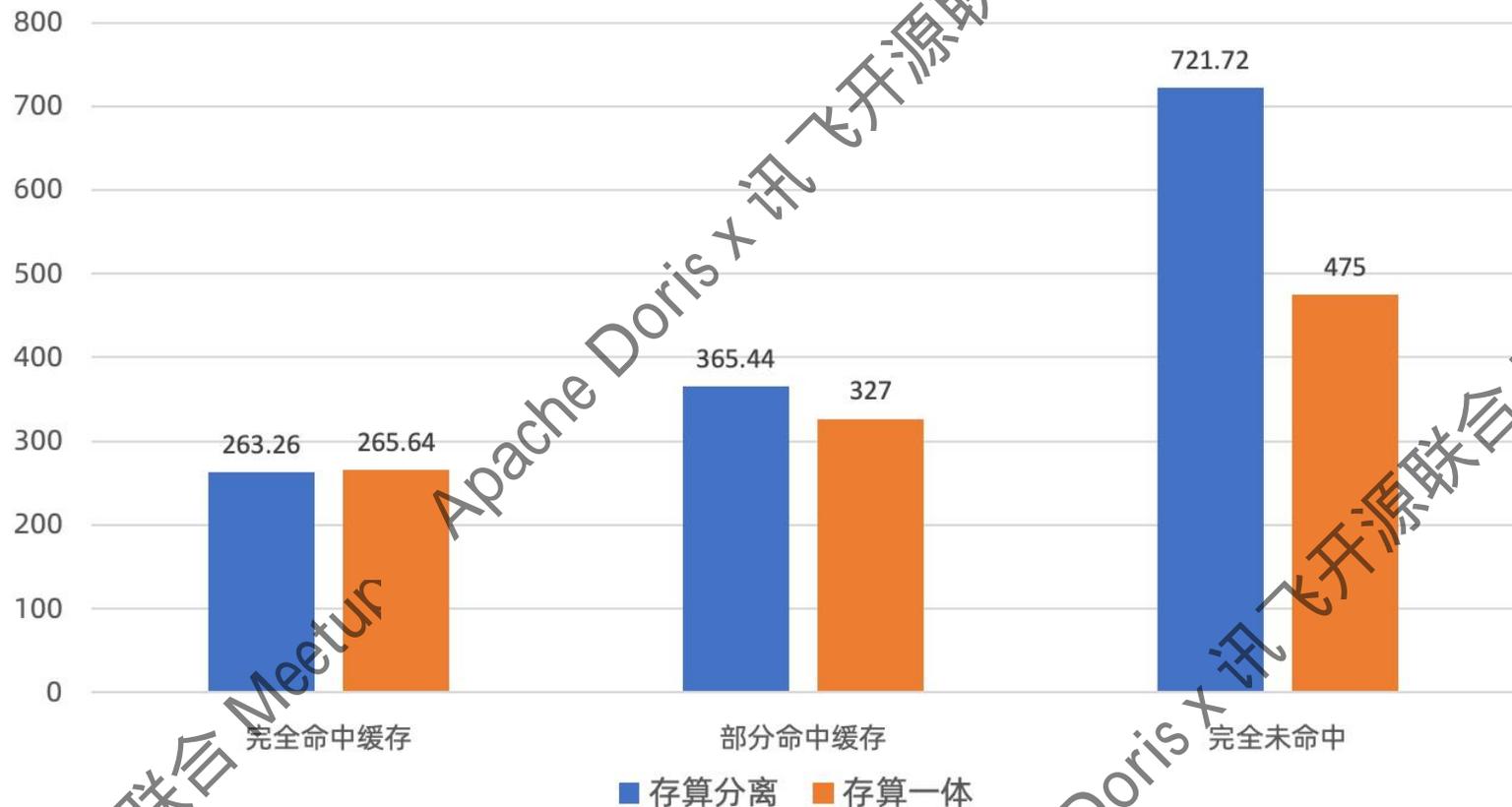
- 多个 Computer Cluster 独立
- 不命中 Cache 时，从 S3 读数据
- 命中时，从本地 Cache 读数据
- 弹性资源大幅降低成本



查询性能对比

单位 (s)

