

**Apache Doris 上云最佳选择：
阿里云数据库 SelectDB 版**

CONTENT 目录

01 阿里云 SelectDB 版介绍

02 阿里云 SelectDB 版核心特性

03 阿里云 SelectDB 版案例介绍

Part 1

阿里云 SelectDB 版整体介绍

Apache Doris 是全球流行的开源实时数仓产品



Open-Source Real-Time Datas Warehouse

2013 Project Creation

2017 Open Source

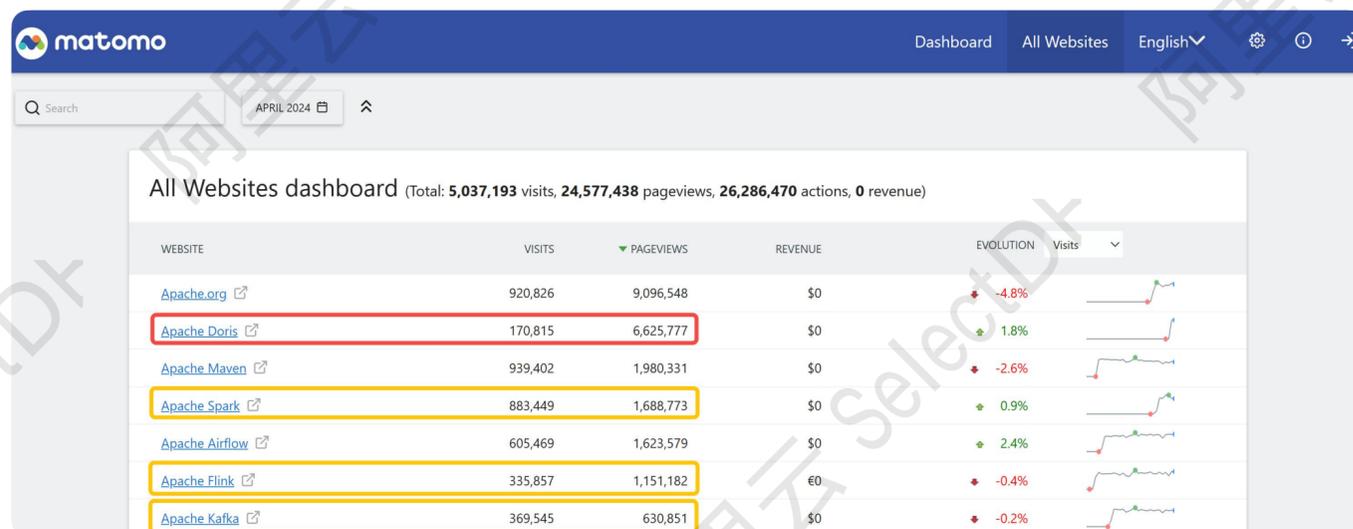
2022 ASF Top Project

11k+ GitHub Stars

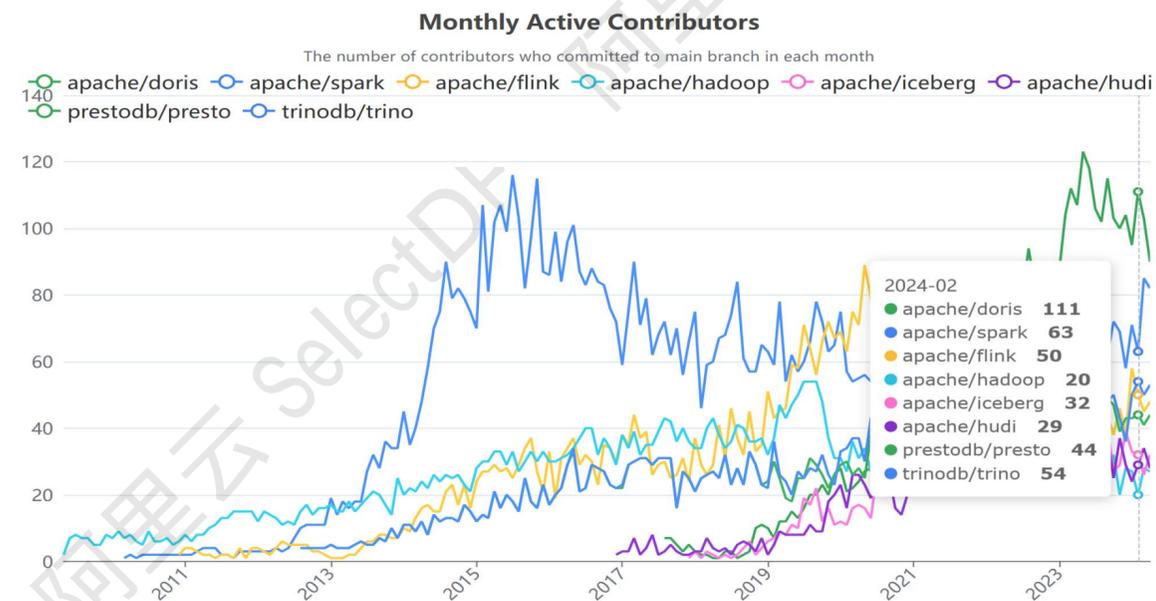
620+ Contributors

5000+ Enterprises

Apache 所有项目中 官网每日浏览量 No.1



开源大数据、数据库项目中 月活开发者 No.1



阿里云数据库 SelectDB 版： 基于 Apache Doris 构建的商业化云原生实时数仓产品

1

可靠服务

云数据库 SelectDB 是阿里云一方产品，
