

云原生时代， Apache Doris 的存算分离全新架构

飞轮科技 CEO、Apache Doris 创始人 马如悦

目录

- Apache Doris 与 SelectDB
- 为什么需要存算分离
- SelectDB Cloud 与 Apache Doris 全新存算架构

目录

Apache Doris 与 SelectDB

Apache Doris: 最流行的开源实时数据仓库



Apache Doris 是一个基于
MPP 架构的开源数据仓库，
支持对大规模实时数据上的极
速分析。

2013

项目正式建立

2017

项目正式开源

2022

商业化公司成立并
成为ASF顶级项目

9300+

GitHub Star

550+

贡献者

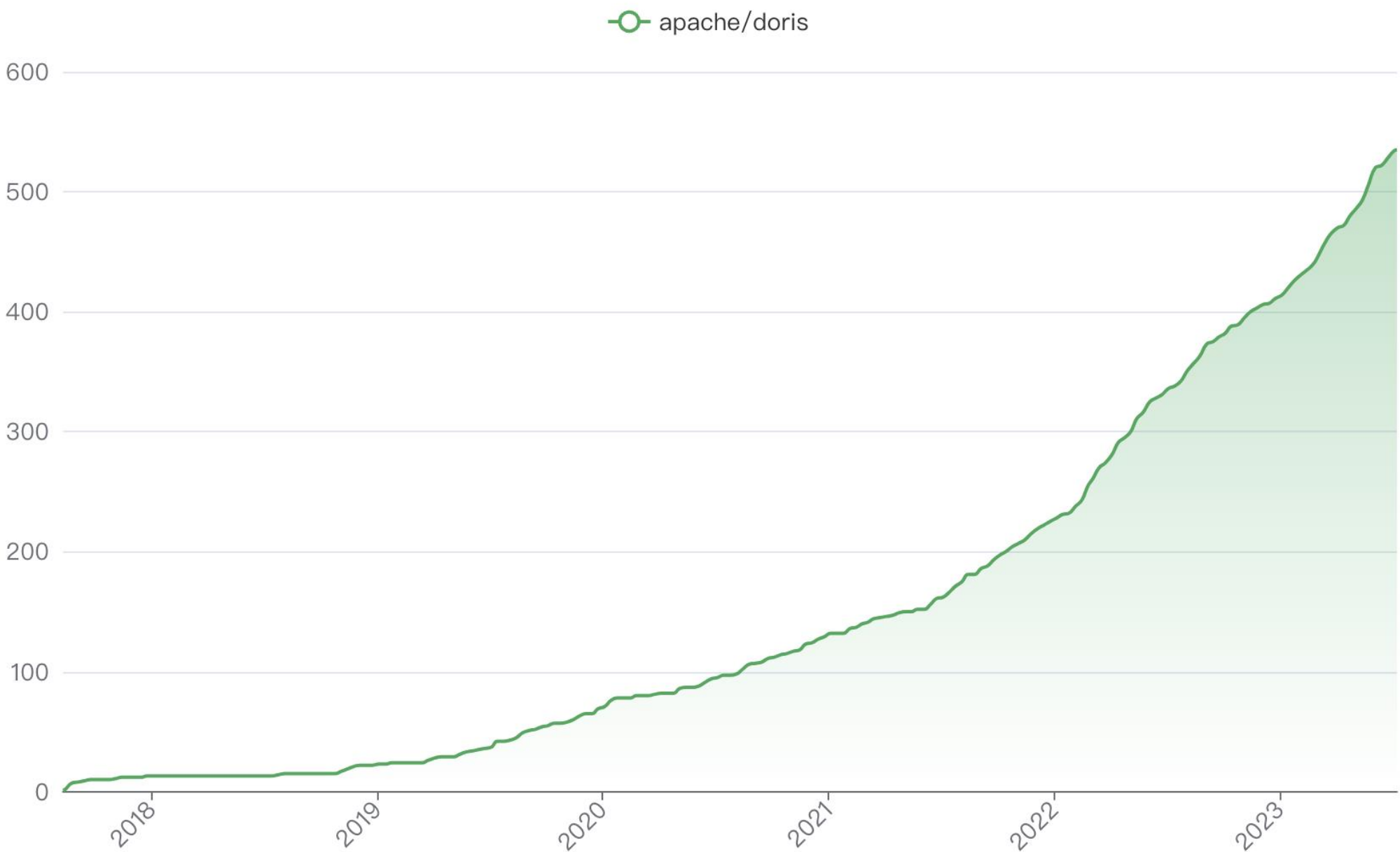
2000+

使用企业

Apache Doris 是当前全球最活跃的大数据项目之一

累计贡献者

GitHub Contributor Over Time

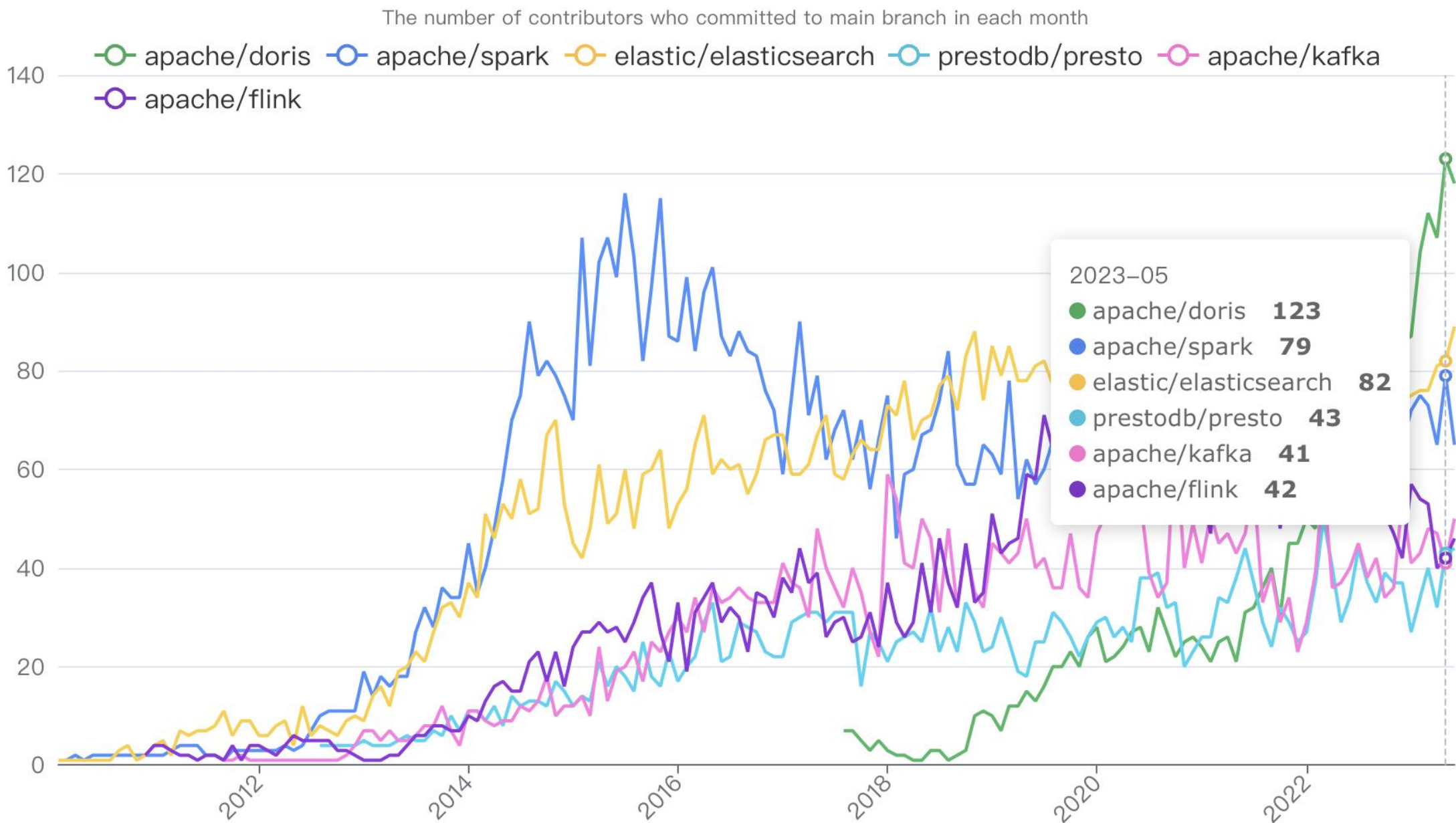


550+

累计贡献者超过 550+ 人，同比新增贡献者超过 120%，并仍处于持续上升的态势。

月度活跃贡献者

Monthly Active Contributors



Top1

在全球大数据开源项目排行中活跃贡献者数连续 10个月位列 Top1，已超越 Spark 最活跃时期。

Apache Doris 被 2000 多家中大型企业生产系统使用

Baidu 百度

美团 美团

ByteDance

Tencent

京东

MI xiaomi

Tencent Music

HUAWEI

Alibaba.com

miHoYo
TECH OTAKUS SAVE THE WORLD

NetEase Games

luckin coffee

360 数科

三七互娱
37 Interactive Entertainment

快手

huya 虎牙

Douyu

知乎
www.zhihu.com