TPC-DS 1TB

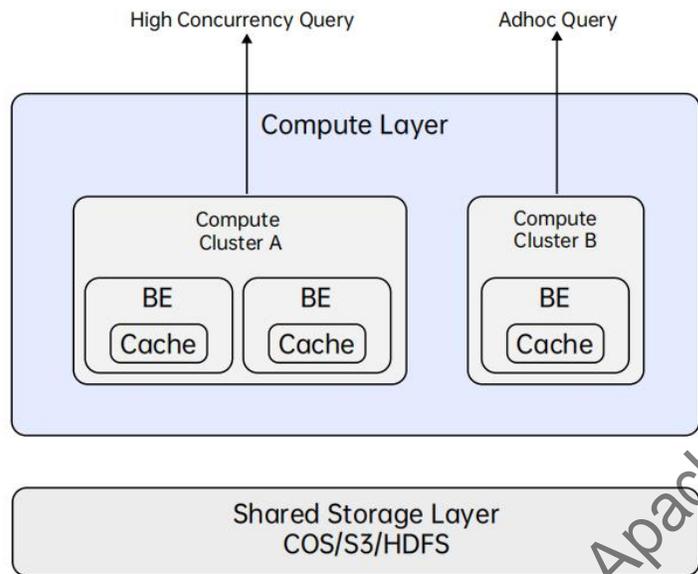


缓存层级

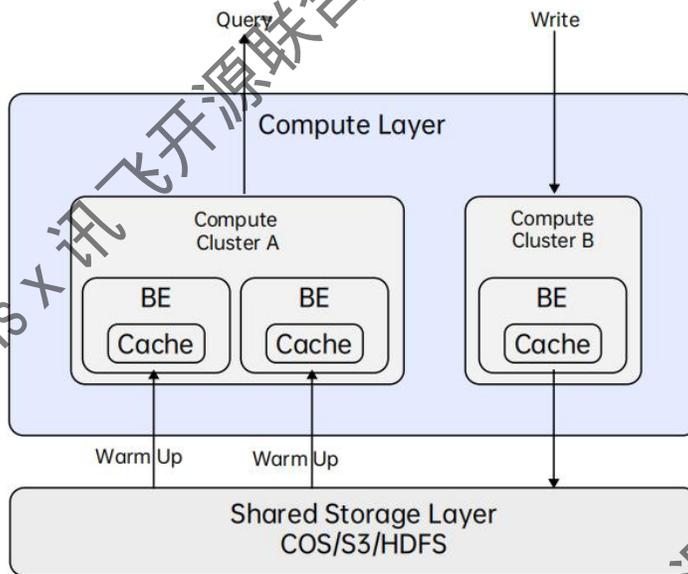
- Remote Storage
- Local File Cache
- System Page Cache
- Doris Page Cache