阿里云提供服务及 SLA 保障

2

强大支撑能力

云数据库 SelectDB 由阿里云产研及 Apache Doris
主创团队（完成 90% 新增代码）联合开发和支持

3

完善生态对接

云数据库 SelectDB 完整适配
Apache Doris 开源生态以及阿里云产品生态



Part 2

阿里云 SelectDB 特性介绍

阿里云数据库 SelectDB 版：云原生存算分离



存算分离与弹性计算

- 核心存储：对象存储 OSS，实时扩容
- 计算集群：分钟级弹性，支持多集群
- 缓存存储：通过云盘缓存加速，秒级扩容

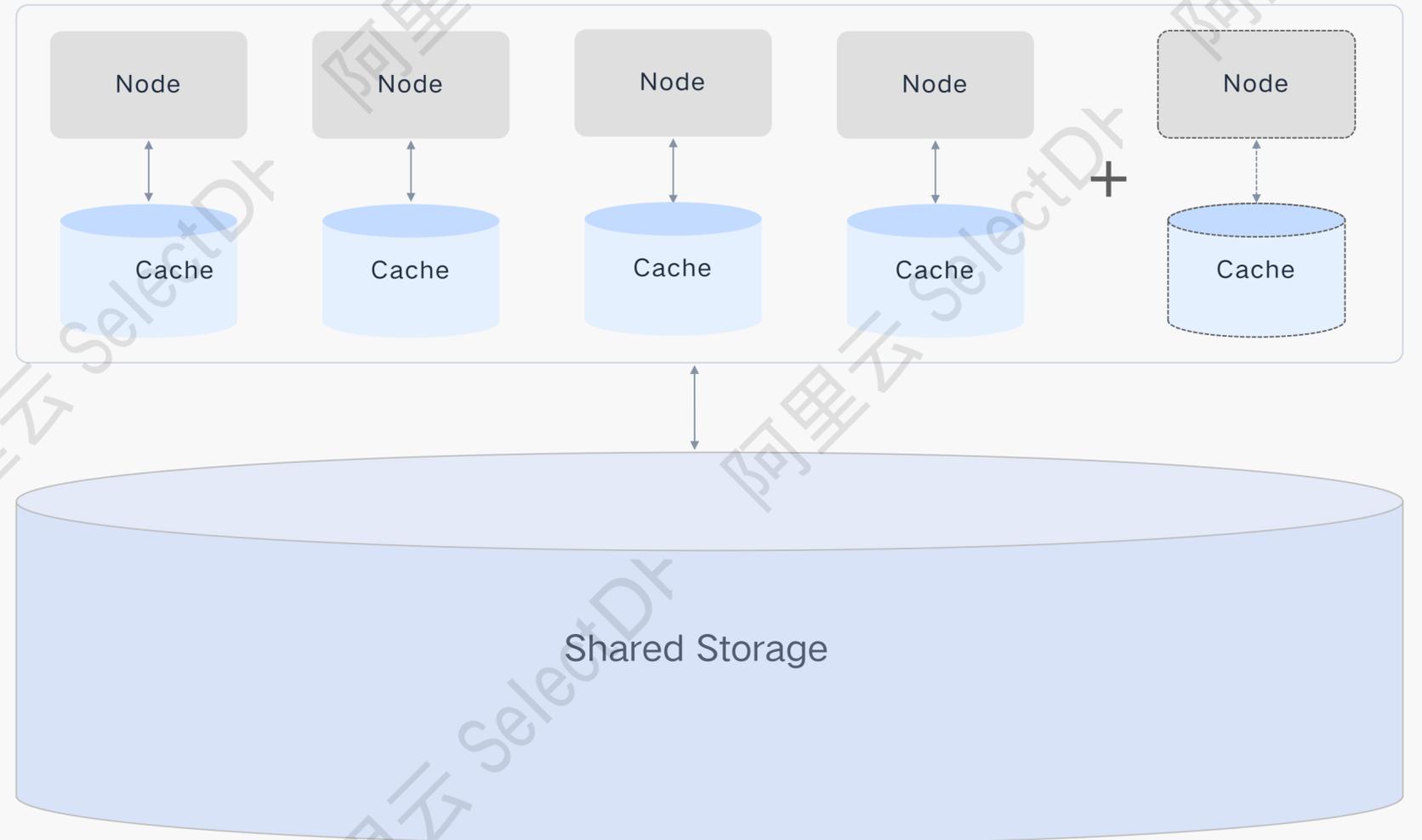
极致性价比

- 存储：数据分片单副本，相比存算一体降低 66%~；对象存储单价相比云盘低 80%+
- 计算：单副本消除冗余数据写入、合并开销，结合弹性，计算资源可降低 50%~

阿里云数据库 SelectDB 版：无全量迁移的实时弹性

实例各层资源实时弹性表现

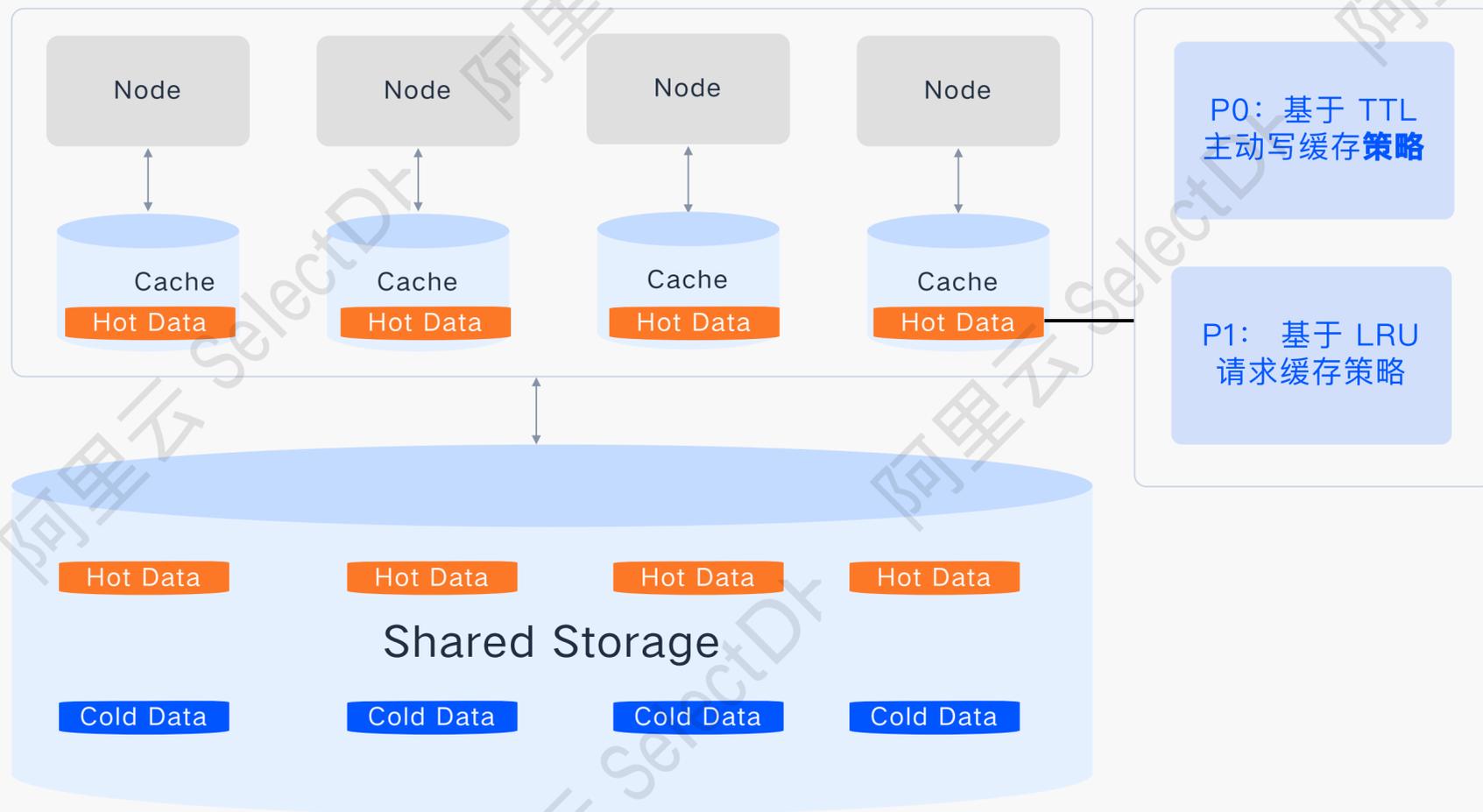
- 计算层：分钟级完成水平扩展和缓存迁移
- 存储层：Serverless 按需实时扩容
- 缓存层：基于云盘扩容，无感秒级