keep

BOSS 直聘

航旅纵横
Umetrip.com

Moka

银联商务
China UMS

中信证券
CITIC SECURITIES

招商局 集团有限公司

杭银消费金融公司
HANGZHOU HONGYUAN FINANCIAL COMPANY

马上消费
WWW.MSXF.COM

嘉实基金
Harvest Fund
远见者 稳健

中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA

中国民生银行
CHINA MINSHENG BANK

拉卡拉

中国平安
PING AN

YTO express

YUNDA EXPRESS

ZTO 中通快递
ZTO EXPRESS

安踏

特步

华泰证券
HUATAI SECURITIES

澳优
Ausnutria

货拉拉
HuoLaLa

满帮集团

CAI NIAO 菜鸟

China unicom 中国联通

中国电信
CHINA TELECOM
世界触手可及

中国移动
China Mobile

海底捞火锅

Bondex 海程邦达

长安汽车
CHANGAN AUTO

慧策

飞轮科技是基于 Apache Doris 的商业化公司



开源 + 商业 双轮驱动
服务用户、客户、开发者

大力将开源 Apache Doris 发扬光大，并基于 Apache Doris 推出商业产品，实现开源和商业的良性增长飞轮

SelectDB 是基于 Apache Doris 的商业化产品

SelectDB Cloud

全托管的 Apache Doris 云服务。
已经上线阿里云、AWS、华为云和腾讯云。

SelectDB Enterprise

自管理的私有化部署软件，部署在物理机/虚拟机、K8s 或者私有云上。

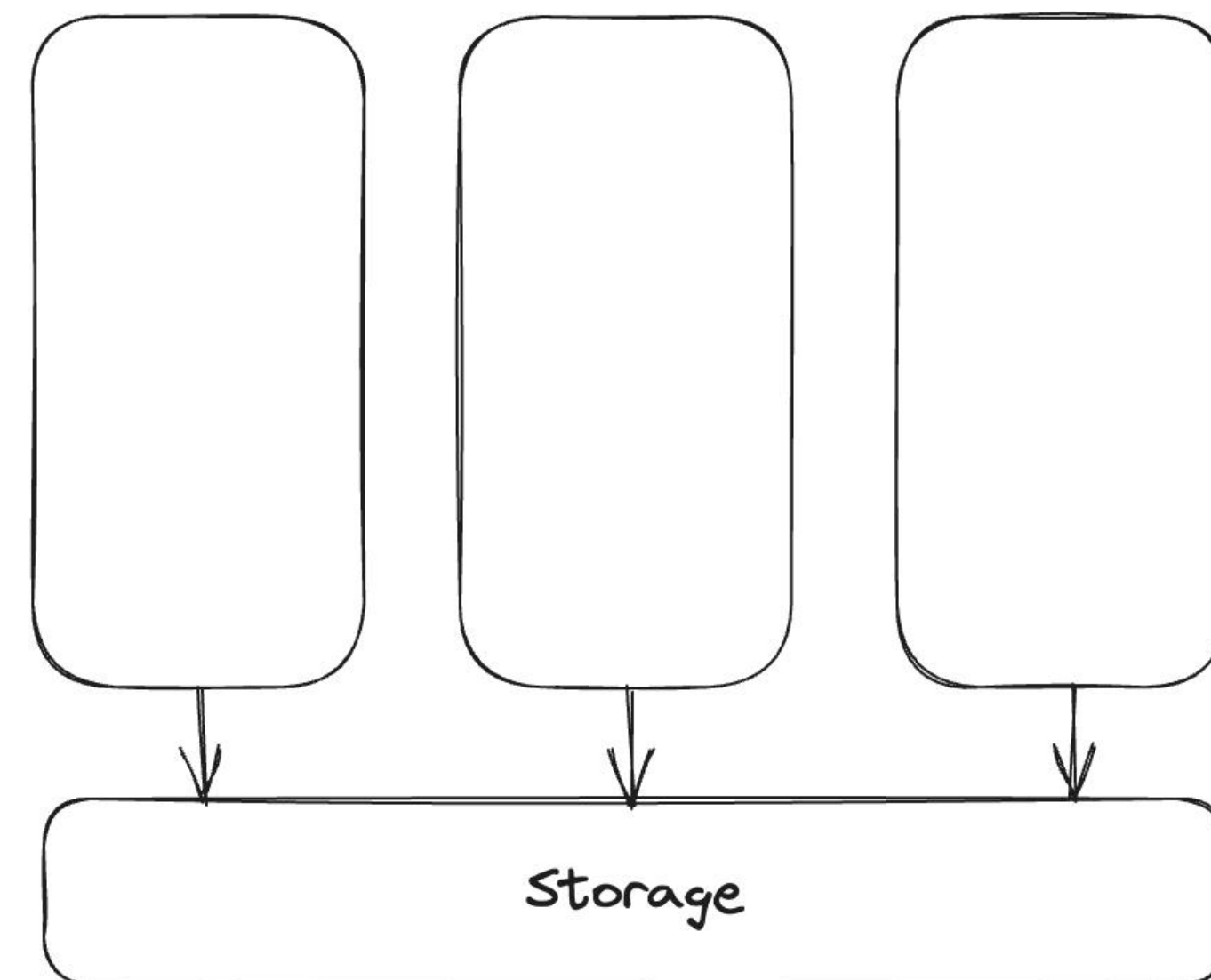
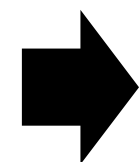
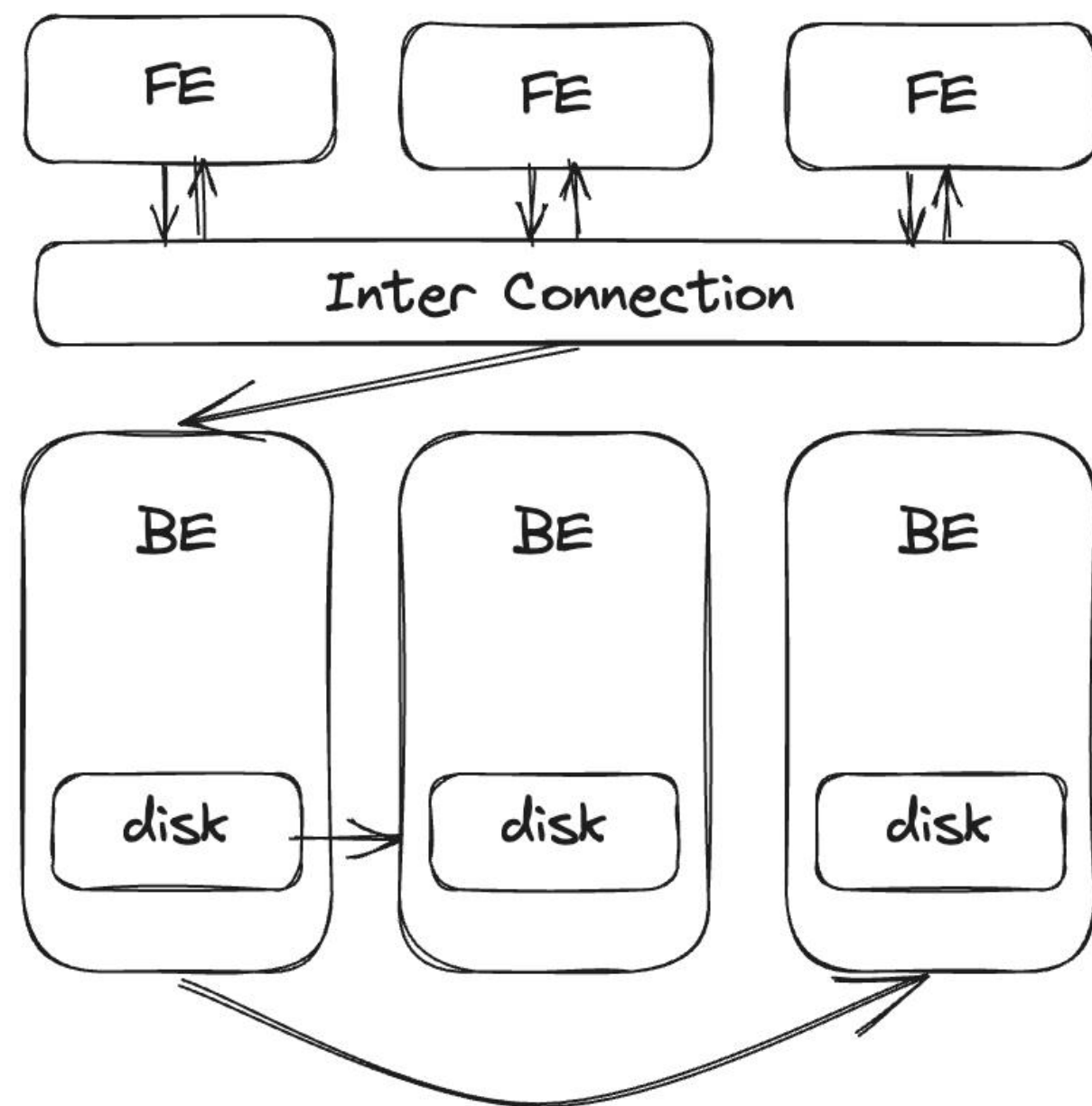
Apache Doris