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%。

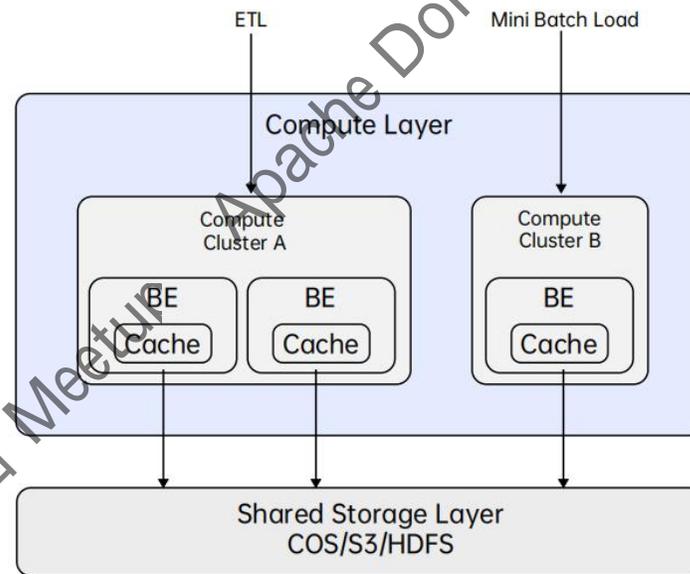
最佳实践



- **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