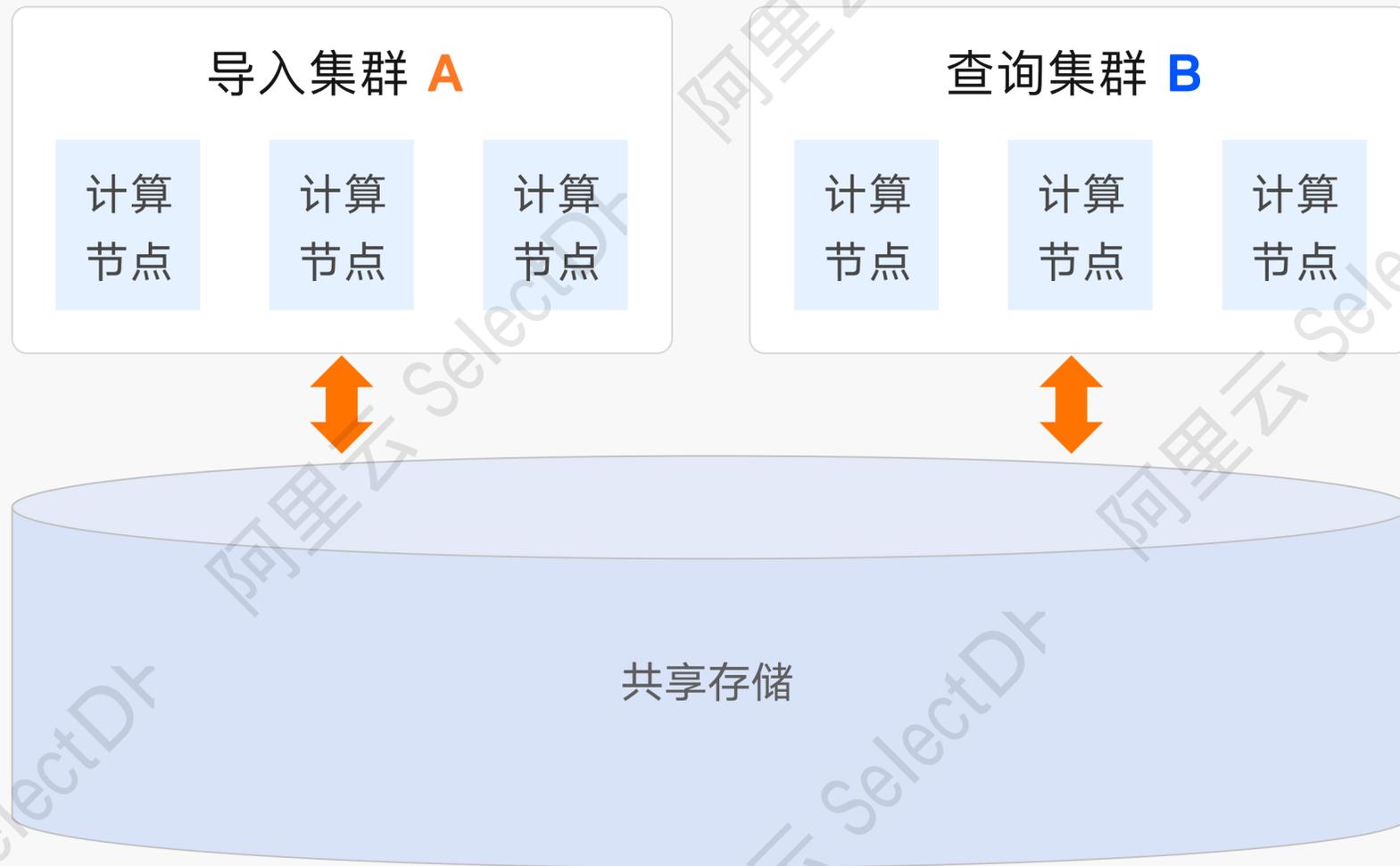
阿里云数据库 SelectDB 版：单副本写&读写 Cache 加速性能

SelectDB 读写缓存策略

- 基于 TTL 缓存策略，指定写入数据缓存时间
- 基于 LRU 的默认读写缓存策略，加速读写
- 指定表和分区预热，提高查询效率
- 单副本写入，降低 IO，加速写入性能