开源的实时数据仓库引擎

目录

为什么需要存算分离

存算分离 vs 存算一体



- 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

- 为什么还需要存算分离?
- 什么样的存算分离架构更能贴合用户需求?
- 要克服那些挑战?

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

可靠性与负载隔离

- 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

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，资源消耗更低

从实际需求出发，2.0 版本降低成本提升弹性



低成本

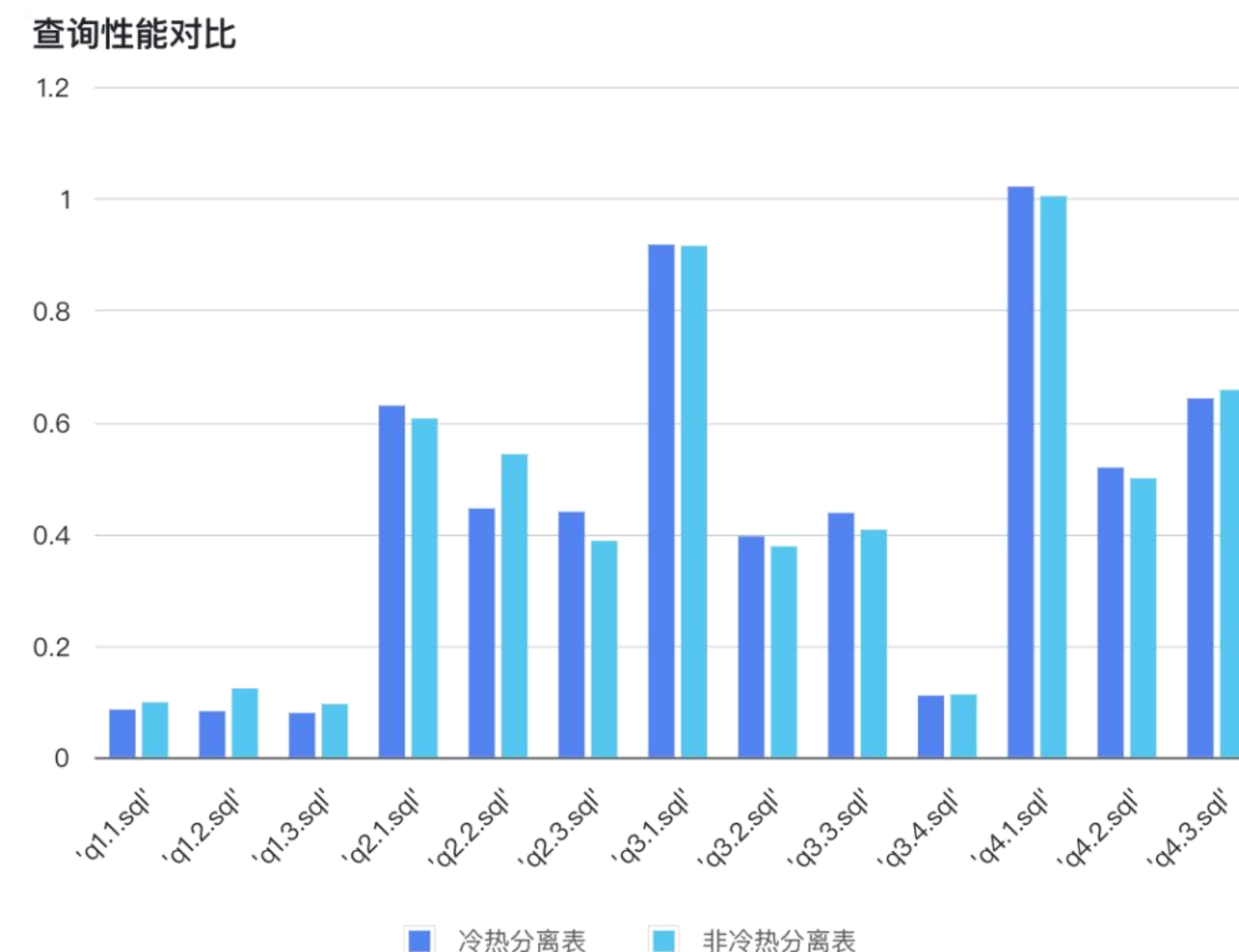
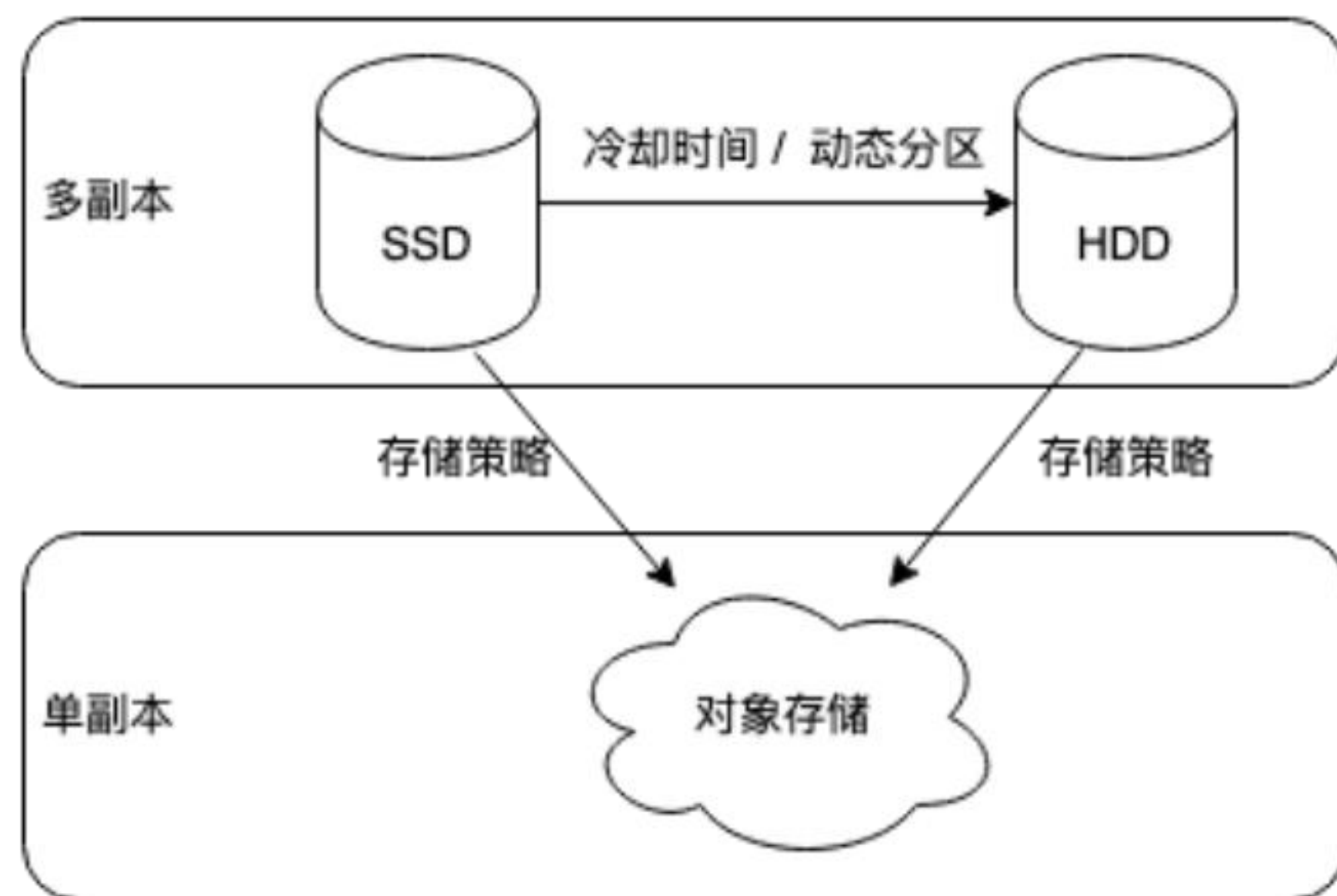
- 引入对象存储节省冷数据资源
 - 增加弹性计算节点