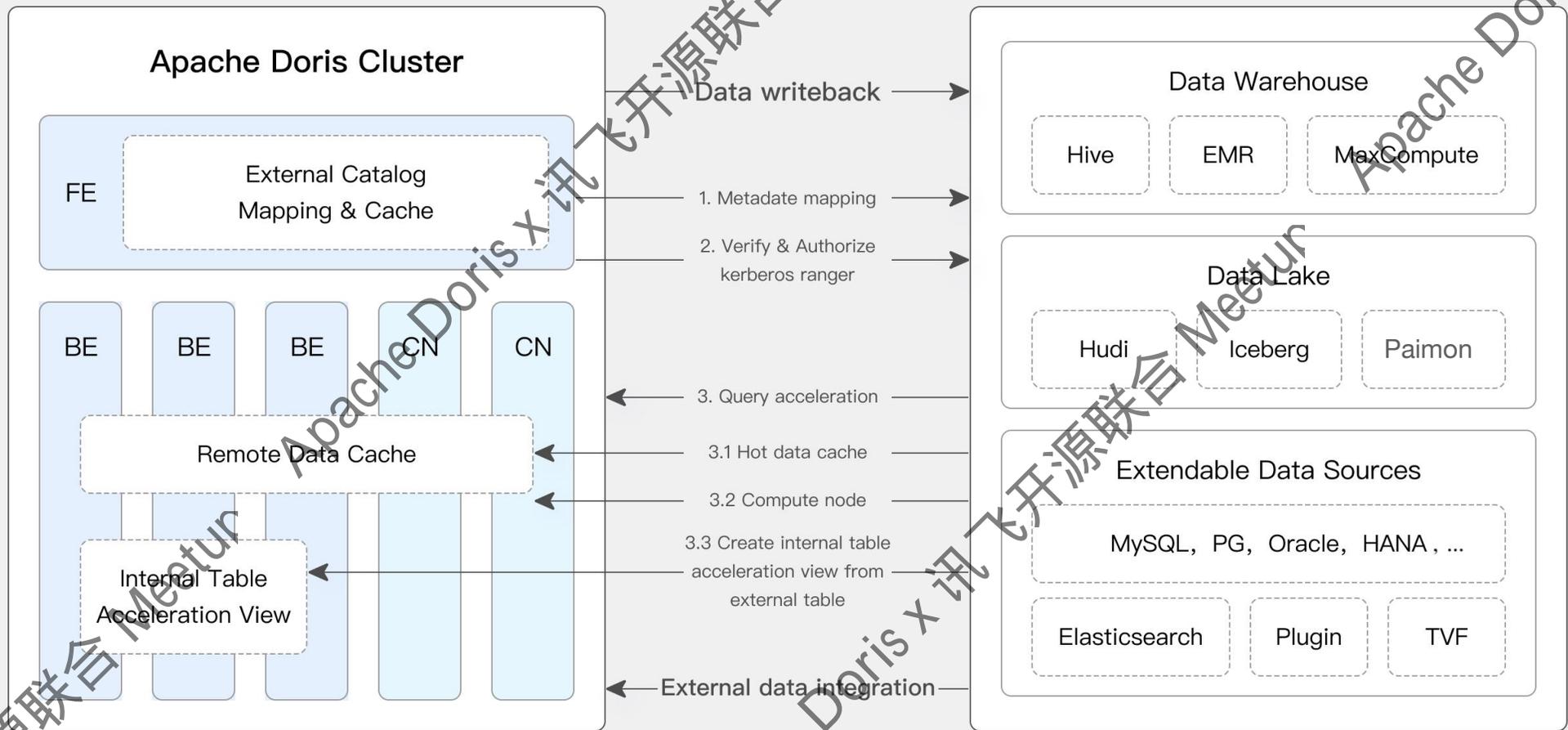
01 存算一体 or 存算分离

02 如何设计面向未来的架构

03 更多 3.0 版本特性揭秘

04 社区发展规划

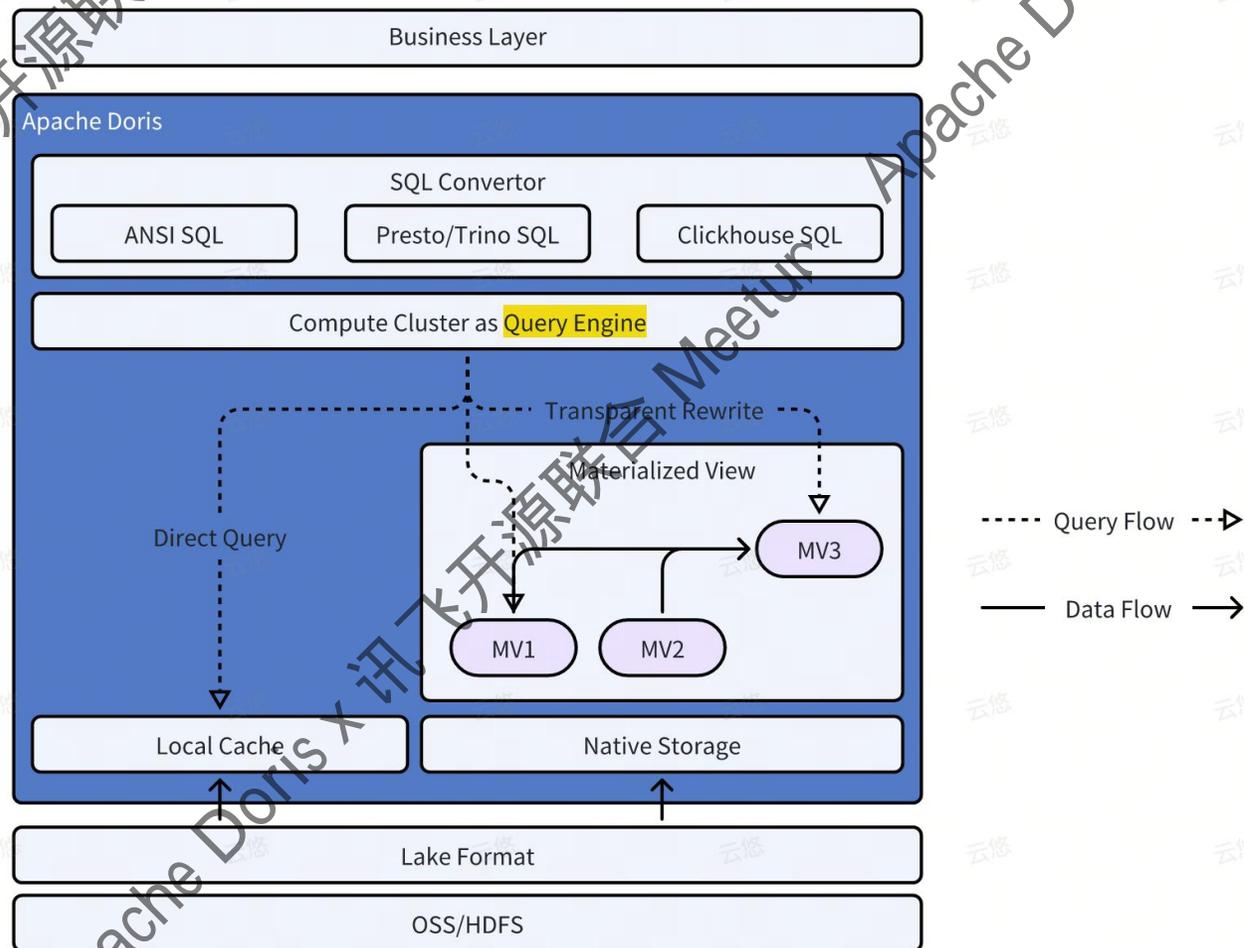
3.0 特性 - Lakehouse 再进化