阿里云数据库 SelectDB 版：单实例多计算组资源隔离



隔离性和一致性保证

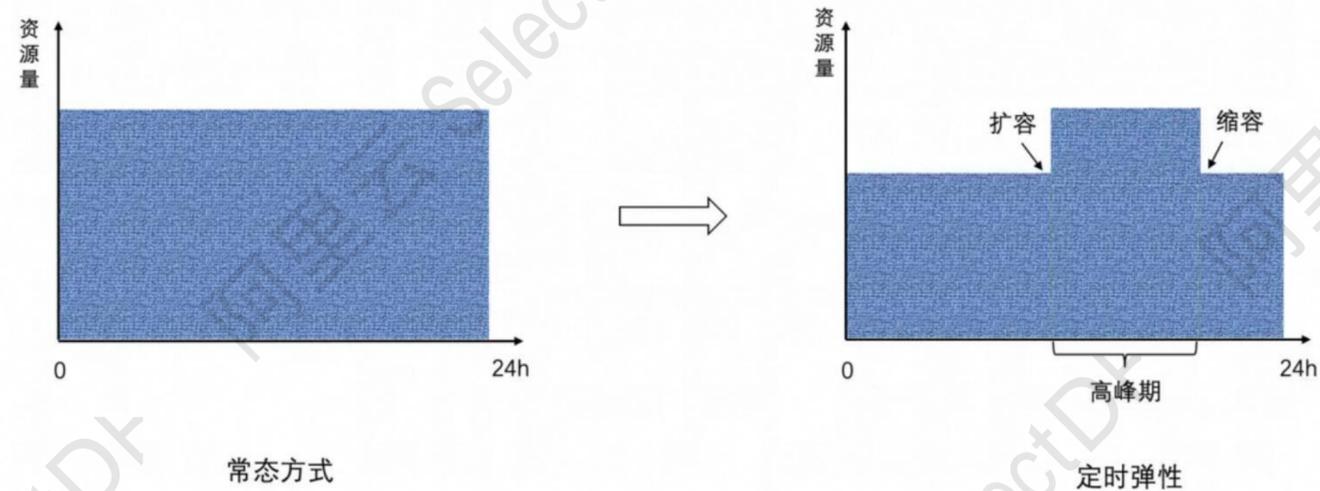
- 多集群元数据和数据共享，数据强一致
- 多个计算集群都可以读写
- “热数据”缓存数据预热