逐步迭代

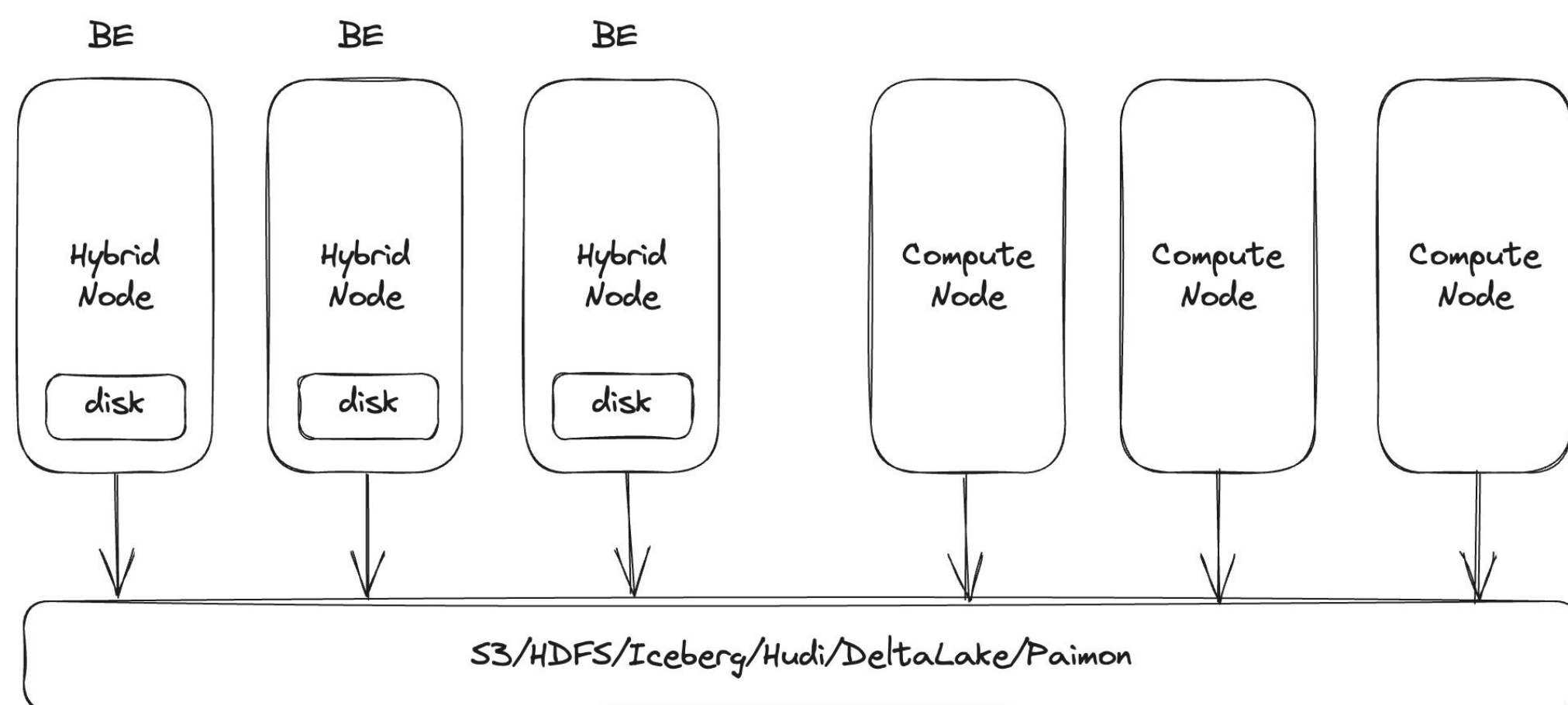
- 绝大多数用户已采取存算一体架构
- 升级过程中需要保证对已有架构的兼容

2.0 冷热数据分层，实现冷数据存储分离



- 根据将冷热数据分别存储在成本不同的存储介质上，从原本的 SSD->HDD 增加到 SSD -> HDD ->OS 三层；
- 云磁盘的价格通常是对象存储的 5-10 倍，如果可以将 80% 的冷数据保存到对象存储中，存储成本至少可降低 70% ；
- 通过冷数据 Compaction 实现数据的高效压缩，提供冷数据 Cache 加速冷热数据查询，降低成本的同时保持性能不受影响；