3.0 特性 - Lakehouse 再进化

阶段1: 湖仓加速

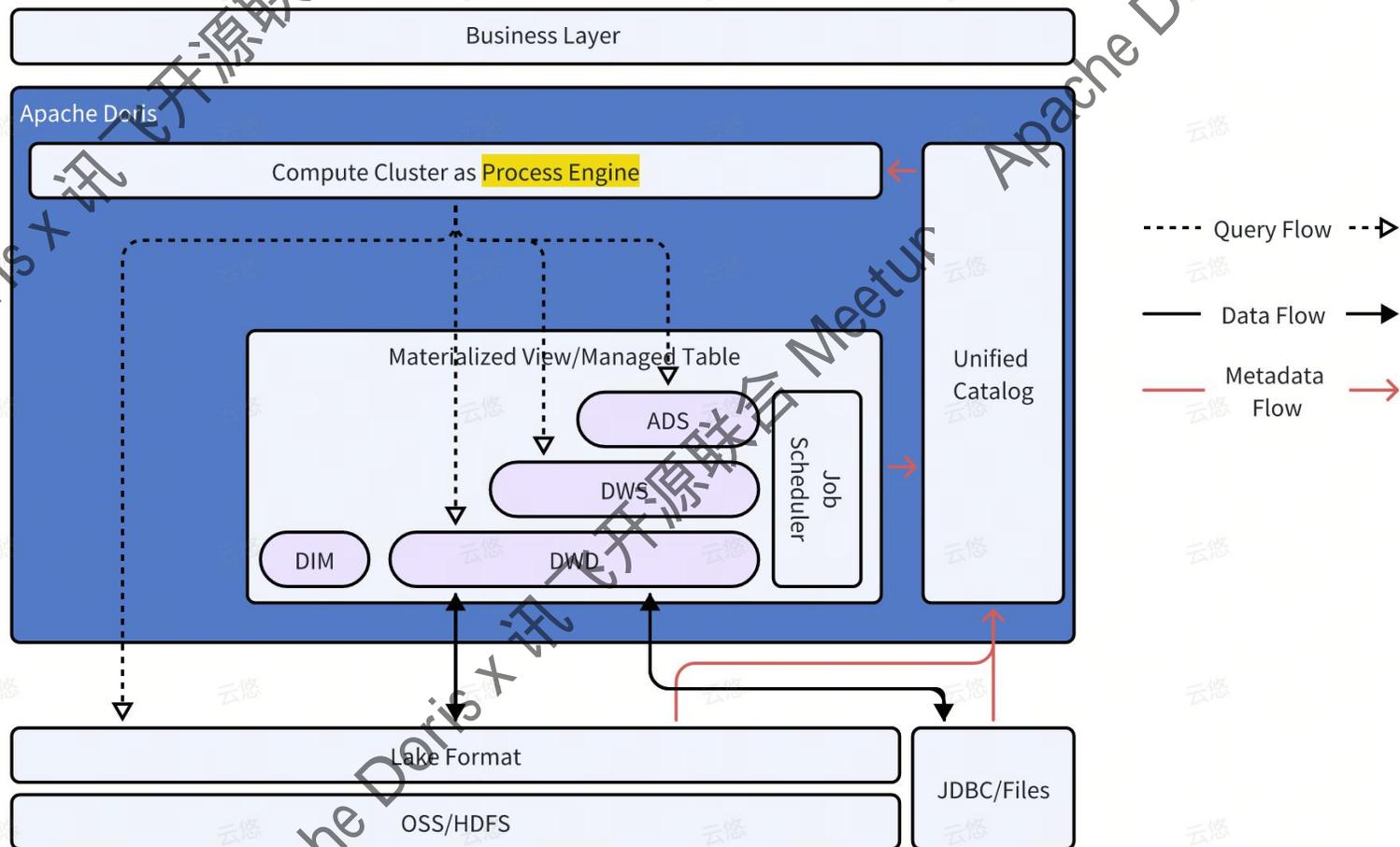
- 丰富的数据源连接
- 联邦分析能力和SQL方言兼容
- 物化视图和透明改写



3.0 特性 - Lakehouse 再进化

阶段2：湖仓数据处理

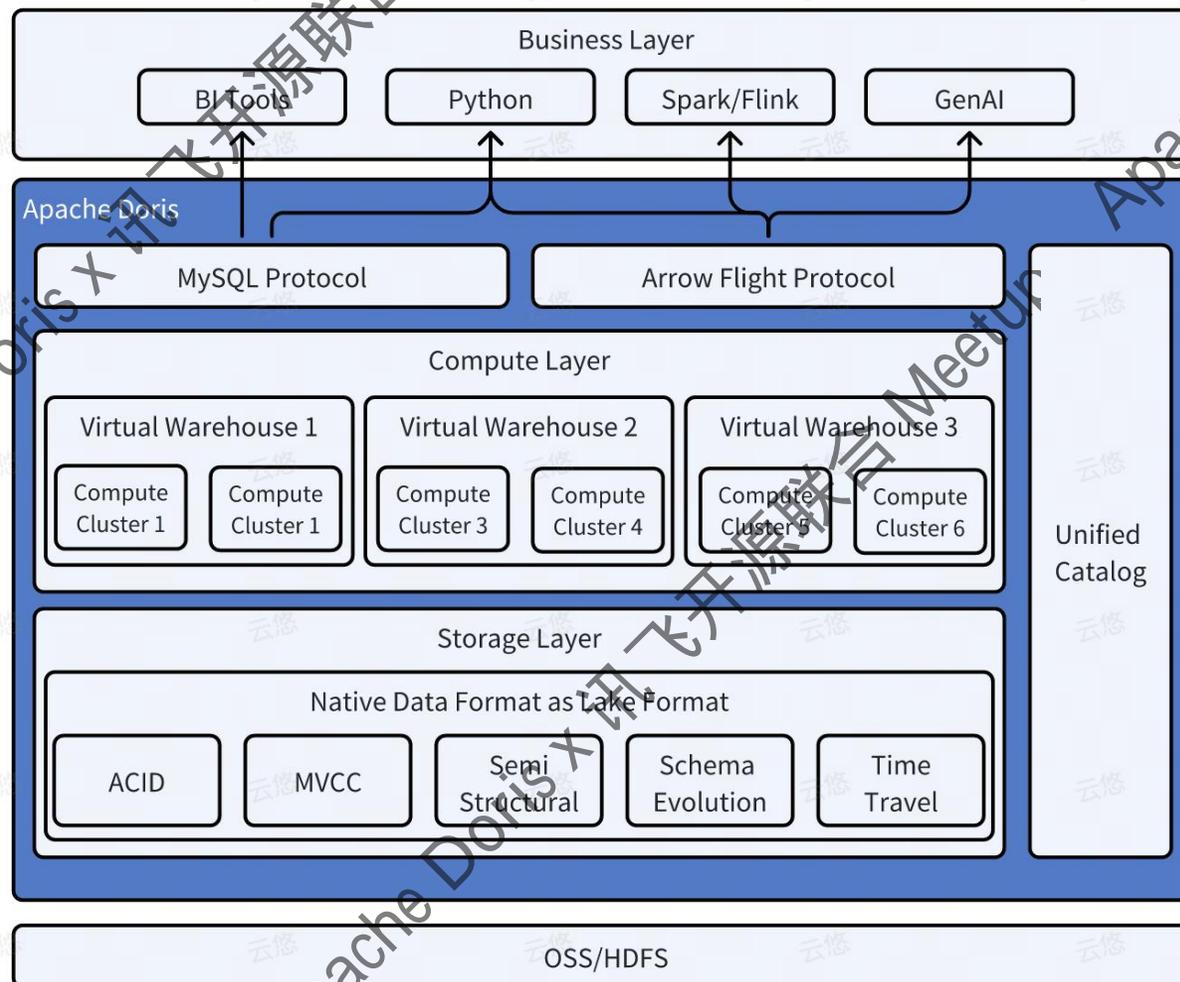