典型使用场景

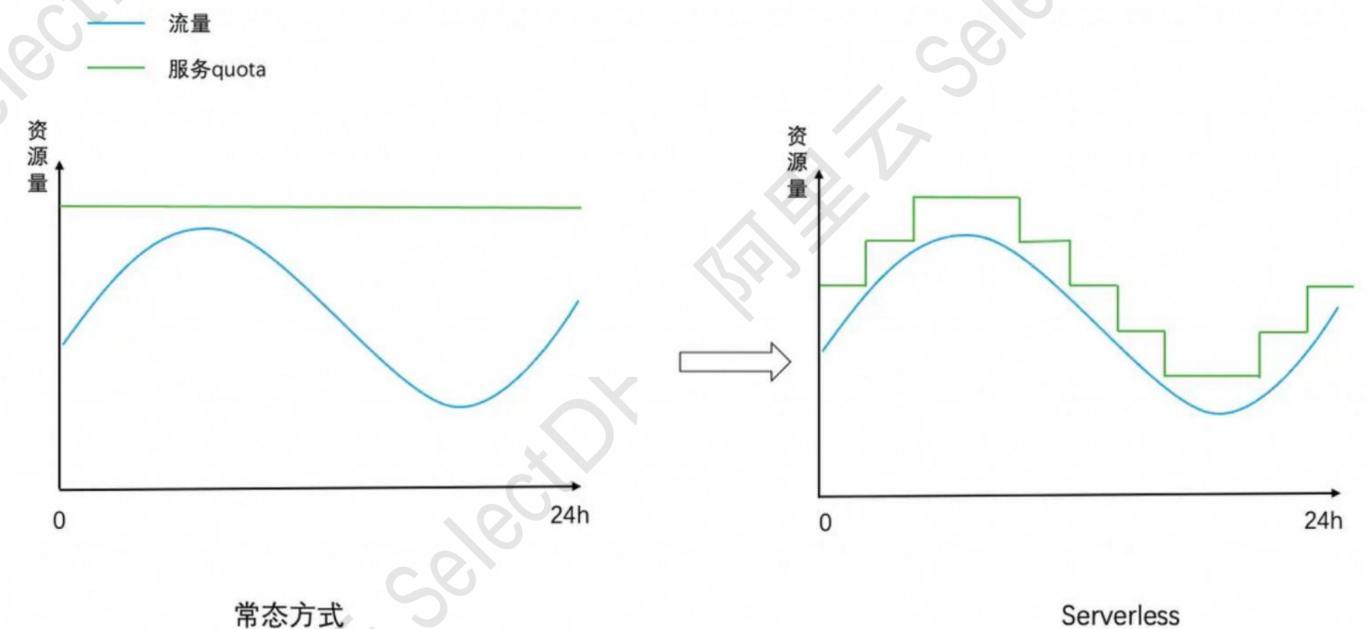
- 导入工作负载与查询负载隔离
- 大查询负载和在线点查询的负载隔离

阿里云数据库 SelectDB 版：支持云原生秒级弹性和 Serverless

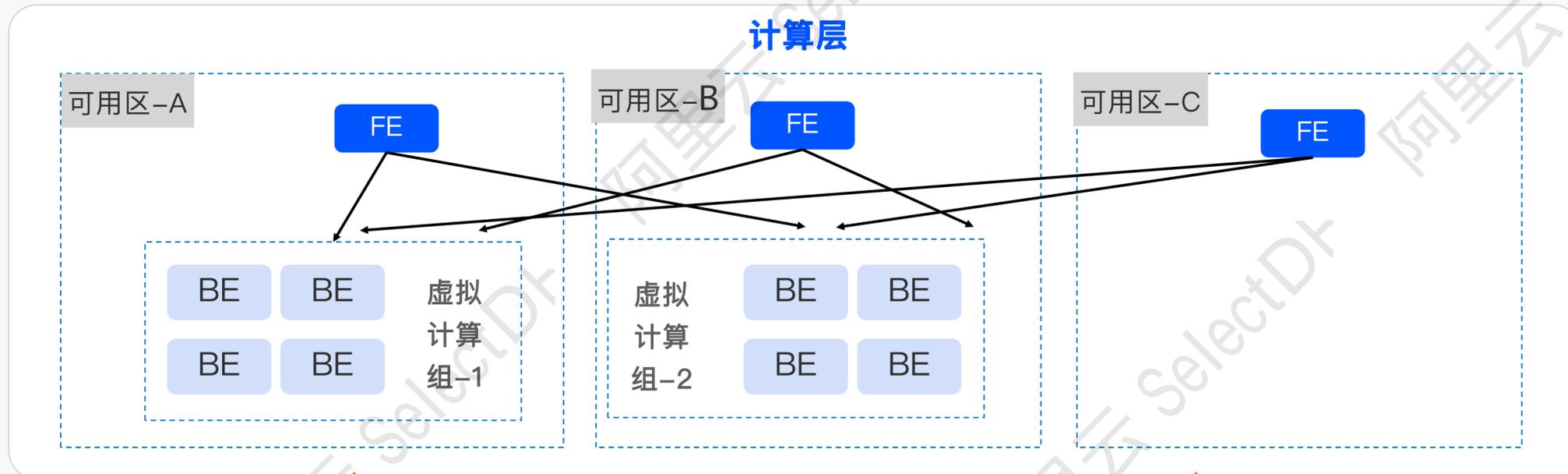
阶段一：计算基于业务周期秒级定时弹性（已发布）



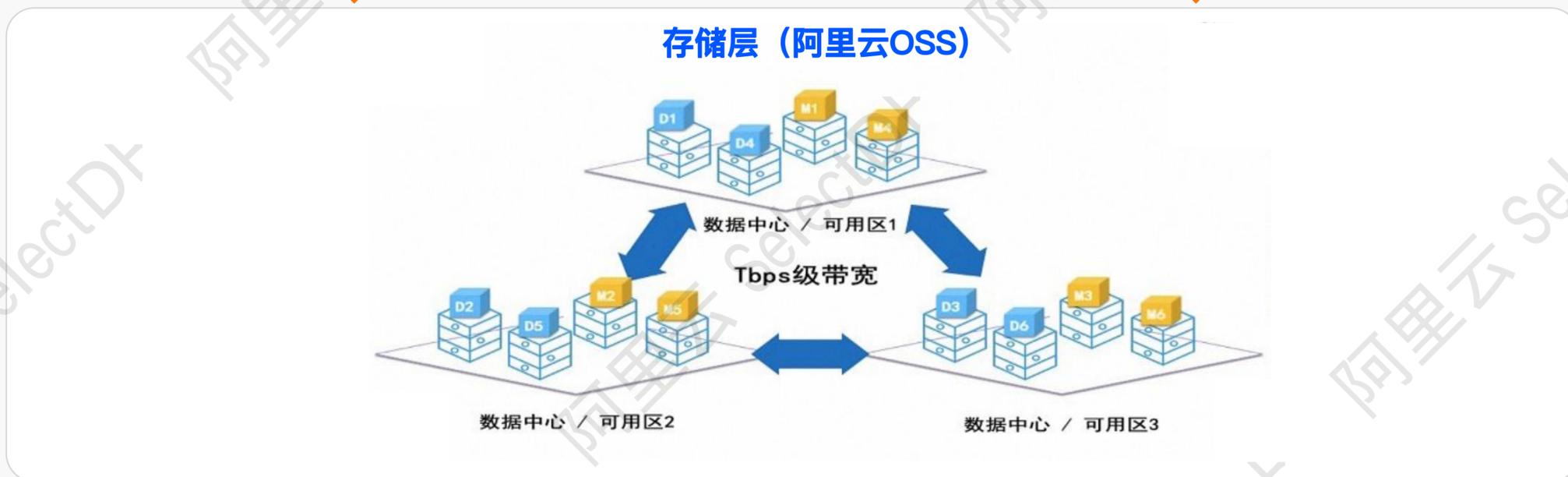
阶段二：基于负载 Autoscaling（开发中）



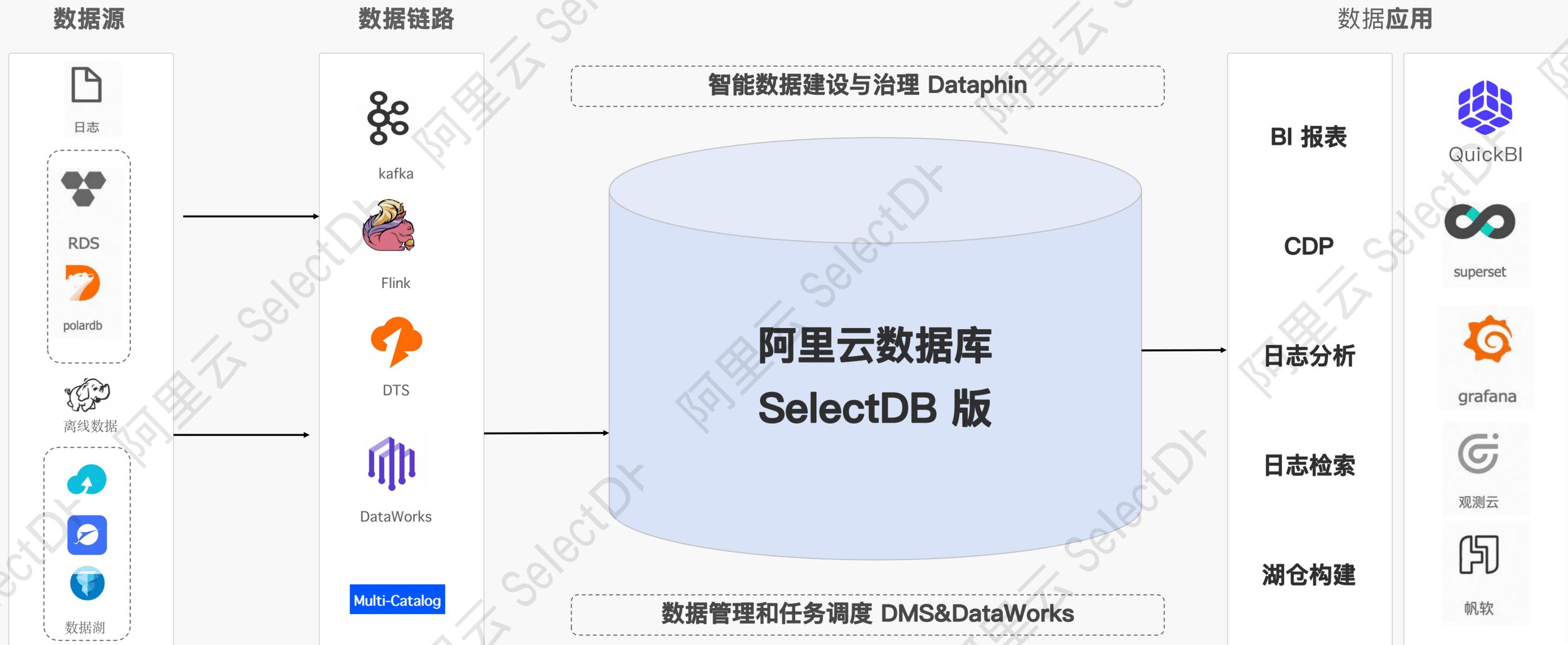
阿里云数据库 SelectDB 版：跨可用区高可用（邀测中）



- FE 跨可用区高可用部署
- BE 多虚拟计算组多读多写，跨可用区高可用
- 存储三可用区强一致高可用

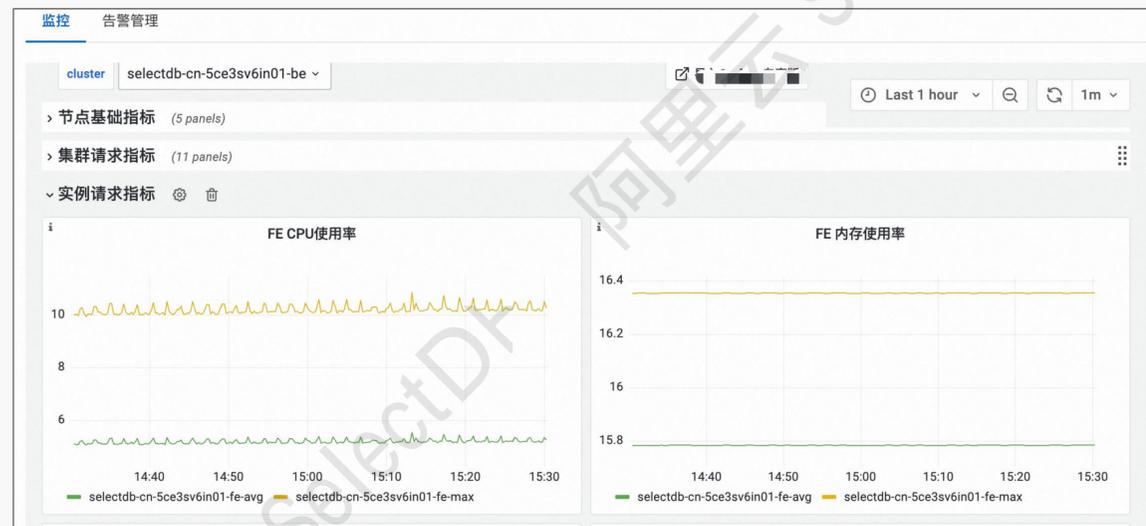


阿里云数据库 SelectDB 版：具备全数据链路生态支持



阿里云数据库 SelectDB 版：全面完善运维支持

资源和引擎全面监控告警



支持自定义参数保持自建集群平滑迁移

The parameter configuration interface shows a table of parameters for a cluster named 'new_cluster (selectdb-cn-5ce3sv6in01-be)'. The table has columns for '参数名' (Parameter Name), '参考值' (Reference Value), '类别' (Category), '当前值' (Current Value), '可修改参数值' (Modifiable Parameter Value), and '描述' (Description). A '提交参数' (Submit Parameters) button is visible at the top right.

参数名	参考值	类别	当前值	可修改参数值	描述
disable_storage_page_cache	false	Query	false	[true false]	
disable_storage_row_cache	true	Query	true	[true false]	
doris_remote_scanner_thread_pool_threa...	96	Query	96	[0-10000]	
doris_scanner_thread_pool_thread_num	48	Query	48	[0-10000]	
enable_query_memory_overcommit	true	Query	true	[true false]	
enable_stream_load_record	false	Load	false	[true false]	
enable_vertical_compaction	true	BackgroundTask	true	[true false]	
es_http_timeout_ms	5000	Query	5000	[0-100000]	
es_scroll_keepalive	5m	Query	5m	[0-14400]	

一键平滑版本升级