2.0 弹性计算节点，实现计算分离



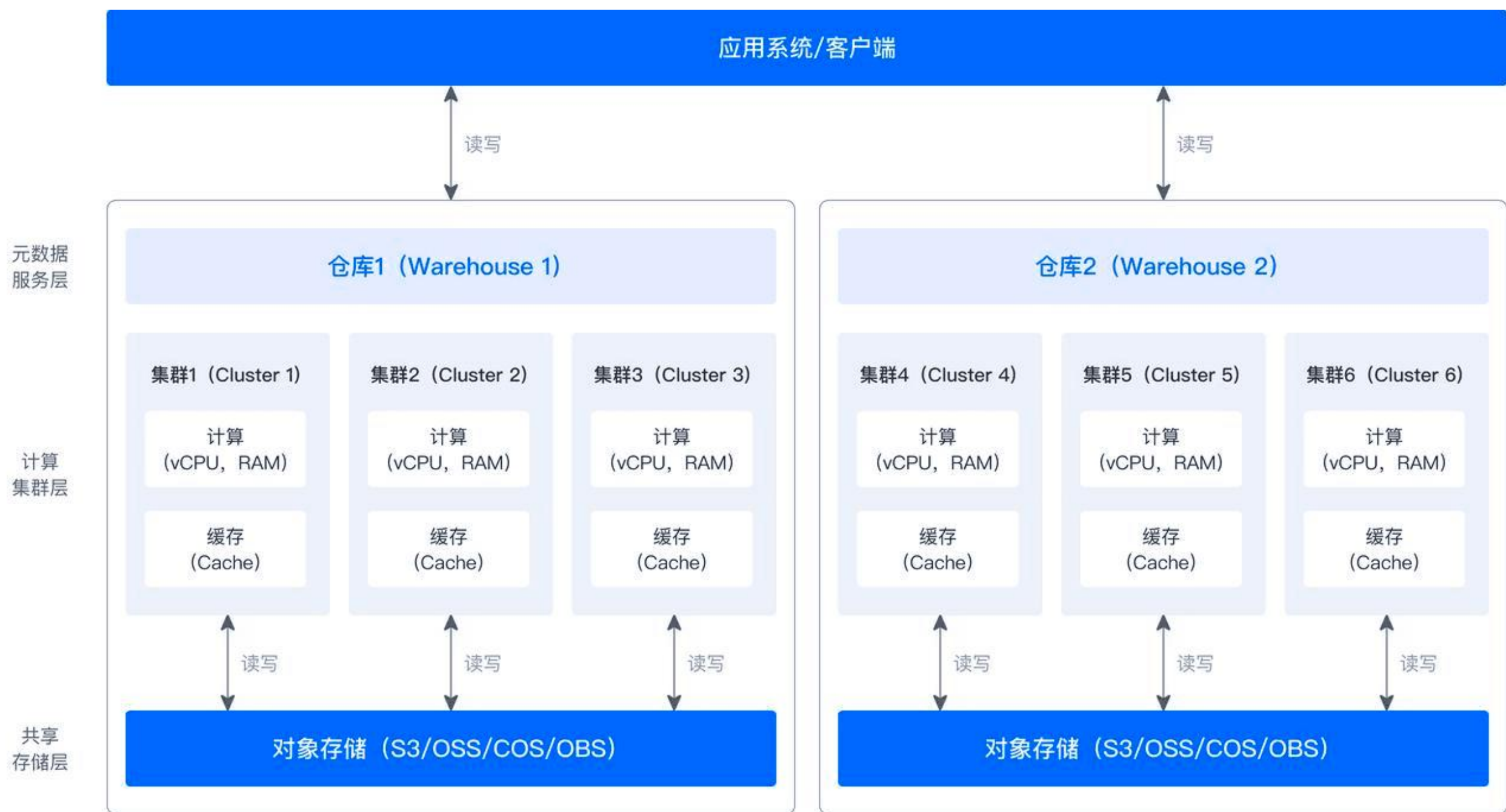
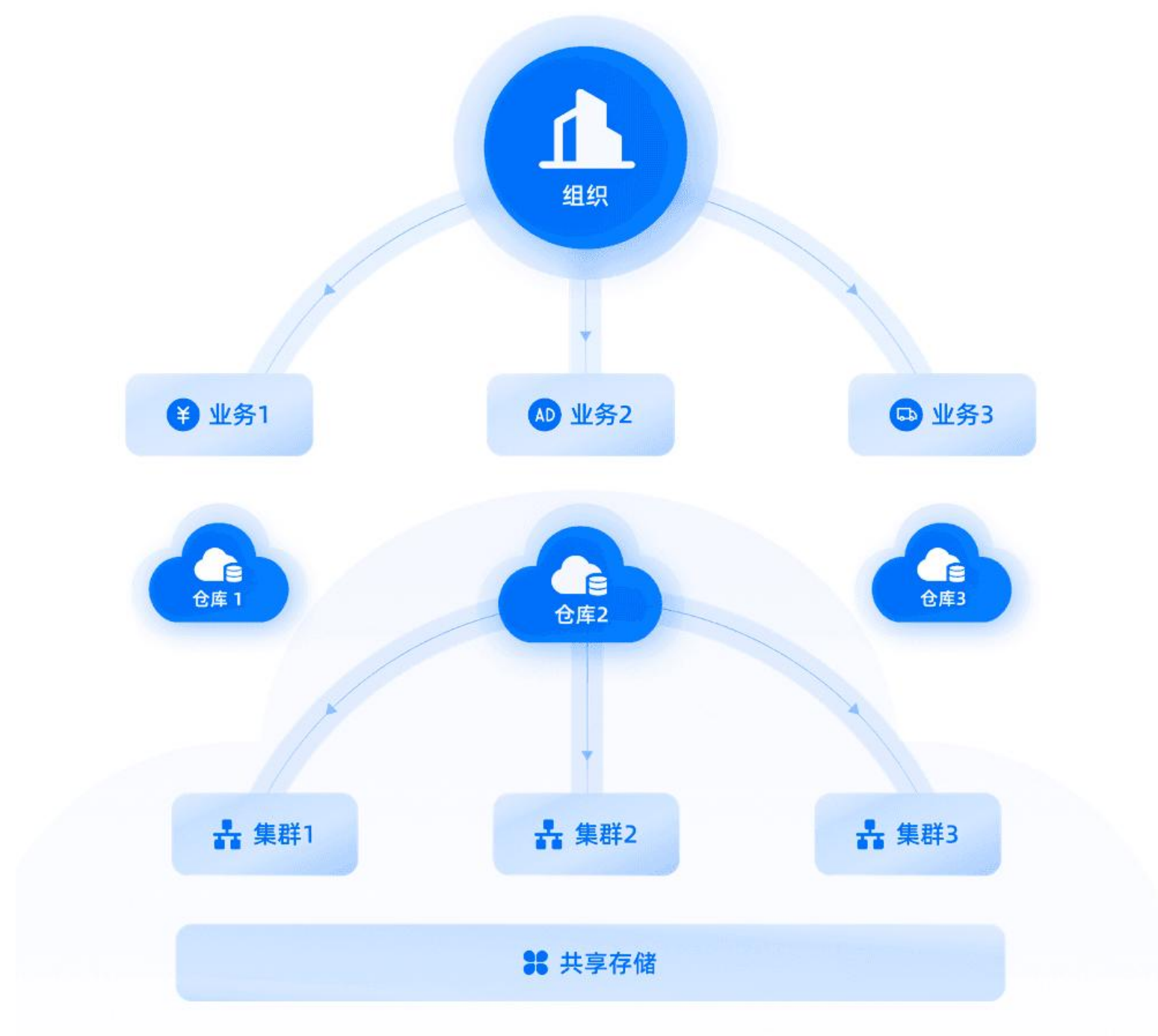
- 引入了无状态的计算节点 **Compute Node**，专门用于执行计算分析；
- 不保存任何数据，集群扩缩容时无需进行数据分片的负载均衡，具有明显高峰的场景中可以灵活扩容、快速加入集群分摊计算压力；
- 用户数据往往存储远端存储中，执行查询时查询任务会优先调度到 **Compute Node** 执行，以避免内表与外表查询之间的计算资源抢占；