- 统一元数据管理
- 数据湖写回
- 物化视图和数据分层加工



3.0 特性 - Lakehouse 再进化

阶段3：湖仓一体

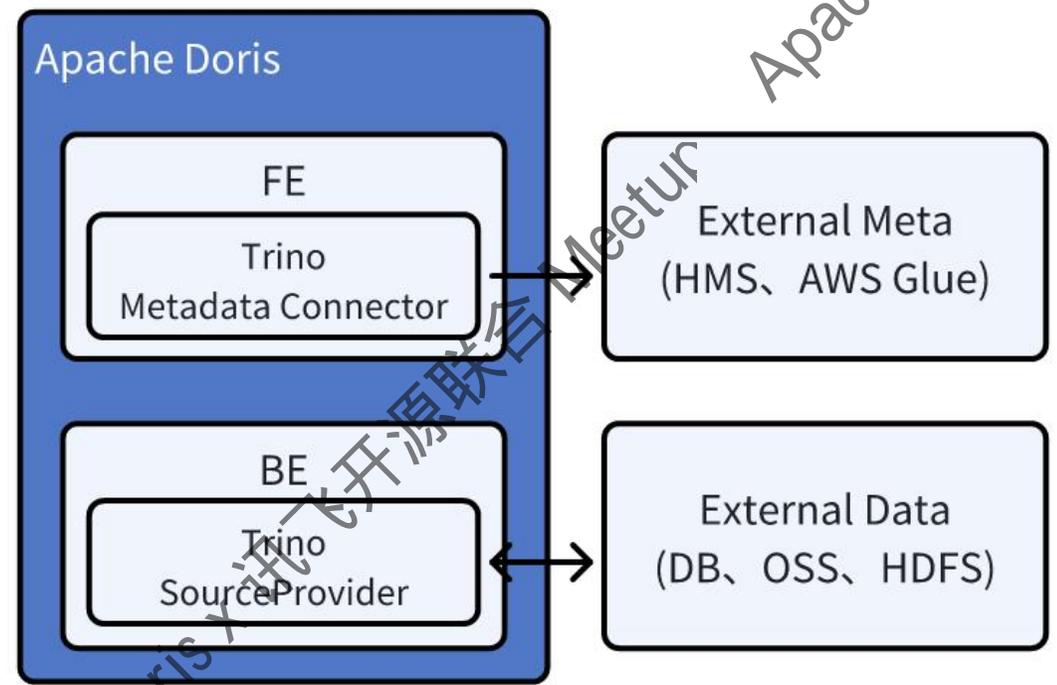
- 内外表特性的统一
- 存算分离架构



3.0 特性 - Trino Connector 兼容

即插即用、快速对接

- DeltaLake
- BigQuery
- Kudu
- Redis
- Kafka
- ...



参考文档: <https://doris.apache.org/community/how-to-contribute/trino-connector-developer-guide>