The version upgrade dialog box provides information about the upgrade process. It includes a warning: '小版本升级会导致实例闪断，部分读写请求异常；大版本升级需要停机切换，实例可能中断服务10分钟左右。请您尽量在业务低峰期升级，并确保您的应用具有自动重连机制。具体操作，详见升级内核版本' (Minor version upgrade will cause instance flash disconnection, some read/write requests are abnormal; major version upgrade requires downtime for switching, the instance may interrupt service for about 10 minutes. Please upgrade during low business activity and ensure your application has an automatic reconnection mechanism. For specific operations, see the upgrade kernel version). It shows the '当前版本' (Current Version) as 3.0.8 and the '目标版本' (Target Version) as 3.0.11. The '升级方式' (Upgrade Method) is set to '立即升级' (Upgrade Immediately). Buttons for '确定' (Confirm) and '取消' (Cancel) are at the bottom.

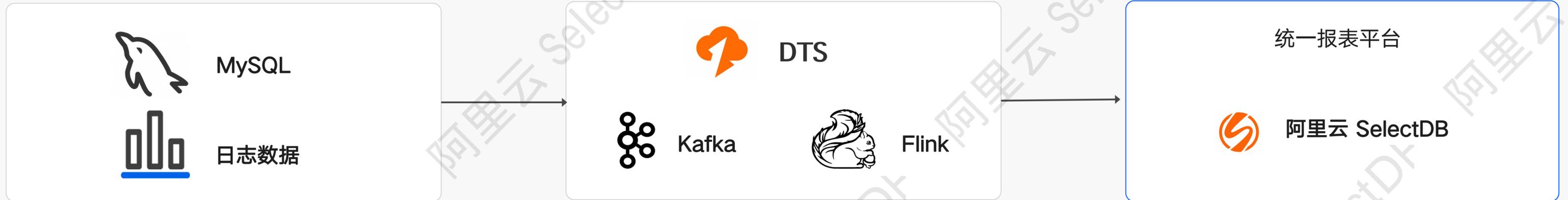
定时/手动：秒级集群平滑扩缩

The elastic scaling configuration interface shows options for '分时弹性伸缩' (Time-based Elastic Scaling) and '手动伸缩' (Manual Scaling). A note states: '至少2条规则且有不同目标计算资源，才可以实现应对业务高低峰的分时弹性伸缩。' (At least 2 rules with different target computing resources are required to achieve time-based elastic scaling to cope with business peaks and troughs). A table for '分时弹性策略' (Time-based Elastic Strategy) is shown with columns for '规则' (Rule), '执行周期' (Execution Cycle), '执行时间' (Execution Time), '目标计算资源' (Target Computing Resource), and '操作' (Action). One rule is configured: '规则 1' (Rule 1), '每天' (Daily), '02:00' (02:00), 'selectdb.xlarge(4核 32GB)', and '取消 保存' (Cancel Save). Below, the '手动伸缩' (Manual Scaling) section shows the current cluster name 'selectdb-cn-5ce3sv6in01-be (new_cluster)', current resource 'selectdb.2xlarge(8核 64GB)', and target resource 'selectdb.2xlarge(8核 64GB)'.