目录

SelectDB Cloud 与 Apache Doris 全新存算分离架构

即将合入 Apache Doris 2.1 版本

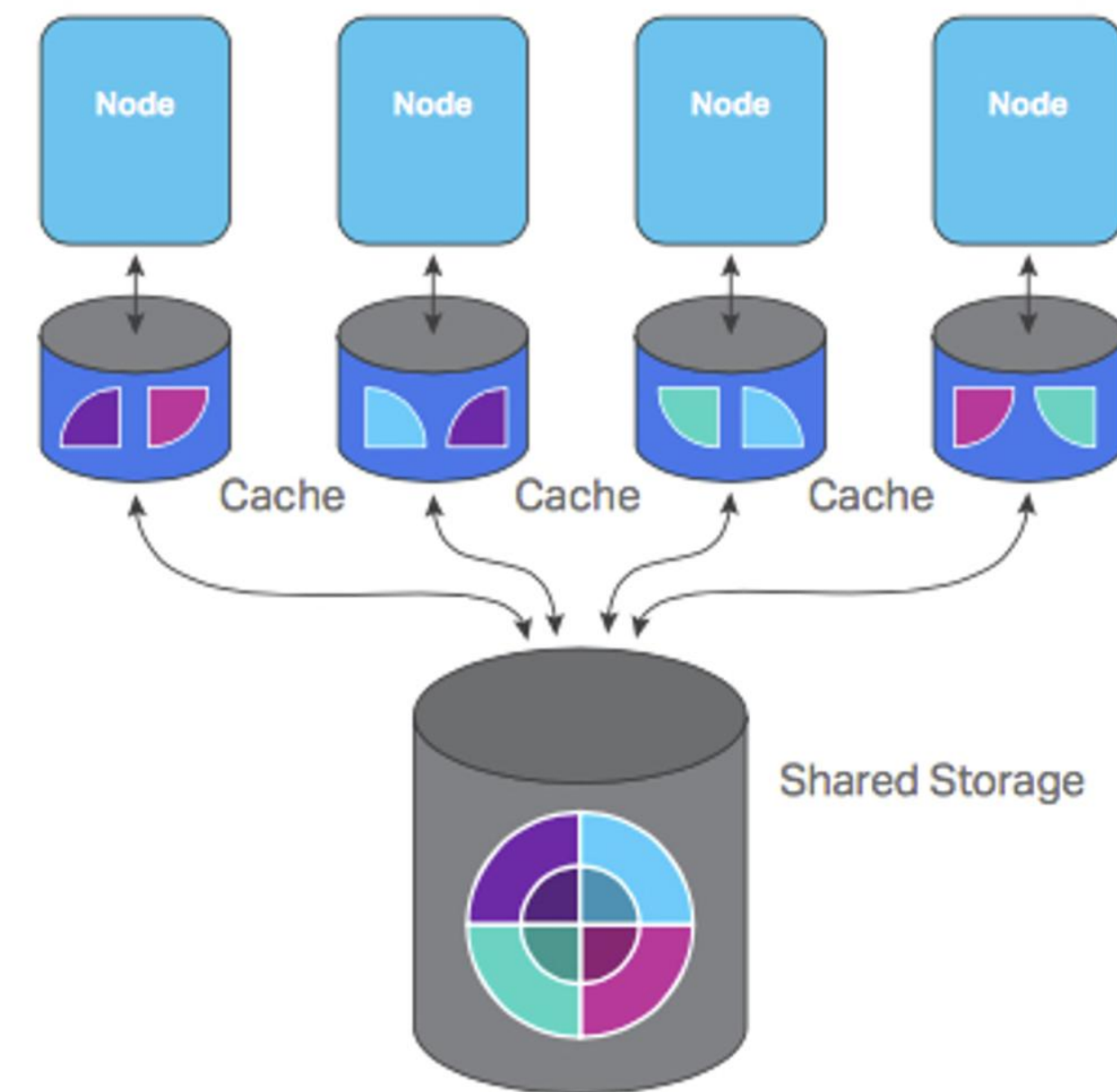
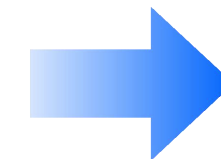
SelectDB Cloud 核心概念与云原生架构



共享存储与本地高速缓存



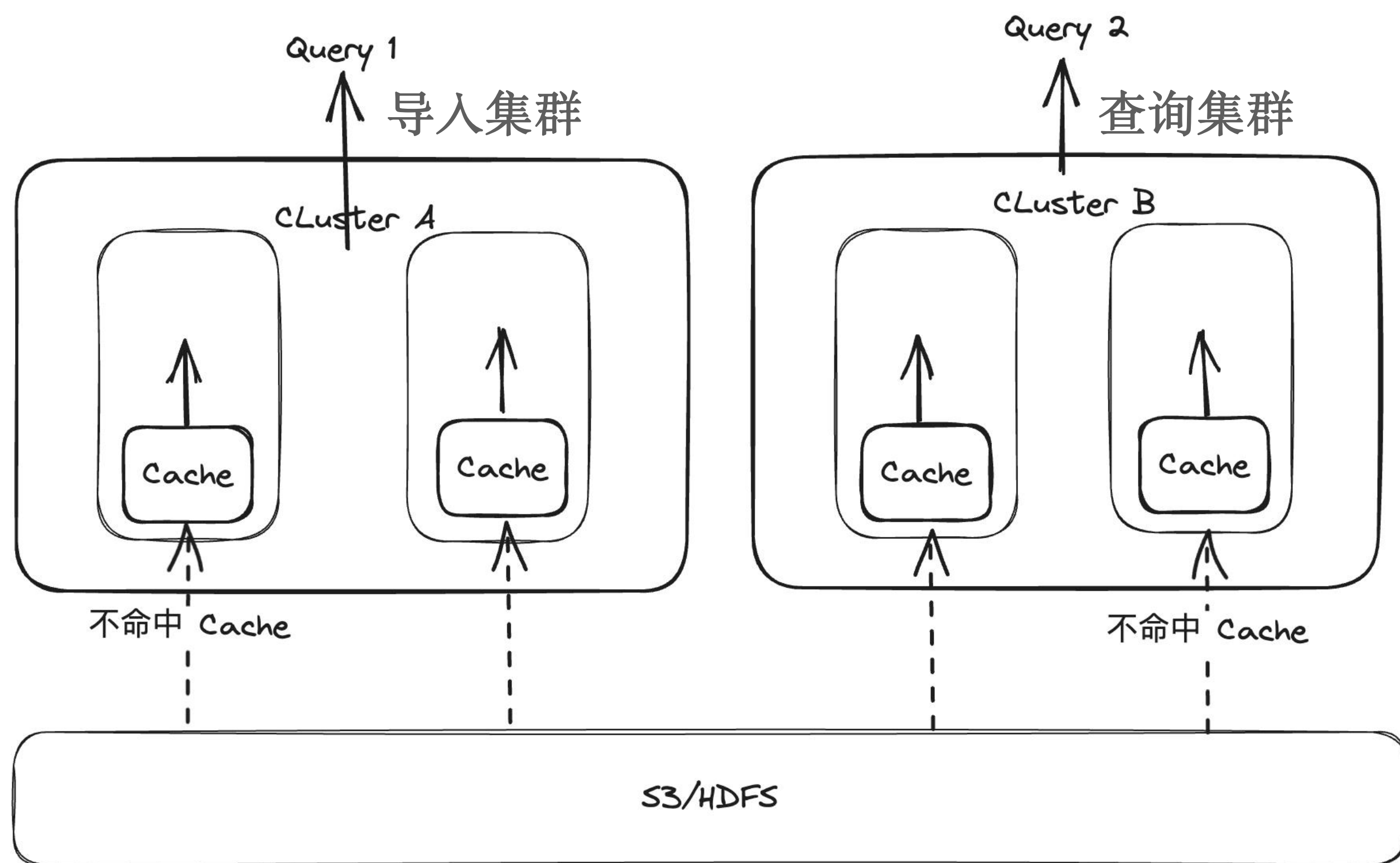
谓词无法下推，网络传输数据量大



对频繁查询的数据或者手动设定的数据进行缓存
新增节点和删除节点，需要处理缓存预热或者迁移

计算与计算分离：多计算集群

多计算集群使用举例



- 元数据和数据共享，数据强一致
- 多个计算集群都可以读写
- 热数据缓存实时同步
- 使用场景
 - 导入的工作负载与查询的负载进行隔离
 - Ad-hoc 的大查询负载和在线点查询的负载间相互隔离
 -