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

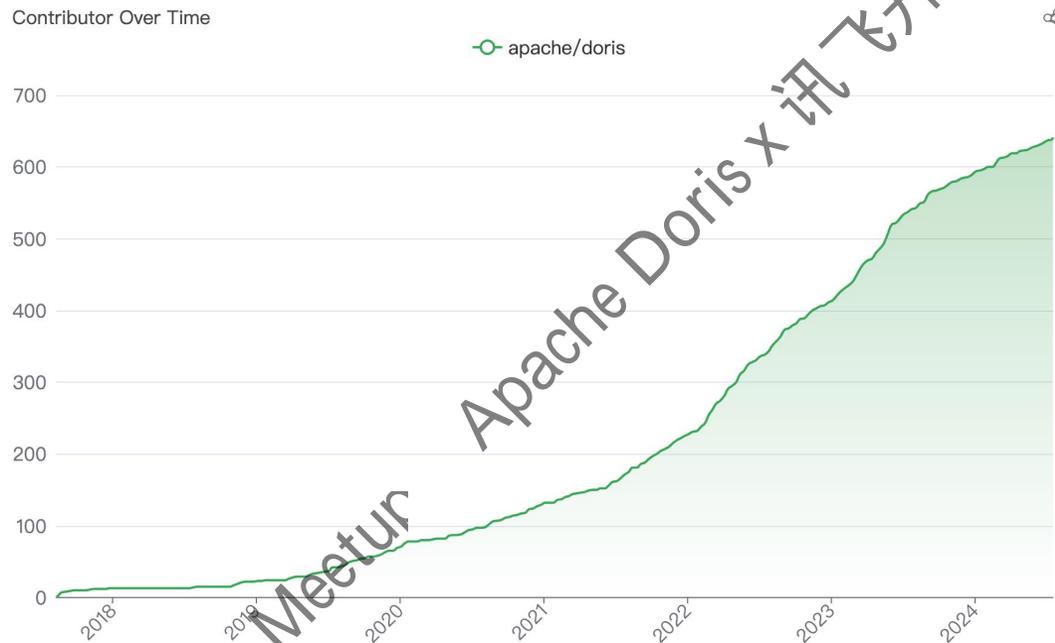
03 更多 3.0 版本特性揭秘

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: 截止2024年3月

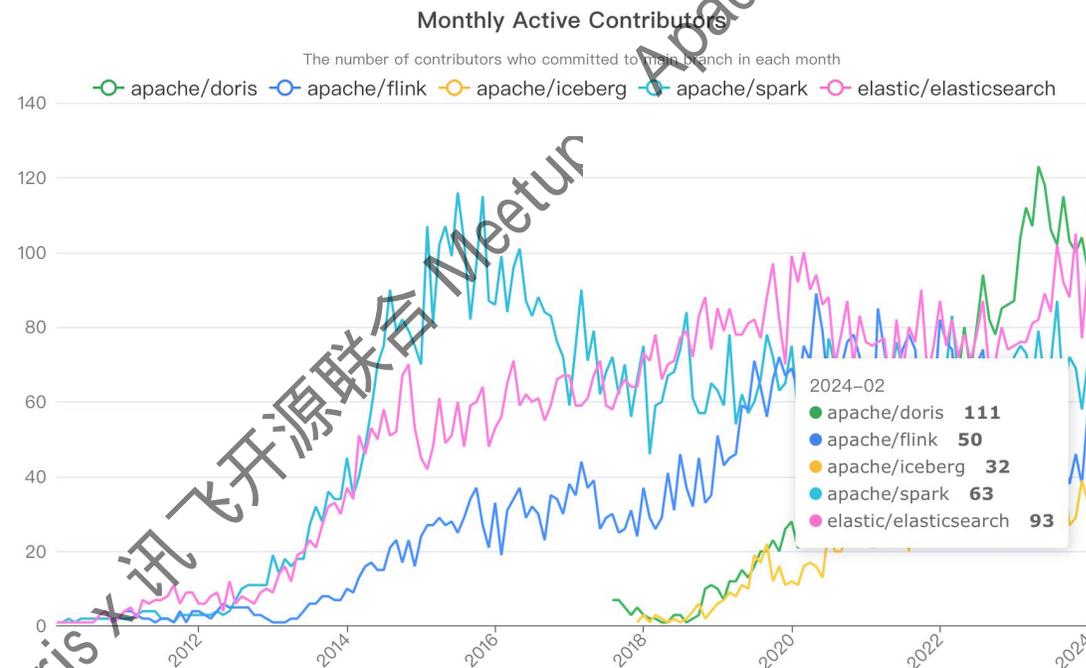
累计贡献者



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**

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



Doris 小助手

- 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

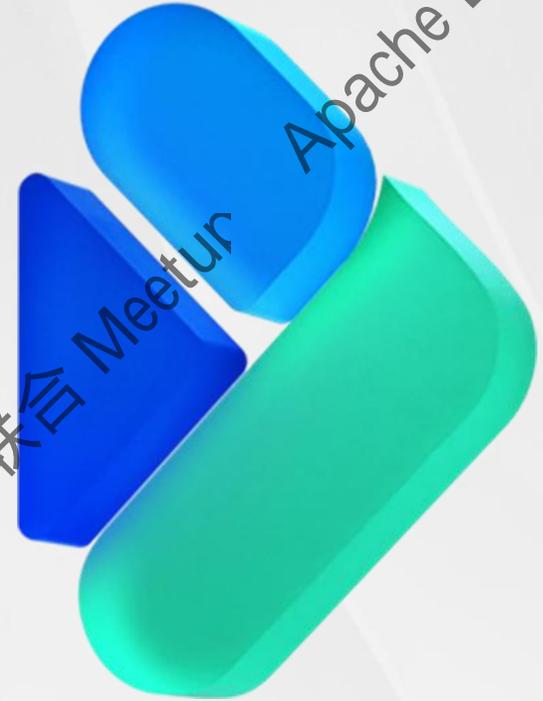
- Doris 小助手

如想进入社区用户社群请备注 [加群](#)



扫一扫上面的二维码图案，加我为朋友。

Thanks !



讯飞开源联合 Meetur

Apache Doris x 讯飞开源联合 Meetur

讯飞开源联合 Meetur

Apache Doris x 讯飞开源