Part 3

阿里云 SelectDB 版场景和案例介绍

场景一：高并发实时报表分析&特征平台



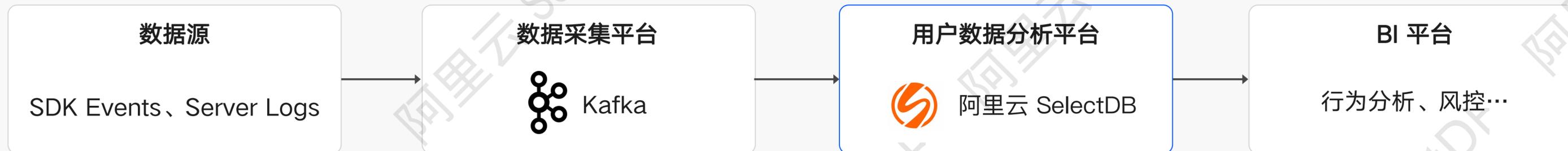
案例1：头部在线教育公司的统一报表分析平台

- 面向客户的报表平台，典型的 **高并发报表分析** (customer facing analytics)，对并发和响应延时要求高
- 采用 SelectDB 作为高并发报表分析解决方案，替代 **MySQL**、**Apache Doris**，快速解决客户线上并发痛点
- 查询：**500w+ 查询/天**，99分位查询延时 **100ms**，查询峰值 **QPS 500~**，压测峰值 **1000+**，导入：数据流 **秒级** 延时，持续追加写入

案例2：头部互联网金融公司的实时特征平台

- 面向 ToC 客户的实时特征服务，要求实时更新背景下，保持高并发、低延时的查询能力
- 采用 SelectDB 作为实时特征查询服务，查询 **QPS 4000~**，响应延时 **20ms**，支持实时更新

场景二：CDP 平台-用户行为和画像分析



头部互联网内容分发公司的用户数据分析平台

- 面向公司全平台的用户行为与画像分析平台，满足事件、留存、漏斗、圈人等用户行为分析需求
- 使用 SelectDB 作为实时解决方案，替换原有离线架构(Kafka + Hive + Spark + MySQL)，“颠覆式提升数据分析效率”
需求交付方式从 开发离线作业 转变为 BI 工具自助分析，交付周期从 7天 下降至 0天
十亿级用户的用户行为数据，数据延迟从 1小时 降低至 15s，查询耗时从 分钟级 降低至 秒级
助力客户一个季度对接公司内 20+ 条新业务，助力平台在全公司内快速扩张
- 目前已在海内外已扩展至 10~ 集群、1000+核，业务仍在快速增长中，已成为公司统一的用户数据分析平台

场景三：日志搜索分析

阿里云 SelectDB 提供相比 Elasticsearch 10 倍性价比的日志搜索分析方案

2 倍 查询性能提升

- SQL 查询语言，支持 Join，相比 DSL 简单，表达能力强
- 高效的执行引擎和优化器，高效的索引与文本分析

5 倍 写入吞吐提升

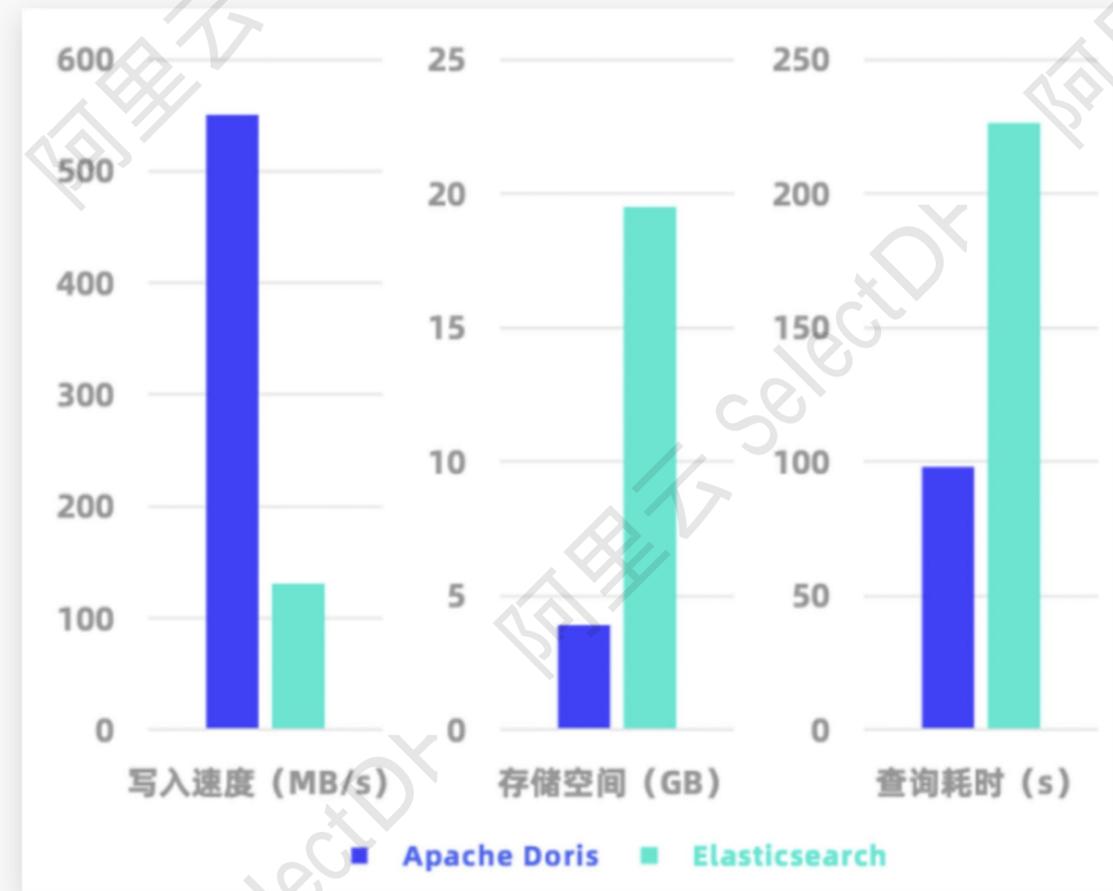
- 向量化指令，提升数据解析，索引构建的性能
- 简化去掉正排等索引结构，降低索引构建开销