计算与计算分离：多计算集群演示

WH_Guide
ALBJ67HH

集群

连接

监控

用量

设置

集群

Cluster_Import

运行时长 0 天 0 小时 42 分钟
创建时间 2023-07-23 17:57:25

运行中

⏸ 停机 ⏹ 重启

Cluster_Write

运行时长 0 天 0 小时 40 分钟
创建时间 2023-07-23 17:58:14

运行中

⏸ 停机 ⏹ 重启

Cluster_ETL

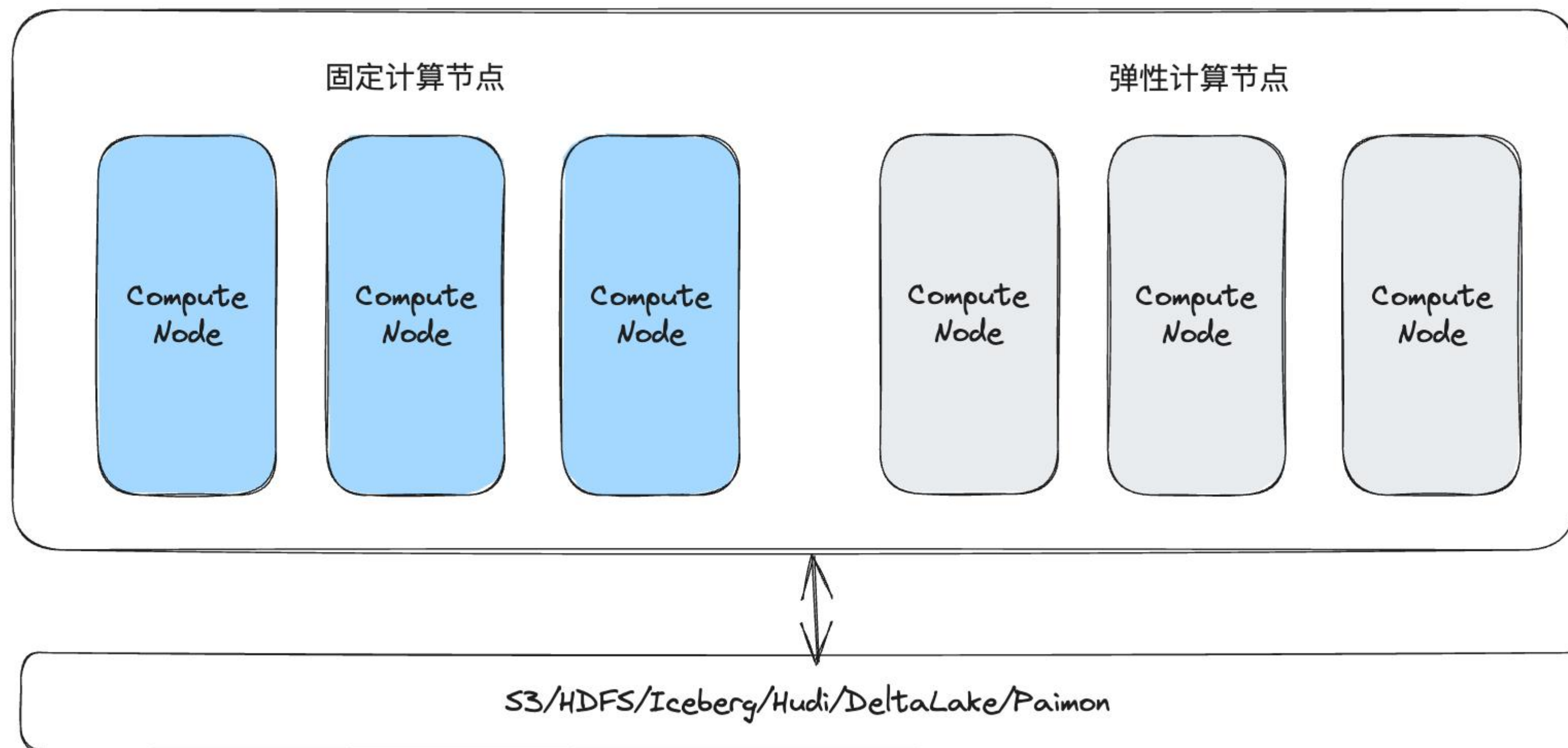
运行时长 0 天 0 小时 4 分钟
创建时间 2023-07-23 19:12:17

运行中

⏸ 停机 ⏹ 重启

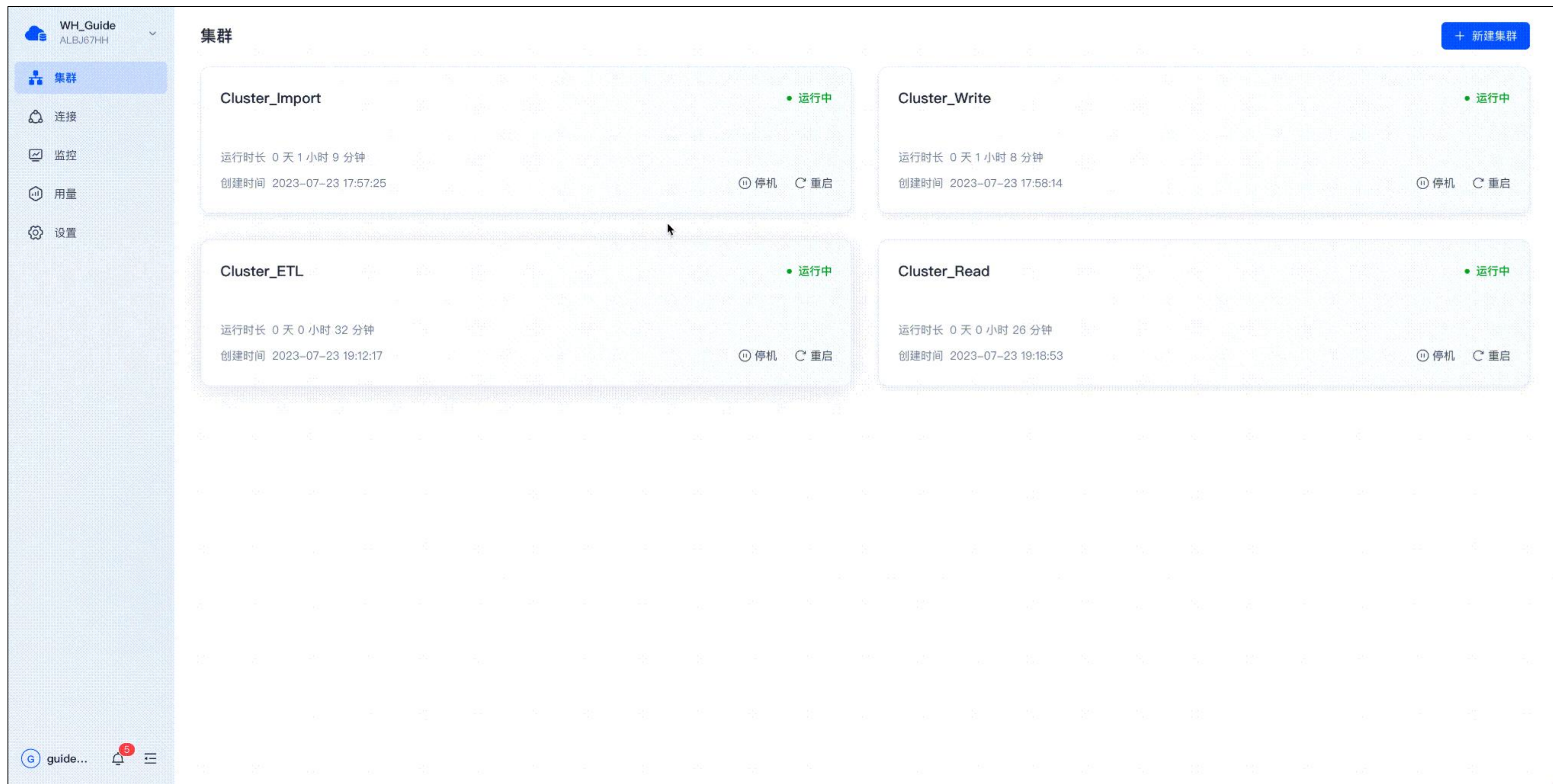
+ 新建集群

计算集群内的计算弹性扩缩容



- 手动扩缩容
- 分时扩缩容
- 集群自动停机

计算集群内的计算弹性扩缩容 – 手动扩缩容演示



计算集群内的计算弹性扩缩容 – 分时扩缩容演示

WH_Guide
ALBJ67HH

集群

连接

监控

用量

设置

< Cluster_Read • 运行中

...

集群伸缩

转包年包月

基础信息

集群 ID
c-7uyo4c2qeotq9rj9hd

集群名称
Cluster_Read

集群计算资源
4 vCPU, 32 GB 内存

集群缓存空间
200 GB

计费方式
按量计费 转包年包月

其它信息

创建者
guide@selectdb.com

创建时间
2023-07-23 19:18:53

启动时间
2023-07-23 19:20:02 设置自动启停

运行时长
0 天 1 小时 29 分钟

guide...

9

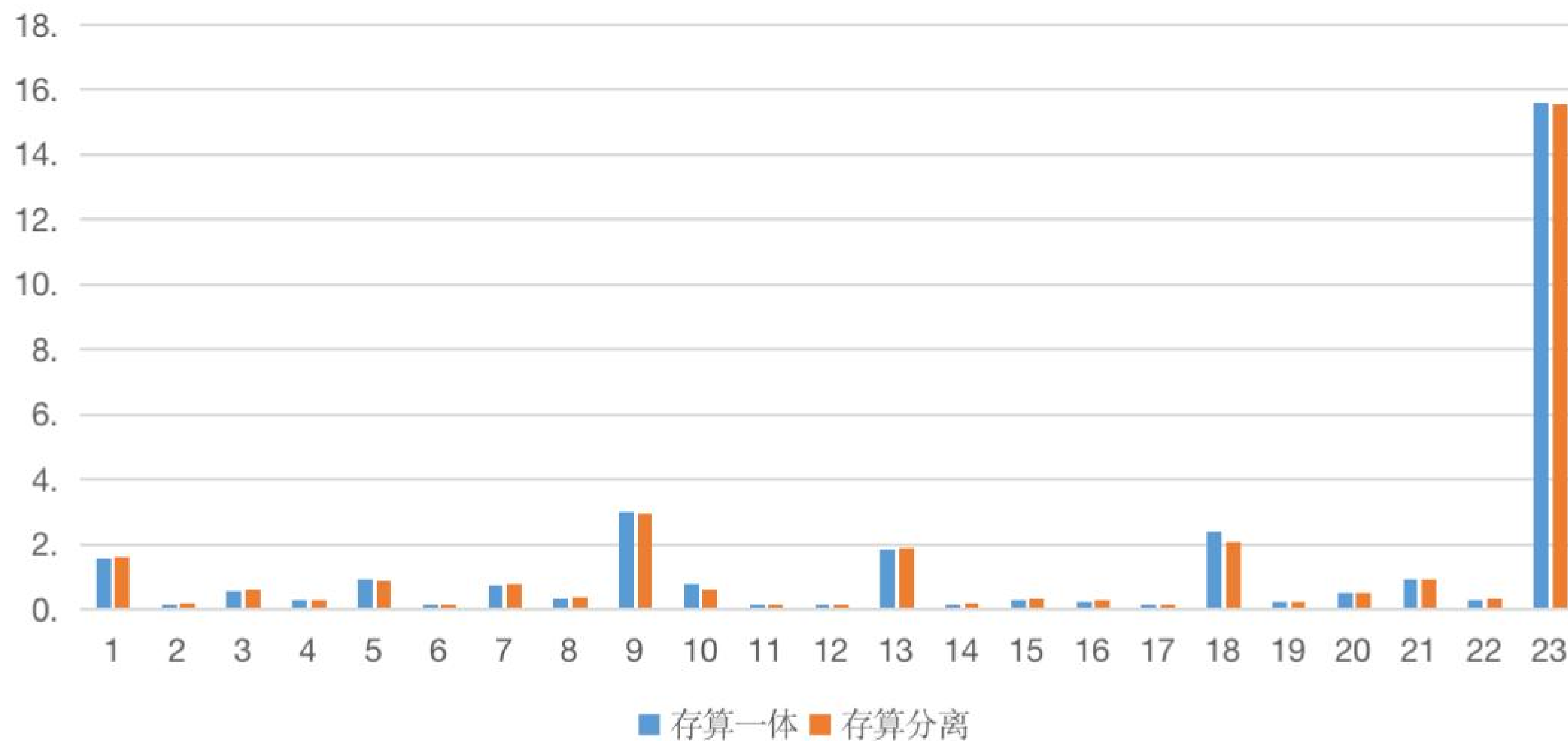
≡

计算集群内的计算弹性扩缩容 – 集群自动启停



存算分离 vs 存算一体 性能对比

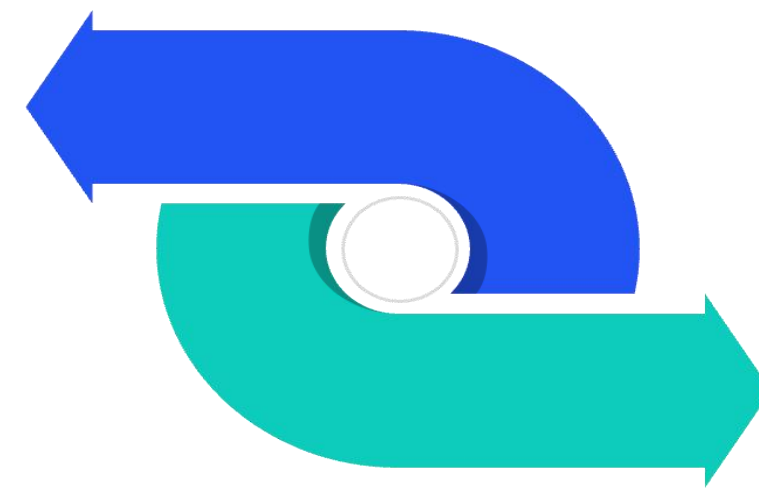
TPCH-SF100



两种架构的适用场景

存算一体

- 对性能要求更高、对查询时延更敏感
- 不需要极致弹性扩缩容
- 没有可靠的共享存储可用
- 仅简单体验或开发测试
- 业务线独立运行Doris，无专职DBA
-



- 成本敏感型，追求性价比
- 已经在公有云上构建服务
- 拥有可靠的共享存储服务
- 要求极致弹性，需要运行在容器平台上
- 有专职的团队维护数据仓库平台

存算分离

THANKS



软件正在重新定义世界

Software Is Redefining The World