80% 存储成本降低

- 简化去掉正排等索引结构，减少倒排数据量 30%
- 列式存储和 ZSTD 高效压缩，提供 5-10 倍压缩比
- 冷热分层，大幅降低冷数据成本

稳定性 大幅提升

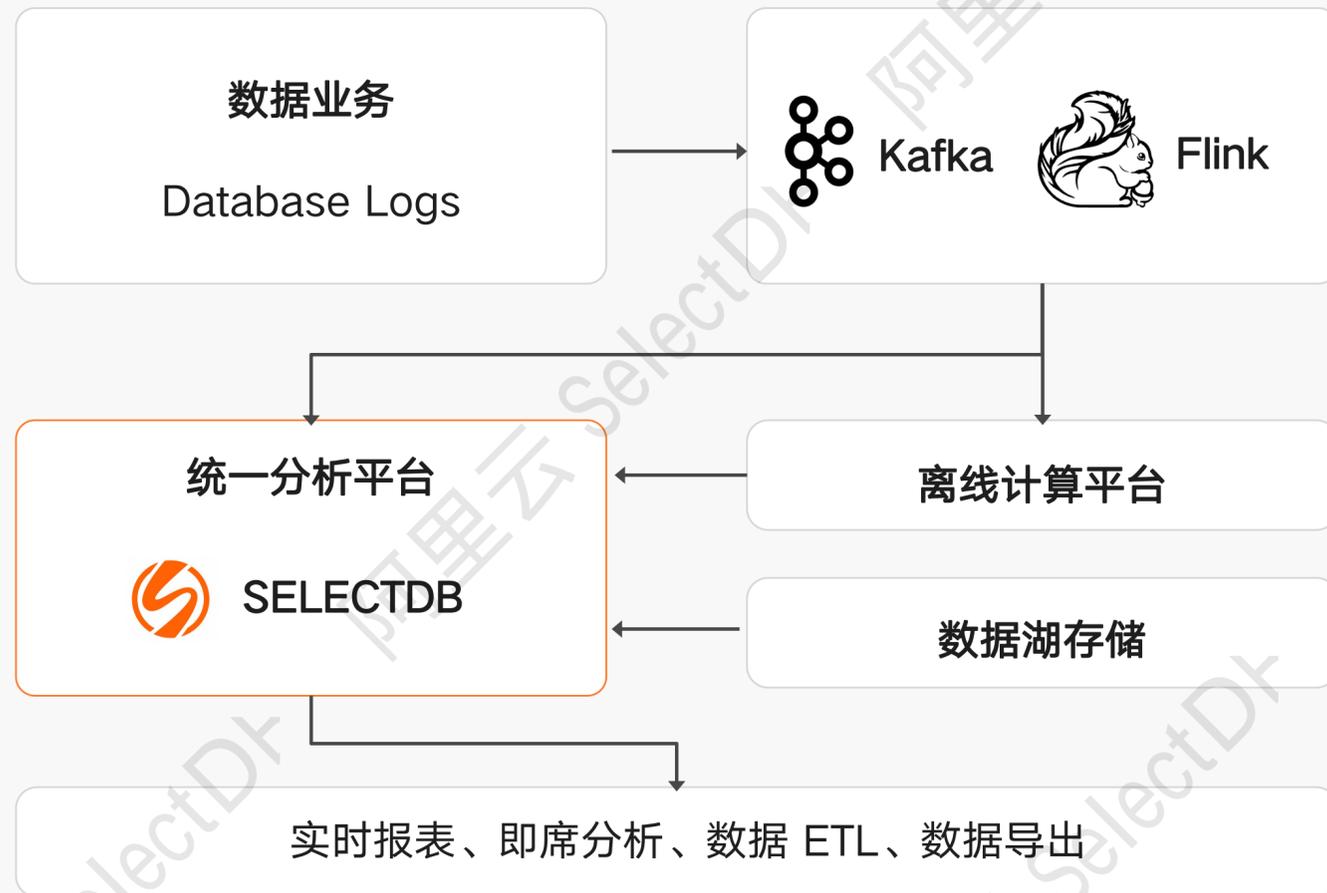
- 资源隔离和大查询限制
- 高效内存管理，避免 Java GC 影响



用户场景从日志搜索分析从 ElasticSearch 到 SelectDB 的效果

- 观测云：使用阿里云 SelectDB 作为日志检索分析方案，提供完整的可观测性产品，资源成本降低 70%

场景四：统一数仓分析平台



波司登零售公司的统一分析平台

- 公司统一的大数据分析平台，支撑公司生产、销售、库存、渠道等全链路数据分析需求，包含高并发实时报表、即席查询、数据ETL处理、数据导出等需求
 - 采用阿里云 [SelectDB](#) 作为统一分析平台，替代其原有组合式架构，大大降低架构复杂性，性价比提升 3倍+
- 实时报表：性能提升 3倍+，并发达 6000，响应延时 20ms
- 即席查询：性能提升 1倍+，响应延时稳定在 1秒 以内
- ETL 处理：性能提升 2倍+，相比Spark类传统系统提升 8倍+
- 数据导出：性能提升 2倍+，10GB数据全量导出耗时 30s级
- 混合负载：持续稳定运行，报表可达 4000+ qps，即席查询平均响应时间在 1秒 内
- 支持客户千亿级的统一分析平台，可轻松扩展满足客户未来 10 倍级的业务增长

扫码加入阿里云 SelectDB 产品交流群

